

**PROCEEDINGS OF THE 26th INTERNATIONAL
CONFERENCE ON ENGLISH TEACHING AND LEARNING
IN THE R.O.C.**

第二十六屆中華民國英語文教學研究國際研討會論文集

**Compiled by
Department of Foreign Languages and Literature
National Tsing Hua University
May 2009**

**國立清華大學外國語文學系彙編
中華民國九十八年五月**

Preface

I am delighted to welcome all of you to the 26th International Conference on English Teaching and Learning in the ROC and to the campus of National Tsing Hua University. Some of you have travelled just a few moments to be with us, others from various parts of the island, yet others from as far away as China, Japan, Australia, and North America. Whatever your investment, I hope you will find this conference rewarding.

As you can tell from the number in the title, this conference has a relatively long history; it was the first conference devoted to English Language Teaching held annually in Taiwan and is unique in being held at different universities each year. National Tsing Hua University was privileged to host the 8th as well as the 13th of these conferences. At the beginning, these were one-day local conferences held on a Saturday in May. Later, as interest as well as caliber of research was raised, international speakers were invited and the conference was spread over two days to accommodate the increasing number of excellent papers. With the co-sponsorship of the English Teaching and Research Association, under the leadership of Prof. Wuchang Vincent Chang, this conference has become the most prominent research-focused annual conference in ELT in Taiwan.

With the theme *Holistic Teaching and Learning: Crossing the Boundaries of ESL/EFL*, this conference asks what holistic teaching and learning are and explores ways of integrating the four skills into both ESL and ESL teaching and learning in traditional and more creative ways that cross previously assumed boundaries. The four plenaries, two workshops, two panels, colloquium, and over 70 papers address these and related issues in ways we hope will stimulate participants toward further research in these fields.

We would like to thank the English Teaching and Research Association, the National Science Council of Taiwan, the Research and Development Office of NTHU, the Humanities and Social Sciences Research Center of NTHU, as well as the many participating publishers for their support, including the Crane Publishing Company Ltd. for publishing the CD-ROM of the conference papers. Thanks also to the conference organizing committee, our FLLD office staff, and the many student helpers. Most importantly, a big thank you to the presenters for providing the content for the conference and contributing to the already high-quality research on ELT coming from Taiwan, and to the participants for providing a lively audience to discuss the ideas presented during these two days.

Johanna Katchen, Chair
Department of Foreign Languages and Literature
National Tsing Hua University
April 30, 2009

TABLE OF CONTENTS

Plenary Speakers

Burns, Anne, Becoming a Holistic Teacher: Discovering our Practices Through Action Research	1
Murray, Denise E., Transforming the Role of ICT in Pedagogy	12
Sasaki, Miyuki, Changes in Japanese Students' English Writing Ability, Strategy-use, and Motivation during Their 3.5-year University Life	17

Panel Discussions

張武昌, 周碩貴, 廖美玲, 葉錫南, 陳秋蘭, 游毓玲, 黃金誠, 我國英語教育政策與教學之影響	19
李玉華, 呂信, 施雅俐, 余綺芳, 如何幫助「英文(一)」課程低成就同學的學習	21

Colloquium

Chen, Howard Hao-jan, Developing Statistic-based and Rule-based Grammar Checkers for Chinese ESL Learners	31
Gao, Zhao-ming, Developing an English Tense and Aspect Checker Using Natural Language Processing Techniques	42
Liou, Hsien-chin, Chen, Yu-ru, Yu, Li-tang, and Chang, Jason S., Development of an Online Concordancer for Primary School Students	43

Workshops

O'Brien, Christopher J. and Chen, Mei-hsiang, Assessing the Learning Management System <i>Ecto</i> for EFL Professors	49
韋金龍, 資訊融入大專進階英文教學實務工作坊：以「網路英文」為例	52

Paper Presentations

Chang, Ching-fen, Lin, Yu-ting, Lee, Ming-chien, and Liu, Yi-wen, Beyond Classroom: Exploring Activity Theory to Informal English Speaking Sessions	62
Chang, Fang-chi, Structured Writing Activities for EFL Junior High School Students in Taiwan	73
張秀珍, 遠距中英口筆譯教學個案分析	83
Chang, Shu-i, Chang, Fang-chi, Shen, Tien-cheng, and Hsu, Hsiu-fen, Early English Writing Activities in EFL Grade-one Classes in Taiwan	93
Chang, Yu-ying and Liu, Wei-ling, Professors' Use of Questions in English Lectures: An Interdisciplinary Comparison	101

Chen, Aleck Shih-wei and Shih, Patti Yi-ping, Rime Familiarity Effect on Spoken Vocabulary Acquisition	113
Chen, Hsiu-chuan, Literature Discussions in an EFL Reading Class	117
Chen, Howard Hao-jan, Examining the Scoring Accuracy and Feedback Quality of an Automated Essay Evaluation Program	121
Chen, Jin-shan, Conversational Narrative Inquiry into Learning to Teach English	136
Chen, Lei-yi and Chen, Yu-mei, Psychosocial Sources of Anxiety in Academic Oral Presentations	142
Chen, Margaret, Vocabulary Games as a Memory Enhancement Device	154
Chen, Mei-chen and Cheng, Meng-fang, Effects of English Proficiency and Gender on Reading Strategy Use	169
Chen, Mei-chen and Huan, Sin-yi, Use of Vocabulary Strategies in Relation to Major and Gender	187
Chen, Mei-chen and Lee, Yen-hua, Teacher Questioning Strategies and Wait Time	203
Chen, Mei-chen and Wu, Chen-ni, Effects of Major and Gender on Language Learning Strategy Use	222
Chen, Mei-hua, English Collocation Competence of College Students and Learner Factors	239
Chen, Yuh-mei, A Comparison of Two Assessment Instruments of English Oral Performance	245
Cheng, Chun-en and Hu, Chieh-fang, Chinese EFL Learners' Context Spelling Knowledge: Position Constraints and Sequential Dependencies	258
Cheng, Wen-wen and Lee, Mei-ling, Students Perspectives toward the Management of a Self-Access Learning Center	271
Chin, Shao Joyce, Investigating the Summary Writing Performance of University Students in Taiwan	278
Chiu, Chia-hui, The Application of Summarization Technology in an EFL Classroom	295
Chiu, Chu-yao, The Effects of Interactive Reading-Writing Journal in EFL College Writing	307
Chou, Shiao-ting and You, Yu-ling, Effects of Reciprocal Teaching on Reading Comprehension in an EFL Classroom	318
Chou, Su-yu and Chen, Yuh-mei, A Study of Cloze Test Items in Scholastic Aptitude English Test and Department Required English Test	332
Chu, Po-ying, Holistic Language Learning through Genuine Learner Autonomy	343
Dai, Mei-hsia, Examining How Meaning, Form, and Context Interact in the Text of News Leads	351
董大暉, 新聞專業英語時體教學對新聞中譯英的影響	357

Hou, Hsiao-i and Li, Min-yu, A Needs Analysis of Culinary Arts Majors' EFL Learning in Taiwan's Context	366
謝佳玲, 余榕毓, 中英網路溝通對比分析及其教學意義	374
Hsin, Ai-li and Lin, Yi-ru, A Study on the Acquisition of English Expletives by EFL Learners in Taiwan	387
Hsu, Liang-fong, Kao, Shin-mei, and Carkin, Gary, A Holistic English Classroom through Drama: Applied Drama Techniques, Process Drama, and Readers Theatre	399
Hsu, Yu-chieh, A Study of English Listening Proficiency and English Oral Proficiency	406
Huang, Hui-ling and Chen, Wen-hui, Corpus-Based Remediation for Technology-Students' Vocabulary and Reading Comprehension	417
Huang, Hui-ling and Chen, Wen-hui, The Effects of Storytelling Approach on English Learning Motivation of Low Achievers in Junior High School	428
Huang, Wen-chi, English Preservice Teachers' Perspectives, Computer Self-Efficacy and Computer Anxiety between Cultures	444
Hung, Shu-mei, Beyond Communicative Competence: Using Instant Messenger in Intercultural Exchange	452
Kuo, Chia-chi and Brown, Charles, Teaching Strategies in Online Writing Instruction by Pre-service Teachers	464
Kuo, Chih-wei and Kuom Chih-hua, Titles of Research Articles: Syntactic Structures and Communicative Effectiveness	479
Lan, Mei-hua, A Comparison of Effects of BLOG-based and Paper-based Portfolios on Writing Attitudes and Performances	491
Lang, Hui-ling, Integrating Discourse-based Approach into English Writing Instruction	509
Lau, Ines, The Advantages of Task-Based Learning Pedagogy in Macao	517
Liao, Ming-chu and Wang, Hung-chun, Grammar Instruction and Corrective Feedback: Cross-regional Student Perception Differences	535
Lin, Chun-cheng and Lin, Chih-cheng, Learning Alone at a CALL Center: What Frequent Users Tell Us	546
Lin, Grace H. C., Ho, Max M. H., Chien, Paul S. C., Listening, Speaking and Reading for Writing: Holistic Pedagogies for an Advanced Writing Course	560
林和蕙, 王旭, 字典音標教國小學童英語發音：以協同行動研究為例	570
Lin, Ya-ling and Chang, Fang-chi, Picture-Book-Making in EFL Writing Class: Third-Graders' and Teacher's Perception	584
Liu, Angie H.C., Scoring Collaborative Oral Interaction in EFL Classrooms: What is Fair?	597

Liu, Yi-chun, Holistic Learning of Academic Writing: Impacts of Service Learning on Writing Identity and Knowledge Construction	608
Ma, Tung-mei and Li, Shu-fan, Bridging Test Construct and Beneficial Washback Effects: Revising the GEPT High-Intermediate Reading Test	612
Min, Hui-tzu, Peer Review Training and Teacher Feedback	625
Pai, Chia-wei and Liou, Hsien-chin, Comparing Synchronous and Asynchronous Computer-mediated Peer Review of EFL College Students	630
沈富源, 運用指示詞與議論文之文本結構教導申論題寫作：以研究所入學考試試題為例	644
Shieh, Wen-yuh, Perceptions of English Language Use at Tourist Hotels in Taiwan	655
Teng, Mei-hua and Hsu, Yueh-kuei, Effective Teaching in Large Multilevel Classes: Beliefs and Practices of Experienced EFL Elementary School Teachers	666
Tsai, Han-min, Examining Metacognitive Performance Between Skilled and Unskilled Writers in an Integrated EFL Writing Class	678
Tsai, Yea-ru, Wu, Chia-pei, and Hsu, Cheng, Facilitating EFL Reading Strategy Use Through a Course Management System	690
Tseng, Jun-jie, English Teachers' Conceptual Change in CALL: A Reflective Approach	702
Tseng, Wen-ta, Validating a Pictorial Vocabulary Size Test Through Rasch Modeling	707
Tso, Wei-fung, Advertisement, Critical Thinking and ESP Writing	718
Tsou, Wen-li, Effective English Courses for University Students	727
Tung, Hsing-cheng, The Historical Review of ETJCEE/DRET Grammar Tests	731
Wang, Chang-o, Senior High School Students' Perceptions of L2 Classroom Climate	759
Wang, Hung-chun, Passive Voice in Research Articles: Revisiting Rhetorical Functions	774
Wang, Liang-chun, Triggering Classroom Interaction with Teachers' Questions: A Taiwan Case Study	785
Wang, Pei-ling, An Integrated Freshmen English Instruction through Technology	809
Wei, Feng-fan and Chiu Tsuo-lin, An Analysis of Collocational Errors in Chinese Learner English Corpus	826
Witton-Davies, Giles, A Holistic Approach to Fluency: Seeing the Wood for the Trees	836
Wu, Hui-ju and Chu, Hsi-chin, Comic Strip as a Supplement of Reading: Effects of Presentation Sequence	850
Wu, Su-yueh, Multiple Instruction Strategies for Increasing Learners' English Vocabulary Retention	857

Yang, Ping-che, Liou, Hsin-chin, and Chang, Jason. S., WriteAhead: An Abstract Writing Assistant System for Academic Writing	864
Yang, Tzu-yin and Su, I-ru, Effects of Chinese Transliteration Systems on English Literacy Learning	871
Yu, Li-tang and Liou, Hsien-chin, Concordance-enhanced Vocabulary Instruction on EFL Elementary School Students	885

Becoming a Holistic Teacher: Discovering our Practices Through Action Research

Anne Burns

Macquarie University

anne.burns@mq.edu.au

How can we gain more holistic and rounded perspectives on what we do as teachers in the classroom? How can we continue to extend and develop our professionalism as English language teachers? How can we develop a 'life-long learning' attitude to our professional practices? Learning about teaching is not confined to initial teacher training. It extends across our whole teaching career and involves curiosity about the way we teach and how our students learn. It also involves developing understanding of the practices and policies that underpin the work we do in our classrooms.

Action research has become popular as a way for teachers to investigate their classrooms in order to understand their teaching practices at a deeper level. But what is action research and how can it be applied in our classrooms? What are the benefits for teachers of becoming investigators of their own classrooms? In this talk I will describe action research as an approach to doing classroom research as well as the processes and methods that can be used to carry it out. To illustrate my description, I will draw on the action research work of teachers in various countries, including Asian countries, to illustrate what teachers say and do about action research. My hope is that I will inspire teachers in the audience to consider doing action research as a way of seeing their classrooms and their teaching more holistically and of refreshing their professional lives.

INTRODUCTION

Learning about teaching is not confined to initial teacher training. It extends across our whole teaching careers and involves curiosity about the way we teach and how our students learn. It also involves developing understanding of the practices and policies that underpin the work we do in our classrooms.

Over the last two decades the demand for expert and competently trained English language teachers has increased dramatically. There are numerous factors contributing to this situation. Governments all over the world see the English language skills of their populations as vital if they are to communicate on the global world stage. English is used as the international language of trade and economy and being able to use English effectively increases a country's ability to compete and develop socially and economically. Therefore, there is growing demand worldwide for competent and knowledgeable English language teachers who can respond rapidly, flexibly and imaginatively to the changes that will continue into the future.

As a result the training and continuing professional development of teachers has become an area of considerable importance in the field of English language teaching. Two forms of change have motivated the way this field now conceives of second language teacher education (Burns & Richards, 2009). The first is the internal change within the profession itself as recent theoretical concepts and research on teacher education have made an impact on teacher education and classroom practices. Change internal to the profession has brought with it more contextualised, holistic and socioculturally based ways of conceptualizing how teachers form and develop across their careers. The second is the external change that has

accompanied demands by government educational policies for greater accountability, standards more centralized control over the training and assessment of teachers.

In this presentation my focus will be on the first form of change within our profession and particularly on the roles of self-reflective and holistic approaches to teachers' initial education and lifelong professional development. In particular I will highlight the concept of action research and offer some ideas that teachers can put into practice individually or in collaboration with their colleagues.

A recent quote from a teacher researcher and educator (Santana-Williamson, 2001, p.35) sets the scene for my discussion,

Nowadays... teaching is seen more as a holistic activity in which teachers constantly try to discover things that work, discarding old practices and taking on board new ones though a process of decision making, reflection, analysis and assessment.... They should conduct action research, reflect on their teaching and on their students' learning, and use their knowledge and experience

In recent years, action research has been promoted as a way for teachers to become involved in investigating what happens in classrooms and reflecting on their practice. However, teachers are not always familiar with this type of research and the way it can be used in their own professional development. During the many workshops I have given in different countries in South-East Asia, the participants have raised numerous questions about action research. They are questions that I am asked frequently and so in this talk, I will highlight them and offer some ideas and responses. Let's begin by thinking about what happens when a teacher does action research. I will start by looking at year 3 classroom in a public high school high in Indonesia. The example is drawn from the work of my former PhD student, Dewi Rochsantiningsih, and I am grateful to her for permission to use this example (Rochsantiningsih, 2005).

ACTION RESEARCH: AN EXAMPLE

The high school where the teacher taught was located in an area where the students were not high achievers and there was a low student attendance. According to the teacher, whom I will call Miss Yasmin, the students showed a slow process of teaching and learning and low motivation. She also said there were discipline problems, student laziness and classroom management. The teacher was a dedicated and committed teacher and was concerned about these problems. She was looking for ways to change the situation for the students. In her words:

...there are so many things I would like to solve and improve. But I feel powerless.

The teacher found out that an action research group was about to start working with Dewi as the facilitator. She volunteered to join the group and to work with other teachers. Her aim was to investigate new teaching strategies to help her overcome the problems of poor attendance, what she saw as laziness on the part of the students and their passive attitudes. There were five students who were often absent from class and the teacher was particularly concerned about these class members. She decided to talk openly in her classroom with her students about these classroom challenges. She also invited them to share their ideas about teaching and learning in her class. This is what she stated:

Negotiating the lessons with the students means involving them as members of a team. And this means giving them trust and responsibility as well.

The teacher then distributed questionnaires that she had prepared to the students which asked them to state what kinds of activities they wanted to do in class. She found out that they wanted more variation in teaching materials and for the materials to be presented in a more interesting way. They also wanted to vary the location of the class. She told the students the result of the questionnaires and they agreed:

- to hold the class in different location (under the teak tree behind the classroom, in the library, the hall)
- to keep to class rules for getting to the location on time and not disturbing other classes
- that the teacher should choose the materials but vary her approach so that the students were not “being told” only.

The teacher was immediately encouraged at the students’ reactions – some of them slowly began to voice their opinions and seeing her encouragement, others joined in. The students began to look excited even though some of them kept silent. The teacher said the “good atmosphere in my class was inspiring”. In the next lesson, they worked in the shadow of the trees in groups. Students discussed vocabulary and answered questions about the topic they were discussing and then did written exercises based on their discussion. The teacher noticed that some students were excited and eager to discuss, but others were distracted by the new location and did not participate very actively. In order to address this problem, in the next lesson the teacher reminded the students about the rules they had agreed and asked them to think about these rules when they met in their new location of the library.

When she arrived at the library all the students were ready, waiting for the lesson. She had decided to vary the material and activity by teaching the students a song. They worked in groups to read the words of the song and then to listen to the song on a cassette and identify some missing words. The students and the teacher finished the lesson by singing the song using all the completed lyrics. The teacher observed that the students seemed to keep their own rules much more and the students who had been inattentive last lesson were much more involved. Also all the students attended the class, even the five students who were frequently absent. She found the atmosphere in the class and the dynamics amongst the students to be very positive and cooperative.

Miss Yasmin came to several conclusions about her research:

Negotiating the teaching-learning process with the students improved the unfavourable class situation...Since the negotiation provided me with data of student preferences, it helped me to become more relaxed and focused to prepare teaching. I could also expand my roles as a teacher. Now I could also place myself as a friend in teaching: not only as the one who always orders and instructs. Positioning the students as “humans” and not as “objects” made my work lighter and easier.

I have outlined this teacher’s research to illustrate some of the processes of action research and to show how a teacher can use it productively to reflect on his or her own practice. To expand this discussion I will now raise several questions that come up frequently about doing action research.

WHAT IS ACTION RESEARCH?

Action research is one of a group of activities associated with the idea of reflective teaching. Advocates of reflective teaching argue that the voices of teachers have been absent too long from the research literature. They argue that teachers should be recognised as thinking professionals who can ‘both pose and solve problems related to their educational practice’ (Zeichner & Liston, 1996:4).

One key feature of action research is that the research *done by the participants* themselves in a particular social situation, rather than by outside researchers. This is a noticeable departure from other types of research and it highlights a central difference in the purpose of the research. The main point of action research is to find out more about what goes on in your *own local context*, which for teachers is the school or classroom where they teach, in order to *change or improve current practice*. In other words, action research is to do with deliberately reflecting on and *intervening in* your current practice. This is different from other forms of research, which may aim to hypothesise, describe, analyse, interpret or generalise - but not to change the research situation.

In action research, problem posing is achieved through a *reflective research cycle* of planning, action, observation and reflection, where the researchers:

- Plan: Identify a focus area of their practice that presents a ‘puzzle’, problem or question
- Act: Decide what they will do about this situation
- Observe: Collect information systematically about what happens when the plan is put into action
- Reflect: Analyse what the data they have collected are telling them about their practice
- Plan a new course of action: Act as necessary to change or improve the practice

(See Kemmis & McTaggart, 1988)

At this point the participants can enter the cycle again by trying out the changes they have planned to see what happens. In the example of Miss Yasmin, the teacher went through several cycles where she learned more each time about what to do next in her class. Her plans for new actions were based on the information she had collected in the previous cycle. The questionnaires gave her new information about the students’ preferences; her observations of the lesson under the trees identified the need for further negotiation and discussion and the third cycle showed positive results that would help her to plan more activities. In broad terms, then, action research involves investigating, analysing and changing our practices by collecting information systematically in a specific social situation.

WHAT KINDS OF RESEARCH QUESTIONS SHOULD YOU ASK? MUST YOU HAVE A HYPOTHESIS?

The kinds of questions asked in action research will be different from those in the scientific approach to research. Action research does not set out to prove or disprove hypotheses, but to broaden and deepen our insights into a challenging issue in our own situation. People who conduct action research may well develop ‘hunches’ or ‘working hypotheses’ about the way they or their students behave, and as they collect data they may change or adapt these hypotheses. But these are not the same kinds of hypotheses that are set up in scientific research where teaching or learning variables are carefully controlled.

Frequently, in action research the question you start out with is not the question you still need to ask at the end of the process. Sometimes you do not have a specific question in mind when you begin your action research; the questions emerge as you observe and reflect on

your classrooms over a period of time. But whether or not you have a specific question when you begin the research, it is useful to have some kind of focus. There are some useful ways to develop ideas about the kinds of questions you may want to pose, especially if you can work with other colleagues and share your ideas:

- Keep a diary or brief notes of teaching, learning or administrative activities in your workplace over a chosen period of time (e.g. a week, month). Read over the diary at the end of this time and identify some of your key thoughts, ideas or concerns.
- Brainstorm some starter statements:
I don't know enough about how my students...
My students don't like... Why is this?
I'd like to find out more about what my students do when they...?
- Make a list of questions about things in your workplace that have puzzled you for some time. Ask other teachers for their 'favourite puzzles'. (How do they compare with your own?)
- Observe (preferably over a period of time) a typical situation in your teaching context. What stands out for you from your observations? What research questions or issues do they suggest?
- Find a favourite article in a journal or teacher newsletter. Think about how the issues it presents can be related to your classroom. What questions or issues does the writer address? What questions or issues does the writer leave out that you would like to know more about?

At a recent workshop I ran in Thailand, these were some of the questions that teachers suggested:

- What kinds of groups will work best in my classroom?
- What strategies or activities can I develop to encourage learners to speak more?
- What are some effective ways of teaching grammar?
- What kinds of communicative activities work best in my classroom?

A colleague who worked with me in a project some years ago, Pamela McPherson, provides a very good example of how the focus of action research questions might change. Pam worked with learners who were enrolled in adult ESL classes in Australia:

My group was diverse in all the ways that make adult immigrant classes so interesting to teach. Ages ranged from 22-58 with equal number of males and females. They came from 15 different countries and spoke 17 different languages. Most had come to Australia because their country of origin was now unsafe for them... My concern was with the wide variation in the levels of spoken and written English...I was uncertain how to manage the class and I felt my planning was very 'hit and miss'...I decided to read the literature on managing mixed-ability groups and to talk to teachers in [my centre] and in community organisations and primary school education about strategies they used...

As a result I decided to focus on developing materials and activities at different levels and to observe the response of the learners to these materials. I documented these observations [using a journal and drawing up diagrams of classroom interaction] and began to realise how much I tended to 'control' their learning by dispersing materials at 'appropriate' levels. When I allowed the students to take control, they worked with the [materials] in different ways which they found personally effective...

However, at this point I became concerned about another aspect of the class. I observed that the students would not cooperate to undertake joint activities. They were also starting to express exasperation, boredom, irritation and once, near hostility, as I brought to the classroom lessons and activities [about personal experiences] I thought were interesting and relevant, but which they were not prepared to participate in... I decided on a strategy of individual consultation. I spoke to each student about what they were learning, how they learning and how they could develop their skills. I documented their comments and followed with activities designed to enhance their requested learning areas. I also documented comments on their reactions to my classroom activities...

I began to see emerging patterns and to uncover the reasons for the rejected activities. Student comments and reactions indicated that discussions that revolved around cultural or social difference were not acceptable... On a class excursion, I learned that the students were aware of deep ethnic, religious and political differences because of their experiences of the part of the world they had just left [former Yugoslavia]...I suddenly realised how difficult it had been for them to maintain the veneer of courtesy and civility when I was introducing activities which demanded that they expose and discuss the differences they were attempting to ignore!
(Summarised from McPherson, 1997)

ISN'T THE DATA COLLECTED IN ACTION RESEARCH TOO SUBJECTIVE?

This question is a frequent criticism, but it misses a major point about action research. The answer to the question is really both *yes* and *no*. Action research data *is* subjective in the sense that it comes from real and personal perceptions and situations in the researcher's workplace. Consequently, the personal views and actions of the researcher and other participants are the part of the focus of the data collection process. There is no attempt to control the classroom or the school as it actually exists. This is an accepted aspect of qualitative or interpretive approaches to research.

However, in another sense action research is *not* subjective. When doing action research we are focusing on collecting data in a systematic way and using data to try to identify what themes or trends emerge that will give us new insights. These insights may be in contrast with what we just assume to be the case, when we have no data to rely on.

A technique that is used to decrease the subjectivity of this kind of data is *triangulation*. This means getting data from more than one source. The researcher then cross-checks these different sources of data to see whether they are pointing to the same themes, interpretations and conclusions. For example, Ms Yasmin collected data from her own journal, the discussions she had with her students, and her classroom observations. These sources enabled her to build up a more complete picture that would not have been available if she had just relied on one set of data. Another way of guarding against subjectivity is to check our interpretations of our data with other colleagues or with the participants in our research to see whether they would reach the same conclusions.

WHAT ARE SOME EFFECTIVE WAYS TO COLLECT DATA FOR ACTION RESEARCH?

Generally, in action research we use the types of data collection methods that are common in qualitative approaches to research. This does not mean that our data cannot be analysed quantitatively, however. Test scores, percentages and totals are all ways in which we might show the findings from our data numerically. I have found it convenient in the workshops and discussions I have had with teachers to classify the most common methods into two types: *observational* methods and *non-observational* methods.

Observational methods, as the term suggest, involves observing and recording what is going on. This is a central method in ethnographic and qualitative research. Teachers I have worked with have collected data on their observations by:

- Making factual notes of the events they observe by recounting exactly what happened
- Writing diaries about the events plus their own evaluations or comments
- Video or audio-recording activities in the classroom
- Making transcriptions of parts of these recordings
- Drawing maps of the classroom that show how the learners interact (photographs can also be useful)

Non-observational data, on the other hand, tries to get ‘below the surface’ of the observations to find out what people believe, think, value or are able to produce. These techniques are commonly used:

- Developing questionnaires or surveys that ask people’s views through closed, ranked or open questions
- Conducting interviews or discussions with key participants, such as learners, other teachers, parents
- Collecting students’ written texts to find out what they can do before or after certain types of instruction

(See Burns, 1999 for more extensive descriptions of these methods)

Many techniques that are used in communicative language teaching can ‘double up’ as research methods. It would be fairly straightforward, for instance, to develop a survey that learners can use as an activity in class that also relates to the research questions you are interested in. For example, if you are interested in researching group work, you could ask the students to use a survey to collect information from each other about what they like or dislike about working in groups. Similarly, asking learners to keep journals on topics related to your research issues would provide additional data that could also be a language learning activity. One teacher I worked with developed a very neat way of finding out about her learners’ experiences and strategies for learning English in her class. She asked them to write letters at the end of their course to the students she would be teaching the following year. The letters focused on topics like strategies they had developed for learning English outside the classroom, what they should do to learn effectively in her class, and what they had learned about her approach to teaching English. This provided her with data for her research, as well as a writing exercise for her current students and a set of reading texts for her to use with her future students.

HOW SHOULD THE DATA BE ANALYSED?

Data analysis involves describing the *what* and the *why* of research. In action research, as in qualitative research in general, you are likely to find yourself initially analysing the data as you go along and responding to what it seems to be telling you, as Ms Yasmin did in her research. This is because the research is very much located in your own context of practice and part of what you find is continually being evaluated against your previous impressions or assumptions. As you move through the cycle(s) of action research, you will be evaluating what you do, forming hypotheses or questioning what you find.

Although the data analysis process is fluid and ongoing while you are still in the research cycle, there comes a time when you want to develop some overall findings and interpretations. The typical steps for undertaking this *summative* analysis are:

- Assembling all the data
- Reducing and coding the data
- Comparing and contrasting the various sources of data
- Building meanings and interpretations of the data

Depending on the types of methods you use, you might reduce the data into a form that can be displayed by:

- Providing tables of numerical data, for example, responses to survey questions
- Developing categories of different types of behaviour you have observed our students using
- Identifying key words or themes that are repeated throughout the data
- Clarifying patterns of classroom language that are typical of your own or your students' interactions
- Identifying key extracts or quotes that highlight particular kinds of behaviour or viewpoints.

To provide a practical example of data analysis let us look at the research of Catherine Kebir (1994:30), an Australian teacher. Catherine wanted to identify the interaction strategies that her students used when they were undertaking communicative tasks in the classroom in order to understand the range of strategies and their effectiveness. After recording the students' interactions as they completed the tasks, she transcribed their talk and then analysed the different types of communication strategies that were being used:

To analyze the data I underlined instances of communication strategies in the transcript, putting a number in the margin to signify which type. To help with classification, I used different colored pens to indicate for example, word coinage or circumlocution. This left me with a detailed list of strategies in their differing contexts.

By identifying the strategies used, Catherine was able to discuss them with her learners and also introduce them to new strategies that were not in their current repertoire.

HOW CAN I GO ABOUT TELLING OTHERS ABOUT MY RESEARCH?

Action research is typically carried out by teachers, and written accounts that show how teachers are investigating challenges in teaching and learning are usually welcomed by other teachers, especially those who may be working with similar classes. These accounts can give a sense of how to address common teaching problems and situations and they share expertise

and experiences.

However, the way the accounts are written will obviously depend on the scope and purpose of the research. Teachers enrolled in formal courses, such as bachelor or master's programs with an action research component, will usually be required to produce genres of writing that follow academic conventions: typically, the rationale and context for the research, literature review, methodology, data analysis and findings and interpretations and conclusions. For less formal projects, teachers can experiment with narrative styles that 'tell the story of the research' with a teacher audience in mind. This style of reporting through formats where the processes of the research and the personal voice of the teacher are heard are usually more appealing to other teachers. When I worked some years ago with a group of Australian teachers we developed the following outline for reporting the research:

1. Title and author's name
2. The setting for the research – background, context, class, type of students
3. The reasons for the research and the questions or focus
4. The steps taken in the research, the teaching-learning techniques used and the data collected
5. The findings or realisations that came out of these steps and the interpretations made
6. The author's response to the findings and the research
7. Any further details of interest to the reader – readings, teaching material or techniques (these could be included in section 4 or in an appendix at the end).

(See Burns, 1999 for further details)

Over the last decade, several publications containing accounts of teachers' action research have begun to appear (e.g. Burns & Hood, 1995; Burns & Joyce, 2001; Edge, 2001). They provide useful models for writing up action research projects. Of course, written accounts are not the only ways to let others know about your action research. Others possibilities include:

- Posters
- Short synopses or summaries
- Individual or group presentations
- Video or photographic displays

Presenting your research to other teachers is a very valuable professional development activity, however you decide to do it. By writing or talking about the research you get to understand it more yourself and this becomes a reflective aspect of the research process. You also enable other teachers to share and learn about action research and you can also get valuable feedback from them. Finally, telling others gives a sense of reaching a finishing point in the research process.

WHAT IS THE STATUS OF ACTION RESEARCH IN THE TESOL FIELD?

Action research has quite a long and well recognised history going back over the last 60 years at least (see Burns, 2005) and it is used in many contexts (for example business and health) apart from education. However, it has only become better known in TESOL and applied linguistics over the last 15 or so years. Broadly speaking, the more traditional and well established types of experimental research, such as quantitative methods involving a hypothesis, and an experimental and control group, have long been considered to be more acceptable forms of research. In addition, academic researchers and teachers in the applied linguistics and TESOL fields are often trained in the more scientific methods. As a result, many of the teacher researchers I have met in my workshops, particularly those based in universities, have often felt nervous about conducting action research.

While this is understandable, it is also true to say that there has been a distinct shift

recently towards more qualitative and practitioner research in TESOL research (for example, *TESOL Quarterly* introduced guidelines for qualitative research papers in 1995). Researchers who adopt these approaches are interested in exploring the nature of the world in which people operate – and in doing research *with* rather than *on* other people. They see the perspectives of the participants in the research process and context (the insider, or *emic* perspective, as it is sometimes called) as a central part of the research. The approaches, methods and processes associated with this kind of research – for example, ethnography, case study, (auto)biography – are now considered a viable alternative to scientific methods.

Action research is associated most strongly with a qualitative approach to research. It is regarded as particularly useful in situations where teachers work collaboratively with colleagues or with academic researchers to introduce changes in curriculum, materials, classroom interactions, school practices and so on. This way of researching - by investigating realistic and relevant issues in people's workplaces - is often seen as having a more lasting impact than presenting teachers with ideas or recommendations imported from elsewhere and which seem unrealistic to them. Although action research is still gaining recognition in applied linguistics and TESOL as a systematic research approach, it is part of the trend towards more socially contextualised and holistic ways of doing research and is likely to become more widely used in future.

CONCLUSION

In this talk I've tried to provide some brief responses to the questions I have been asked during many professional development presentations. They are questions about action research that seem to come up quite frequently. It is useful to remember that there is no 'one size fits all' approach. Action research is very much a process approach to doing research and it is highly likely to be modified as you go along, according to the problems and issues in your own situation, the time you have to do the research and the types of data you are able to collect. It can be adapted to suit the time period a teacher has available for conducting research. Trying out and experimenting with an action research approach could extend over a period of a few lessons or a few weeks. A more substantial project, especially for a formal course might last for several cycles and continue for much longer periods. There is no "hard and fast" rule about the length of action research.

Action research is a stimulating way for teachers to enter the world of research, to explore their teaching and to develop their careers as lifelong learners about teaching. Perhaps the best encouragement, I can offer for those who want to get started in action comes from two teachers I worked with recently:

My experience of action research is that it is difficult to grasp or explain the concept until one is in the process of doing it. It is in the doing that it starts to make sense and become clear. (Jane Hamilton, cited in Burns, 1999: 20)

Collaborative action research is a powerful form of staff development because it is practice to theory rather than theory to practice. Teachers are encouraged to reach their own solutions and conclusions and this is far more attractive and has more impact than being presented with ideals which cannot be attained. (Linda Ross, cited in Burns, 1999:7)

REFERENCES

- Burns, A. (1999). *Collaborative action research for English language teachers*. Cambridge: Cambridge University Press.
- Burns, A. (2005). Action research. In Hinkel, E. (Ed.). *Handbook of research on second*

- language teaching and research*. Mahwah, NJ: Lawrence Erlbaum.
- Burns, A. & Hood, S. (1995). *Teachers' voices 1: Exploring course design in a changing curriculum*. Sydney: National Centre for English Language Teaching and Research.
- Burns, A & de Silva Joyce, H. (2001). *Teachers' voices 7: Teaching vocabulary*.
- Edge, J. (2001). *Action research: Case studies in TESOL practice*. Alexandria, VA: TESOL
- Kebir, C. (1994). An action research look at the communication strategies of adult learners. *TESOL Journal*, 4, 1: 28-31.
- Kemmis, S. & McTaggart, R. (1988). *The action research planner*. Geelong: Deakin University Press.
- McPherson, P. (1997). Action research: Exploring learner diversity. *Prospect: A Journal of Australian TESOL*, 12, 1, 50-62.
- Santana-Williamson, E. (2001). Early reflections: Journaling a way into teaching. In Edge, J. (Ed.), *Action research* (pp.33-44). Alexandria, VA: TESOL.
- Zeichner, K.M. & Liston, D.P. (1996). *Reflective teaching: An introduction*. Mahwah, NJ: Lawrence Erlbaum.

Transforming the Role of ICT in Pedagogy

Denise E. Murray

San Jose State University

denise.murray@mq.edu.au

Information and communication technology (ICT) has been used in language classrooms for more than two decades. Over this time, classroom use has moved from drill, text manipulation and word processing to more interactive and communicative applications such as email, chat, and web-based programs. However, the full potential of ICT is not often realized in the language classroom, for a variety of reasons, some related to the technology itself, others to teacher adoption of innovation. Research and examples of best practice provide us with a framework for the future direction of ICT in language pedagogy.

EARLY USES OF ICT

Early researchers noted the “mismatch between the capabilities of the standard educational computer system and the main thrust of language pedagogy” (Kenning & Kenning, 1990, p. 16). Because of this mismatch, the use of ICT was considered peripheral to instruction, resulting in certain forms of practice, considered wasteful of class time, being relegated to computer environments outside the classroom or as a source of supplementary or fun activities for language learners. Much early computer-assisted language learning (CALL) software was for drill and practice of grammatical features of English and some was game-based. The former exploited the computer’s ability to provide endless (and often meaningless) repetition without becoming bored—unlike a human teacher. As Salaberry notes, “the majority of CALL applications have not relied on any type of theoretical rationale except for the theoretical framework already built into basic CALL exercises: the computer is the ideal environment for the implementation of the type of Programmed Learning from the behavioristic tradition” (Salaberry, 1996, p. 8). Game-based CALL, while more engaging for the learner, was used by teachers as a break from the more central classroom instruction. As well as CALL software, teachers and learners also used computers as word processors.

None of these early uses engaged learners in authentic language use or encouraged learner autonomy, even though language pedagogy at the time was based on interactionist second language acquisition theories, which focussed on communicative language teaching (CLT) and learner-centeredness (Nunan, 1988). These trends led to a call for a “theory of CALL (computer-assisted language learning)”, as though the introduction of computers is of itself transforming. All technological adoption passes through phases. “Initially, they shape themselves to the contours of custom; ultimately, they follow paths selected through struggles among groups seeking to turn technologies to their own interests” (DiMaggio, Hargattai, Neuman, & Robinson, 2001, p.327). For language educators, that struggle has been over how to drive ICT for pedagogical purposes for “it is not so much the computer but the kinds of tasks and activities that learners do on the computer that can make the difference” (Hoven, 1999, p. 149). What kinds of tasks and activities can transform the role of ICT from patient drill master to a site for authentic language use and support learner autonomy?

A number of new applications of computer technology have transformed its use in language education—most notably communicative applications such as email and chat, and the expansion of the world wide web. However, while these applications have the potential for authentic language use and the development of learner autonomy, this potential is not

always realized.

The following sections draw on data from research I have conducted over the past decade in educational and business settings (Murray, 1988; Murray, 2004; Murray, 2004; Murray, 2005); teacher education courses I have taught using ICT (Murray, 2008); online language training programs I've consulted on; and a number of research studies my colleagues and I have been engaged in over the past five years (McPherson & Murray, 2003; Murray, 2003; Murray, Lloyd, & McPherson, 2006; Murray & McPherson, 2002; Murray & McPherson, 2004; Murray & McPherson, 2005).

CURRENT PRACTICE

I have found that, as teachers have experimented with ICT, they have helped learners engage in activities that enhance learner autonomy and creativity through authentic communication & collaboration. New technologies with an emphasis on communication and self-expression facilitate such innovative approaches. What has been key to the successful implementation of ICT, has been teachers' focus on learner needs as they develop tasks and activities to help learners achieve their curriculum goals. Teachers have used ICT as both tool and tutor, a distinction first made by (Taylor, 1980). The computer as tutor is where the computer is a temporary teacher (for example in some CALL programs), compared with the computer as tool, where the computer provides no such instruction (for example, email or a wordprocessor).

Computer as tool. The three main uses of the computer as tool are tools that provide information to learners, tools that help them communicate and tools that help them organize their work.

The primary tool for information is the world wide web. Teachers in our studies found that learners need to both find their way around the web and select, read, interpret and use what they find. Navigating the web in English (Murray & McPherson, 2005) is not transparent to learners, and even those who have experience in their first language, may not use effective navigational strategies. Many have poor searching skills and find it difficult to decide which of the many results from a Google search might have the information they need. Once they reach a site, they need to be able to read webpages. But, research has shown that webpages differ from other texts, even though they contain text types that also appear in print (Bauman, 1999; Lipscomb, 2002). Furthermore, reading online differs from reading print texts (Thurstun, 2004; Tindale, 2005). The teachers in our studies understood that their learners needed to master cyberliteracy in English and visual literacy, as well as other print aspects of literacy traditionally taught in the language classroom. They therefore explicitly taught the developing conventions of webtexts and carefully scaffolded instruction so that learners acquired the strategies and readings skills needed to successfully search for, select, read and use information for some other non-communicative purpose.

The primary tools for communication are email, chat, blogs, and discussion lists. This group of tools is referred to as CMC (computer-mediated communication). CMC provides a site for authentic language use as learners can interact with each other and with native speakers across time and space boundaries. Research has shown that CMC facilitates collaboration (Debski, 2006; Warschauer & Kern, 2000), and leads to pragmatic competence (Belz & Kinginger, 2002). However, the language modelled in these tools may not be the language learners need in other contexts. Much CMC contains features of spoken language and a simplified register (Murray, 2004). Teachers in our studies used email and discussion lists extensively with their learners. These teachers found they had to teach many of the skills involved in email, as their learners were not familiar with it, even in their first language. Whether learners are familiar or not, they need to be explicitly taught the emerging conventions so that they know with which audiences it is appropriate to use abbreviations

such as LOL (laugh out loud) or emoticons or “flaming”, the use of often abusive language.

The primary tools for organizing work are a wordprocessor, a spreadsheet and a presentation tool. Teachers in our studies used wordprocessing extensively and also encouraged and taught learners to use powerpoint (Murray et al., 2006).

Computer as tutor. When the computer is used as tutor, it takes on the task of instruction, acting in the role of teacher. Many CALL programs tutor learners. However, our research has shown the characteristics of best practice, that is, that closely mirror effective instruction by human teachers. In particular, we have found best practice involves:

- Models
- Explicit instruction
- Feedback

Models need to be provided so that learners know the target. However, as already discussed for the classroom, learners need explicit instruction in the language features of the texts they are learning. Feedback, however, remains a challenge in ICT. While some educators are experimenting with natural language processing (Heift & Schulze, 2007), NLP is still in its infancy and effective only in small domains. Currently, we therefore need to focus on how to provide feedback that does not involve artificial intelligence programming. We have found that, while the computer cannot provide interaction for the learner, a carefully designed CALL program can provide interactivity. Interactivity “can [also] be taken to mean the capacity the package provides for the learners to interact with, interpret, negotiate, and make meaning from the texts available, whether these are orthographic, audio, audiovisual, or visual texts” (Hoven, 1997 p. 11). For feedback to be considered interactive, it needs to be timely, specific, and multimodal. It must be provided to learners immediately they have complete the activity. Although the program cannot know exactly what the learner did, the feedback can be specific by explaining not only why a response is not correct, but also why the correct response is the best answer. Where possible, the response should be written, but also include a visual, aural, or kinesthetic element.

While our studies have shown language teachers transforming the use of ICT in instruction, there remain considerable barriers to its extensive adoption.

TECHNOLOGICAL BARRIERS TO ADOPTION

Few teachers have immediate access to a classroom of computers. Currently, most teachers have to reserve a computer classroom for learners to use. So, they may find themselves wanting to illustrate an instructional point by a visit to the web, but have to wait several days until their turn in the computer classroom. Teachers in our studies also commented on the failures of the technology and the lack of bandwidth or latest model computers (Murray, 2003). Very quickly the hardware and software becomes obsolete and yet schools cannot afford to constantly buy new hardware and software. Until the computer is as readily available as a textbook, teachers will find it a challenge to integrate ICT seamlessly into their instruction.

Current technology is, however, on the brink of bringing together a critical group of features that will help educators integrate ICT in their curricula. These features include broadband, convergence, wireless, and miniaturization. Broadband is essential for tools such as chat or video. Convergence refers to the integration of audio, video, and data communication into a single source that can be received on the same device via the same connection. All four are now happening with cell phones; however, for many users, the fully functional, integrated cell phone is too small and too expensive to function as a classroom interactive “book.”

TEACHER ADOPTION OF ICT

Early implementation of ICT was largely in the hands of the innovators and early adopters (Rogers, 1995), that is, the small percentage (around 16%) who take up an innovation because they are captivated by it and its potential. Such computer enthusiasts, however, often took the computer as the starting point of materials development, rather than the curriculum and its goals and objectives. Our survey of teachers use found that their primary use was wordprocessing, email, and the web (Murray et al., 2006). What is needed for a broader adoption of ICT—both more tools and by more teachers?

Researchers (e.g., Anderson & Nicholson 1995, Ellis & Phelps 2000) and practitioners (Corbel 1996, Murray & McPherson 2001) agree that appropriate and adequate training is required for teachers to exploit ICT. Both Corbel and Murray and McPherson, in surveying and interviewing teachers found that teachers needed training and technical support both before and during instruction. “Perhaps more than any other innovation it calls for support that is ready to hand, or, in the language of management, just-in-time” (Corbel, 1996: 57). Waiting for a technician to come to help causes great frustration to teachers. Yet, most schools cannot afford to have technical support staff ready to hand.

Teachers in our interviews indicated that they were discouraged from using technology because even after training, they still weren’t able to use ICT expertly. This lack of expertise is often exacerbated by learners in the class knowing about the technology than the teacher. But, teachers can use such student expertise since it empowers learners. Further, teachers were dismayed by the amount of time they needed to prepare activities for using ICT, only to find that the website had changed or their school’s computer system was down. Teachers therefore need to be courageous and flexible. And learners, too, need to be patient and flexible.

A FRAMEWORK FOR THE FUTURE

Recent studies of classroom practice have demonstrated that ‘A theory of CALL is a theory of language acquisition; the fact that the technology changes does not mean that the principles of language development do’ (Egbert, Chao & Hanson-Smith, 1999 pp1-2). Technology should not be the driving force in teachers’ decisions about what to do in the classroom. Nor should administrators require ICT without sufficient professional development and technical support. Rather, teachers need to choose from a variety of approaches, methodologies, practices and techniques, depending on the needs of their learners. Teachers need to explicitly teach cyberliteracy and scaffold learning, providing models of authentic language for learner. They need to choose (or develop) CALL programs that also provide explicit instruction and models and that have a high level of interactivity.

“We will fulfill the best use of computers in the classroom when we allow and encourage students to perform the most real tasks possible, to take advantage of the power of modern information and communication technologies to help try to change the world in ways that suit students’ own critical values and the interests of humankind” (Warschauer, 2001, p. 8).

REFERENCES

- Bauman, M. L. (1999). The evolution of Internet genres. *Computers and composition*, 16, 269-282.
- Belz, A. J., & Kinginger, C. (2002). The cross-linguistic development of address form use in telecollaborative language learning: Two case studies. *The Canadian Modern Language Review*, 59(2), 189-214.
- Debski, R. (2006). *Project-based language teaching with technology*. Sydney: NCELTR.
- Heift, T., & Schulze, M. (2007). *Errors and intelligence in computer-assisted language*

- learning*. New York: Routledge.
- Hoven, D. (1999). CALL-ing the learner into focus: Towards a learner-centred model. In R. Debski & M. Levy (Eds.), *World CALL: Global perspectives on computer-assisted language learning* (pp. 149-168). Lisse: Swets & Zeitlinger Publishers.
- Kenning, M. M., & Kenning, M. M. (1990). *Computers and language learning: Current theory and practice*. Chichester: Ellis Horwood.
- Lipscomb, P. (2002). *Reading what? on the Web*. Unpublished manuscript, Sydney, Australia.
- McPherson, P., & Murray, D. E. (2003). *Communicating on the Net*. Sydney: NCELTR.
- Murray, D. E. (1988). The context of oral and written language. *Language in Society*, 17, 351-373.
- Murray, D. E. (2003). Materials for new technologies: learning from research and practice. In W. A. Renandya (Ed.), *Methodology and materials design in language teaching: Current perceptions and practices and their implications* (pp. 30-43). Singapore: SEAMEO Regional Language Centre.
- Murray, D. E. (2004). The Language of Cyberspace. In E. Finegan & J. Rickford (Eds.), *Language in the USA* (pp. 463-479). Cambridge: Cambridge University Press.
- Murray, D. E. (2004). New frontiers in technology and teaching. In C. Davison (Ed.), *Information technology and innovation in language education* (pp. 25-44). Hong Kong: University of Hong Kong Press.
- Murray, D. E. (2005). Technologies for L2 literacy. *ARAL*, 25, 188-201.
- Murray, D. E. (2008, April 3-5). *Making choices*. Paper presented at the Colloquium on "Online teacher education: Communities we make, communities we break" at the 42nd Annual TESOL Convention, New York.
- Murray, D. E., Lloyd, R., & McPherson, P. (2006). *Teacher and learner use of the new technologies in the AMEP*. Sydney: NCELTR.
- Murray, D. E., & McPherson, P. (2002). *Using Planet English with AMEP learners*. Sydney: NCELTR.
- Murray, D. E., & McPherson, P. (2004). *Using the Web to support language learning*. Sydney: NCELTR.
- Murray, D. E., & McPherson, P. (Eds.). (2005). *Navigating to read; reading to navigate*. Sydney: NCELTR.
- Nunan, D. (1988). *The learner-centred curriculum: A study in second-language teaching*. Cambridge: Cambridge University Press.
- Rogers, E. M. (1995). *Diffusion of innovations*. New York: The Free Press.
- Salaberry, M. R. (1996). A theoretical foundation for the development of pedagogical tasks in computer mediated communication. *CALICO*, 14(1), 5-34.
- Taylor, R. P. (1980). *The computer in the school: Tutor, tool, tutee*. New York: Teacher's College Press.
- Thurstun, J. (2004). Teaching and learning the reading of homepages. *Prospect*, 19(2), 56-71.
- Tindale, J. (2005). Reading print and electronic texts. In D. E. Murray & P. McPherson (Eds.), *Navigating to read; reading to navigate*. Sydney, Australia: NCELTR.
- Warschauer, M., & Kern, R. (2000). *Network-based language teaching: Concepts and practice*. Cambridge: Cambridge University Press.

Changes in Japanese Students' English Writing Ability, Strategy-use, and Motivation during Their 3.5-year University Life

Sasaki, Miyuki
Nagoya Gakuin University
sasaki@ngu.ac.jp

The present study reports on the changes in 37 Japanese students' English writing ability, strategy-use, and motivation during the students' 3.5-year university life. It is a follow-up to Sasaki (2004) that heuristically examined the changes in 11 Japanese students' English writing behaviors over 3.5 years. For the present analysis, however, reflecting on the limitations of the cognitive-only approach I employed for Sasaki (2004), I adopted a more social perspective, assuming that students' cognitive changes are bound to be affected by their socio-cultural environments. In Sasaki (2004), the participants' 2 to 8 month overseas experiences had a significant impact on the changes in their L2 writing strategy use and motivation. For the present study with a larger sample size, I realized that not only the participants' study-abroad (SA) experiences, but also the different lengths of the SA experiences could have different effects on the participants' changes. I thus paid special attention to the effects of the different length of the students' study-abroad (SA) experiences on the changes over 3.5 years.

In this plenary talk, I will focus on the results for the following five research questions:

1. Do the participants' L2 writing abilities change according to the different length of their SA experiences?
2. Does their use of "Local Planning" (Planning what to write next)" change according to the different length of their SA experiences?
3. Does their L2 writing motivation change according to the different length of their SA experiences?
4. How do the L2 ability changes interact with the use of the "Local Planning" strategy?
5. How do their L2 writing motivational changes interact with their L2 writing ability changes?

Having asked these questions, I attempted to examine what could be achieved in terms of L2 writing ability development over 3.5 years, and how and why it could be achieved in the way it had been.

PARTICIPANTS

Four groups of Japanese EFL students volunteered to participate in the present study: the At-Home group who remained in Japan during the whole 3.5-year observation period, and the three Study-Abroad groups who spent some time in Canada, the United States, Australia, or New Zealand, mainly between their sophomore and junior years. The SA-2 Group spent two months abroad, SA-4 Group spent four months abroad, and the SA-8-to-11 Group spent 8 to 11 months abroad. When the study began, the students were all 18-year old university freshmen, majoring in British and American studies at the same university in Japan. They had studied English for six years by the time the study started.

DATA COLLECTION

I collected data in the first month of the participants' freshman year, and the third month of their sophomore, junior, and senior years, and I also interviewed them in the eighth month of their senior year.

My quantitative data consisted of

- (1) Argumentative compositions written for randomly selected topics, and
- (2) Participants' retrospective accounts, while watching their own video-taped writing performance, of what they were thinking about when they stopped writing.

And my qualitative data consisted of

- (3) Interviews (20-30 minutes) about L2 writing motivation given after each composition session, and
- (4) Post-senior interviews (40-60 minutes) on changes in L2 writing ability and motivation.

I statistically analyzed the quantitative data, and I analyzed the qualitative interview-data using Yang, Baba's (2004) research framework based on Engeström's (1987) expanded activity system.

RESULTS

The results of the present study can be summarized as follows:

- (1) The students' SA experiences changed the nature, but not the frequency, of their use of the Local Planning strategy.
- (2) The different length of the students' SA experiences had a significant impact on their L2 writing ability changes over their 3.5-year of university life.
- (3) The different length of their study-abroad (SA) experiences had a significant impact on their use of the Global Planning strategy, and possibly on becoming self-regulated writers.
- (4) Only those students who went abroad formed L2-related "imagined communities" that potentially motivated them to improve L2 writing ability.
- (5) Only those students who spent more than four months abroad became motivated to write better in L2, imagining the classes they took abroad even after they came home.
- (6) Only those students who spent more than eight months abroad became intrinsically motivated, and voluntarily practiced to improve their L2 writing, even though such an actions did not directly benefit their future career.

Although these results mainly treat the cognitive aspects of the participants' development, the interview data about their growth and goals as English writers provide evidence that second language writing development is not only cognitive but also social.

REFERENCES

- Sasaki, M. (2004). A multiple-data analysis of the 3.5-year development of EFL student writers. *Language Learning, 54*, 525-582.
- Yang, L, Baba, K., & Cumming, A. (2004). Activity systems for ESL writing improvement: Case studies of three Chinese and three Japanese adult learners of English. *Angles on the English-Speaking World, 4*, 13-33.

我國英語教育政策與教學之影響

Wu-chang Chang^a (張武昌), Shih-guey Joe^b (周碩貴), Meei-ling Liaw^c (廖美玲)

His-nan Yeh^a (葉錫南), Chiou-lan Chern^a (陳秋蘭)

Yu-ling You^d (游毓玲), Chin-cheng Huang^e (黃金誠)

^a 台灣師範大學

^b 長庚技術學院

^c 臺中教育大學

^d 彰化師範大學

^e 輔英科技大學

t22044@ntnu.edu.tw

sgjoe@mail.cgite.edu.tw

meeilingliaw@gmail.com

brianyeh@ntnu.edu.tw

clchern@ntnu.edu.tw

youyl@cc.ncue.edu.tw

cchuang@mail.npust.edu.tw

近年來在英語確定其世界語的地位以後，世界上各個非英語系國家無不積極將英語列為優先學習的外國語言。國內近年來也掀起一股全民學英語的風潮，自將英語教學向下延伸至國小三年級、大專院校推行國際化、一直到公務人員升遷須通過全民英檢，幾乎可以用「全民瘋英語」來形容這股風潮。政府為了提升國民的英語力，並進而提升國家的國際競爭力，行政院於 2002 年提出「2008：國家發展重點計畫」，將營造國際化環境與提升全民英語能力列為培育 e 世代人才的首要工作之一。為達此一目標，教育部在 2005 年至 2008 年的施政主軸中訂定各項行動方案，期使達到提升國民英語文能力與國際化之目標。除了「挑戰 2008」中各項與英語教學相關的政策以外，影響國內近年英語教育的政策還包括了因應教育改革而訂定的「九年一貫課程」與「教科書一綱多本」、廣設高中大學、技術學院升格大學、與將英語教育向下延伸至國小三年級等多項重要的教育政策。

近十年來，國內英語教學的相關研究質與量均明顯成長，但研究主題通常以學生學習動機、學習困難、教材編纂、測驗評量、聽說讀寫教學等為研究探討之重點，以政府教育政策對於國內英語教學現況所造成之影響為研究主題的研究卻很少。因此，本研究小組自 2006 年 8 月起執行為期兩年的整合型計畫，自國小、國中、後期中等教育、技職體系高等教育、至一般體系高等教育，針對政府教育政策對於各級學校英語教學的影響進行全面的檢視與探討。第一年的計畫係以問卷方式大規模調查國小、國中、後期中等教育、技職體系高等教育、以及一般體系高等教育的英文老師，對於政府教育政策對於英語教學的影響之看法。根據第一年問卷調查的結果，研究人員於第二年進一步以訪談的方式探討各級英語文教師、學校行政主管、教育主管機關人員、家長、補教業人員、

業界行政主管對於英語教學的看法以及英語教育政策的建議。因此，研究小組希望透過專題座談的方式，將研究成果公佈以引起更多英語文教師對於英語教育政策的注意，並藉由與會的各級英文老師的參與討論，提供研究人員繼續進行相關研究之參考。

如何幫助「英文（一）」課程低成就同學的學習

How to Facilitate Learning among Low Achievers in the Freshman English Course

Yue-hua Lee(李玉華), Eleanor Shin Leu(呂信), Ya-li Shih(施雅俐), Chi-fang Yu(余綺芳)

東吳大學

jyhlee@scu.edu.tw

esleu@scu.edu.tw

jasmine@scu.edu.tw

cynthiay@scu.edu.tw

近幾十年來由於台灣廣設大學院系，造成入學新生能力普遍低落，英文科目當然也不例外，例如東吳大學 97 學年日間部入學新生 2102 中有 695 人（33%）大學指考百分制裡未達 50 分，尤其因應教育部規定，各學校紛紛設立畢業門檻，東吳大學民國 97 學年入學新生必須通過東吳英檢門檻才能畢業，由於這些因素，課後輔導已不再是有升學壓力的國、高中生才實施的補救機制。

東吳大學自民國 94 學年即開始實施「大一英文」課程的線上補救教學，利用網路環境為低程度班級同學設計了「混成式教學/學習」(blended teaching/learning) 的補救措施，亦即學生每週仍於傳統教室上課四小時，利用東吳大學的網路學園平台 (<http://elearn.scu.edu.tw>) 建置幫助同學吸收閱讀及語練教材的學習資源，學生每週必須上網學習 2-3 小時。此外每班並配置助教實施課輔，助教除了協助任課教師督導全班同學的線上學習，並針對程度特別差或有學習困難同學執行每週兩小時的實體輔導。此補救教學機制實施至今已三年餘，學生人數從民國 94 學年的 487 人、民國 95 學年的 719 人、到民國 96 學年的 905 位同學，每年也都有十餘位不等老師以及十多位助教投入補救教學工作，助教除和任課教師密切聯繫且每月都必須繳交工作報告，任課教師每學期定期舉行會議檢討改進教材設計、線上學習環境、與助教的輔導工作等相關層面的問題。

本次專題座談將有四位任課教師各自花費 15-20 分鐘針對補救教學

線上教材的設計與運用（李玉華）

線上學習環境的優劣與學生回饋（施雅俐）

教師與助教相輔相成的互動（呂信）

補救教學學習成效與滿意度探討（余綺芳）

等四個層面做報告，然後開放剩餘 30-40 分鐘時間拋出下列多項議題與來賓互動，希望經由這樣的討論能夠集思廣益，找出大學裡最適合幫助低成就同學的教學模式，讓莘莘學子受益。

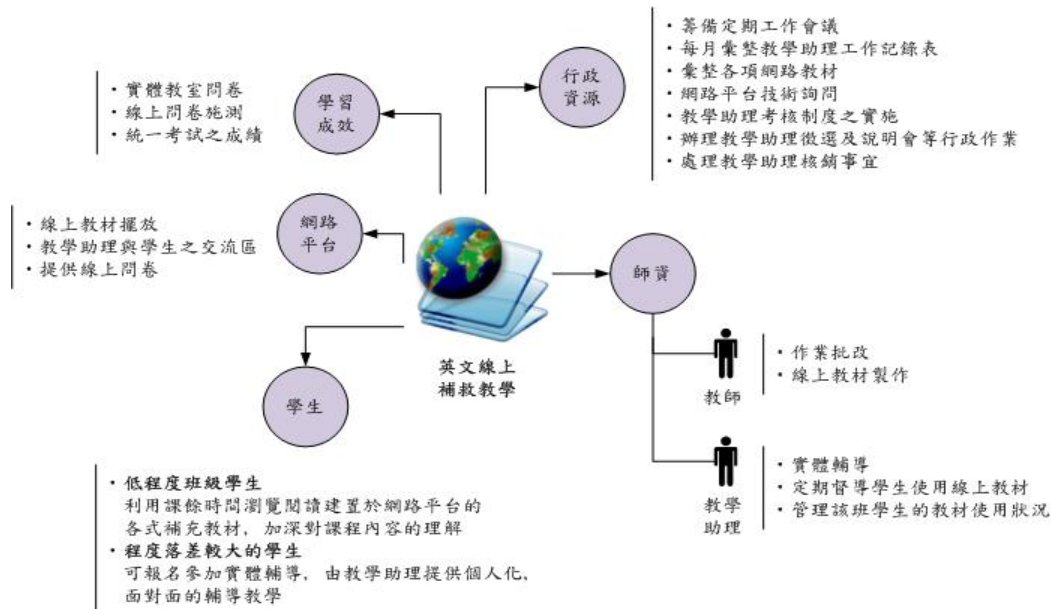
討論議題：

1. 目前線上自學教材主要以補強聽力、閱讀或寫作能力，相較之下，口說的補充學習則較被忽略，主要因素也許是因為科技的限制，也或許是因為缺乏真實溝通的對象，該如何突破限制，運用網路學習平台加強口說教學？

2. 東吳英文（一）線上補救課程與測驗的製作為統一完成後，複製給所有初級班級使用，身為線上補救課程研發製作者，該如何兼顧不同學院系級學生的特殊需求，甚至是教師的個別教學理念？
3. 東吳補救教學計畫的三方主要參與者為「學生」、「教師」及「助教」，教師應如何鼓勵低成就學習者參與額外輔導課程？教師如何指導助教選擇合適的教材或學習活動？教師與助教如何提升學生參與度並營造互動合作的機會，以創造出高動機並具正面的團體輔導動力(group dynamic)？又，如何將負面的社群影響力降到最低？
4. 目前東吳英文（一）線上補救課程之測驗題型有選擇、填充或是非題，起始原因主要為電腦平台設計，也較利於電腦標準化批改。這種看似「行為主義」的練習活動，(behaviorist activities)，著重於「正確性」(accuracy and form)的線上學習，對初級班的補救學習是否有用？又，如何豐富教材或測驗的多元性，以從 'Computer as tutor' 進一步提升為 'Computer as tool' ？

補救教學介紹（余綺芳）

東吳大學自民國 94 學年即開始實施「大一英文」課程的線上補救教學，利用網路環境為低程度班級同學設計了「混成式教學/學習」(blended teaching/learning) 的補救措施，意即學生每週仍於傳統教室上課四小時，利用東吳大學的網路學園平台 (<http://elearn.scu.edu.tw>) 建置幫助同學吸收閱讀及語練教材的學習資源，學生每週必須上網學習 2-3 小時。此外每班並配置助教實施課輔，助教除了協助任課教師督導全班同學的線上學習，並針對程度特別差或有學習困難同學執行每週兩小時的實體輔導。此補救教學機制實施至今已三年餘，學生人數從民國 94 學年的 487 人、民國 95 學年的 719 人、到民國 96 學年的 905 位同學，每年也都有十餘位不等老師以及十多位助教投入補救教學工作，助教除和任課教師密切聯繫且每月都必須繳交工作報告，任課教師每學期定期舉行會議檢討改進教材設計、線上學習環境、與助教的輔導工作等相關層面的問題。



線上教材及測驗的設計與運用(李玉華)

e-learn 平台的設立目的

94 學年開始，東吳大學大一英文學生依指考成績，分初、中、高三級授課，採統一教材。一般而言，英文低成就同學的特質，為自治力差、容易放棄，e-learn 線上

學習平台，希望能藉由每週的線上課程，改變同學的怠惰、被動的讀書習慣，點滴累積，提升英語能力，了解到每天給英文一點時間是必要的，也會比前更不怕英文，同時也因為 e-learn 線上測驗成績佔學期分數的一部分，讓同學於平時按部就班學習，儲存一部份的能量，於期中、期末會考時，不會因壓力過大而完全放棄，並且因為有好的成績表現，提升學生的學習動機與興趣。

線上教材及測驗的設計與運用

94 學年至今，e-learn 教材的呈現方式及測驗題型，依同學問卷結果及使用需求逐年修訂，另外初級班的同學的造句作業，也規定於線上繳交，任課教師並且以追蹤修訂方式線上批改，學生可以即時存取甚或觀摩其他同學改好的作業，成績由電腦自動登錄，方便日後查考。

大一英文課程線上補救教材以大一英文讀本、聽力授課內容為主，於 elearn 教材目錄中詳細列有同學每週須複習的教材及在設定期限內該完成的各項測驗，以下僅就 94 學年至 97 學年的線上教材及測驗做簡單介紹。

94 學年統一教材：Soochow English Reader + True Colors BK2，94 學年線上教材最特別的部分就是由余綺芳老師製作的文法單元，透過串流大師，詳細介紹英文重要基本文法概念，如完整句子和英文片斷的差異、重要時態、比較級、最高級等，讓同學有機會透過生動詳細的解說，複習重要的文法概念，余綺芳老師並且針對每課重要句型，透過不同的標記，做詳盡的分析，幫助同學掌握課內重要文法句型。在測驗方面則有文法測驗、字彙測驗及課外閱讀測驗，其中課外閱讀測驗除了讓同學加強閱讀能力，並有資深老師提供閱讀技巧解析及英文段落及篇章架構解說，幫助同學掌握文句之前後組合關係。聽力部分除了聽力測驗還包含考古題練習，讓同學熟悉考試題型及方式。

95 學年統一教材：Active Skills for Reading BK3 + True Colors BK2，閱讀部份提供 PPT 字彙介紹句型結構分析、重要片語及例句示範及句型結構分析，測驗包括字彙測驗、文法測驗及課外閱讀測驗。聽力部份除了有設有自我評量（不計時，不限定作答次數）並於每次評量之後，附加聽力小撇步（tips on listening）及考古題練習。

96 學年最大的變革是大一英文分開成閱讀和聽力兩門獨立課程，閱讀課本為 Active Skills for Reading BK3，聽力則為 Top Notch BK1，線上教材基本上沿續舊有形式，閱讀部份提供每一課重要單字、針對字彙及主要片語，設計有 PPT 檔，學生並可點選喇叭，聽到該部份的錄音，測驗則有文法句型測驗、字彙測驗及課外閱讀測驗。聽力部份則提供課本對話及 DVD 對話台詞，並加上自我評量，由學生在不限定時間情況下，做反覆聽力練習。

97 學年開始，大一英文初、中、高三級，各級採用不同教材，初級班教材：Active Skills for Reading BK2 + Top Notch BK1，閱讀方面：針對每課的重要單字設計有 PPT 檔，除了單字解釋、衍生字介紹，並藉由生動活潑的對話情境，加深同學對單字的印象，並熟悉其正確的使用方式。文法句型測驗，則於測驗後，除了提供正確答案，並於每題題後附上簡單的相關文法重點解說。聽力部份：除了每課的成就測驗及綜合測驗，並提供課文對話及文字檔供同學利用，補充聽力練習，則是以課本的對話為例，講授各種不同的學習策略及聽力技巧，最後輔以與課本相同主題之課外對話練習，讓同學實際應用每次學到的聽力技巧。

綜觀 94 學年至 97 學年的 e-learn 教材設計，除了注重多元的變化，更照顧到學生學習的需要，由單純的提供教材，陸續加上學習的策略技巧及影音的呈現，但求提高學生學習的樂趣，從而把被迫式的學習轉化成日後自發性地學習。

未來的期許

期望透過 elearn 的平台，以分數為正面誘因，刺激低成就同學的學習動力，紮

實英文的根基，以便日後與「大二英文」的課程銜接更加順利，進而使學生能在大四以前，通過東吳大學英文畢業門檻。在二~三年內，東吳 e-learn 平台將規劃改為 blackboard 方式，每班教材選用的彈性更大，每位初級班的授課教師能依不同學院學生的之特殊需求，選取適合自己班級的線上教材，或者以平台的補充教材為核心，依個人的專長、教學理念，充實自己授課班級的線上補充教材，更符合學生需求。也衷心希望有更多的專兼任老師可以參與線上教材的製作，使得網路教材更豐富、更多元。日後也希望有更多的機會和其他大專院校擔任低程度班級的老師有更多線上教材設計的經驗分享。

線上學習環境的優劣與學生回饋（施雅俐）

線上學習環境的建置

本校現在使用的 e-learn 網路學園平台主要有三大功能：課程內容、課程資訊及課程互動。

「課程內容」目前可說是線上補救教學的主軸，這裡設有 4 個區塊—課程介紹區、教材目錄區、測驗 / 考試區及作業 / 報告區。教師製作的自學教材與測驗練習即建置在這裡，也是初級班學生、教師及課輔助教最常使用的區域。

e-learn 網路學園的另外二大功能—「課程資訊」及「課程互動」，則依各班教師或課輔助教運用平台的程度，被利用及發揮的程度也有所不同。有些班級僅運用「課程資訊」中的「最新消息區」，作為課輔助教與學生間的聯絡簿；「成績資訊區」則可視為積分板，是一種預警制度，讓學生在學期中隨時檢視自己每一項作業、測驗的得分狀況和老師的評語，並與全班成績分布狀況評比。「課程互動」功能中，每班都會利用的區塊則是「郵寄老師、助教區」以及「問卷投票區」，作為學生寄電子郵件與期末進行課堂反應問卷的管道。

該平台還有其他互動性的功能，例如「議題討論」可視為網路論壇，可由教師設定議題，供學生於課後延伸討論，並透過自動轉寄功能，只要有學生上線回覆該議題，全班即可從電子郵件閱讀該文章，可激起回覆意願。不過這類互動性的功能，則因各授課教師教學理念及教學設計的差異，並非每班初級班都會建立此等網路論壇。

線上補強學習機制對初級學生的影響

從學生的回饋資料中發現，線上補強學習機制對初級學生的影響可分以下 4 個面向來探討：

自主學習認知上的提升

一般我們會認為低成就學生比較沒有能力進行自主學習，但是，透過要計分、設定測驗起迄日期並安排出進度的線上補強學習機制，學生起初的確是被要求上線做測驗，漸漸地竟培養出主動讀書的習慣。這樣的混成式補強措施或許可說是培養英語低成就學生發展自主學習的第一步，看似是行為主義的訓練其實也能激發學習者負起自我學習責任的意願，為邁向自我導向學習(self-directed learning)做準備。

學習策略的訓練與培養

英語低成就學生或初級英語使用者普遍缺乏自己尋找學習資源或運用學習策略的能力，他們甚至不清楚自己的問題在哪裡，學習速度亦較慢，因此他們需要借助鷹架支

撐的程度可能要比英語中、高級者要來得多或久一點。本補救教學計畫的線上教材或練習正針對低成就學生的學習型態，而成為這群學生在汪洋大海中的救生圈。除了課堂授課之外，課後還可點閱基本的補充教材或進行練習，加強一般同學普遍都有的字彙文法問題。做造句作業前，與其任由學生從網路上胡亂複製句子，不如提供佳作範本觀摩。我們發現，透過完成線上測驗，有些學生已經開始會自己去查閱字典參考書或與同儕展開討論。在線上聽力方面，除了解決指定教科書中聽力練習不足的問題，更藉由線上聽力教材，杜絕學生考前普遍不知如何準備聽力，僅全憑「實力」去應考的問題。

當然，教師在製作線上測驗時，無論答案詳解設計地多麼周全，仍然會有低成就學生希望一點閱測驗結果，就如同翻譯機般，一字不漏地呈現中英對照；或者，學生竟期望線上測驗要猜中期中、期末考的題目。這方面的問題，似乎可以透過授課教師運用課堂或網路平台的環境，多傳授學習策略或作答技巧，亦即除了教授課文，亦應著重「如何」閱讀、「如何」聆聽、「如何」學英文的能力。

電腦及網路的使用

另一個經常會影響線上學習機制順利運作的因素就是電腦或平台系統的問題，本校即曾發生在開學之初，授課教師及課輔助教大力推廣宣導使用網路平台，而學生在家要上線完成作業測驗時，竟發生連續數日當機問題，大大影響學生上網意願。

英語學習焦慮的降低

從學生的回饋資料中，似乎看不出線上學習顯著提升低成就學生的英文學習興趣。但值得注意的是，不少學生表示如在課後認真進行線上測驗，比起在課堂上的考試，焦慮感較低。如又獲得高分肯定，亦會增進信心與成就感，有了成就感，學英文的興趣自然會提升。

線上環境的限制以及可嘗試的方向

受限於電腦標準化批改，目前線上測驗的題型仍著重於「正確性」(accuracy and form)的練習。同時，統一製作的線上教材，例如單字片語的解釋，也多屬將電腦視為個人家教(computer as tutor)的性質。如要豐富教材或測驗的多元性，並增強溝通能力，未來應可將電腦輔助教學的角色漸漸調整為是一種「工具」(computer as tool)，例如以超連結連結至真實的語料來源(hypertexts)，以訓練閱讀，運用網路上廣大的聽力資料庫(e.g. YouTube, BBC)來練習聽力，或使用 WebQuest 模式進行探索學習，練習寫作、參與討論或簡報的能力。

另一方面，以英文(一)這類共通課程來說，班級數與授課教師人數眾多，本校目前以能力分班、各級有統一教材及統一考試、初級班更搭配混成式補救教學計畫等來做為基本品管策略。而在補救教學計畫中，每班雖配有課輔助教輔導學生使用網路平台，學生仍較在意授課教師是否也有能力親自使用網路平台；是否能搭配課堂進度，隨時調整網路學習的進度，並掌握學生線上學習的狀況。至於線上基本的教材及測驗於統一製作並複製給各班使用後，如要協助學生將線上學習發揮至極致，甚至如果要照顧到不同系級學生的需求，首先應協助授課教師增進電腦輔助教學的能力。當然，這是最具挑戰性的一點，不過也應該是能增進混成式教學效益的關鍵之一。

教師與助教相輔相成的互動（呂信）

94 學年度起，東吳「大一英文」實施能力分班。但三種程度的學生使用同一種閱讀與聽力之教材，期中與期末考也採統一命題方式，這對低程度班級同學是很大的挑戰。為了提升課後輔導的成效，開始實施「大一英文」課程的線上補救教學。為了減輕老師的負擔，每班配置助教一名。助教除了協助任課教師督導全班同學的線上學習，並針對有學習困難的同學實施每週兩小時的實體輔導。至今每年都有十餘位的老師以及十多位助教投入補救教學工作。助教除了和任課教師密切聯繫外，每個月都必須繳交工作報告，而任課教師每學期定期舉行會議，以檢討改進教材、線上學習環境、與協助助教實體輔導等相關的問題。希望藉由學生、教師及助教三方面的努力來提高學生主動學習的動機，進而提升英語能力。

助教的徵選及訓練

在每學年第二學期期中考後，發布助教徵選通知。只要對英語教學有興趣的大三、大四或碩士班學生都歡迎參加。暑假中安排報名、面試時間及進行面試。開學前兩週舉辦工作說明會，讓與會者了解並熟悉學習平台的運作。之後進行英語教學法之培訓，由培訓的表現決定助教任用資格。

助教的工作內容

助教工作主要分為兩大項目，一個是管理網路學園同學使用之情形，另一個是實體輔導。網路學園是補救教學相當重要的一個部份，裡面包含各式線上練習與測驗，提供低程度班級學生自我演練。而助教在此所扮演的角色，就是管理該班學生的使用狀況並且督促學生確實上網練習。各班助教需在學期開始之前熟習網路學園操作要領，包含增刪學員名單、發布公告、上傳教材、管理測驗、批改測驗等。其次助教每週需提供兩個小時實體輔導時間，藉由面對面的輔導教學，助教能夠直接了解並掌握學生學習情形，針對同學之學習盲點給予適度的補強與學習策略指導，以建立學習自信心。

助教與助教間的相互學習

助教於補救教學計畫實施期間，需定期填寫教學工作報告，回報各階段的工作情形與意見回饋。每月的報告包含(1)基本資料：教師姓名、助教姓名、輔導時間、輔導地點、紀錄月份。(2)輔導過程紀錄：本月公告事項、本月進度與輔導狀況、本月學生學習狀況以及測驗情形、本月輔導心得或困難處。每月的報告審核無誤後，除送交任課老師，讓老師了解學生學習狀況外，還需上傳至教學資源中心學生學習資源組網頁，供所有的老師及助教參考。

最重要的是參與每學期期初、期中考、以及期末考前舉行的補救教學會議。會議內容原則上為(1)重要事項宣布(2)教學心得、技巧分享(3)回報補救教學執行中所遭遇的困難，期望能藉由團體討論中腦力激盪的方式，幫助各位助教解決各種教學困難。有時更邀請已畢業的學長姐，尤其是目前當實習老師的，回校分享及經驗傳承。

助教與學生之間的互動

助教與學生之間的互動關係是補救教學中相當重要的一環。助教因年齡與學生相近，只要有熱情、積極的個性，就很容易與學生達成有效的溝通，這是聘用助教之始衷。為了能帶動學習意願低的同學，助教必須應用各種教學方法及技巧，即使再加上威脅或

利誘，最終目的是在幫助學生學習，才能發揮補救教學的功效。

助教參與輔導的過程中難免會碰到一些困難，例如：學生不上線做測驗及作業，即使發 e-mail、打電話也不理，公告點閱次數很低，重要事項需提醒很多遍，或線上測驗出現問題時不願意主動詢問，直到時限已過才要求重做等等。然而助教最大的困難是：實體輔導由學生自願參加，如何達到預期目標（例如：每次 6-10 位同學等）？雖然一開學，助教與學生已訂出最適當的時間來進行輔導，但有一部分同學常常缺課，或來輔導時睡覺、吵鬧等影響他人。這時助教必須和任課教師聯繫，商討對策以提升輔導人數及輔導品質。

雖然助教的困難有待解決，而 96、97 兩學年曾針對修讀大一英文參加補救教學課程的學生，實施補救教學課程助教滿意度調查問卷，其結果得知學生對助教的努力是給予高度的肯定。

助教與教師之間的配合

教師與助教的配合對於同學在網路學園的學習上有很大的幫助，在開學前之補救教學會議中，教師和助教需決定助教到班上和同學見面的時間，並進行事先的溝通。除了彼此認識外，助教務必介紹學生如何使用網路學園系統。在每週上課後教師和助教均可互相連繫，瞭解當週課程內容，是否有測驗需要延期等。課堂上同學提出的問題也可以轉給助教，必要時也可做為實體輔導的教材。助教和教師可以隨時決定實體輔導的名單，該名單可視需要增加或更改，而助教在實體輔導前連繫學生也可以透過教師提醒同學。

在上學期初期常有學生不清楚網路學園系統的使用，教師或助教需要多次地強調其功能及重要性，尤其要提醒學生測驗及作業時間。補救教學的學生們多半較被動，教師或助教利用「**最新消息**」隨時發布訊息，以此加強對學生的溝通，進而提高他們使用網路學園的意願。

實體輔導是針對課程內容對同學們做個別輔導，主要對象是對程度較差的同學，除了鼓勵學生參與外，教師可經由指導助教，以同學的學習困難為要，預先討論安排輔導方式及內容，以提升輔導的成效。學生參與實體輔導的次數及表現，可提供給教師做為成績計算的參考。

教師與助教的互動頻繁，任課教師針對其助教的工作表現，在每學期末進行考核，考核的指標分為：(1)專業能力與工作態度 (2)助教與教師、學生的互動 (3)網路學員課程管理 (4)實體輔導執行 (5)總評。教師送交出的考評結果將為下學期(年)助理是否續聘之重要依據。

成效及滿意度分析（余綺芳）

一般而言，學習成效的兩個指標是學業成績（期中、期末成績）以及滿意度。Knowles (1970) 和 Tough (1982) 指出滿意度是學生對學習活動的感覺，是學生經過學習活動後的心理反應。

統一考試成績

為瞭解補救教學的成效，特別提出 94 學年大一學生統一考試成績結果。由於大一英文課程雖然能力分班但卻使用統一教材、統一考試的方式，因而可以利用統一考試

的標準來檢驗混成式補救教學是否讓低程度班級同學的成績更有進步。94 學年度統一考試的結果顯現低程度班學生上學期期中平均成績 50 分，期末 56 分，下學期期中 61 分，期末 60 分。中程度班學生上學期期中平均成績 64 分，期末 68 分，下學期期中 69 分，期末 70 分。高程度班學生上學期期中平均成績 73 分，期末 76 分，下學期期中 78 分，期末 77 分。(不過由於學校規定大一英文上學期總成績低於 50 分者，就無法在下學期繼續修習大一英文課程，同時下學期初又加入一些重、補修生，低、中、高三個程度的上、下學期總人數都有變動，也略微影響了每個程度上、下學期的總平均分數。)

滿意度分析

於民國 95 年 5 月施測所有大一英文同學，來瞭解大一英文學生學習滿意度情形，問卷共發出 2,217 份，回收 1,576 份，回收率 71.1%。問卷分兩部分：

一部份為所有同學都必須回答，來比較未接受補救教學的中、高程度同學與接受補救教學的低程度同學不同的看法。

題目	低程度	中、高程度
滿意自我的學習態度	64%	60%
認為自己認真用功修課	63%	60%
上課心情愉悅	65%	59%
修習課程會提升英文興趣	60%	56%
課外利用線上學習資源自學	63%	48%
課程安排滿意度	66%	61%
教師製作教材滿意度	72%	61%
認同學校的英文學習資源	77%	70%
認同學校培養學生英文能力	70%	63%
課程提升其英文程度	62%	55%
認同課程有助於未來學習英文	65%	58%

另一部份則由低程度班級同學回答有關補救教學的問題。

題目	低程度
推薦大一英文課程繼續提供線上補救教學	75%
認同教學助理的制度	71%
認為線上補救教學可以有更多學習課程內容的機會	71%
認為「線上補救教學」對學習大一英文課程有幫助	70%
認為線上補救教學可以幫助養成每週固定學習的習慣	68%
認為線上補救教學課程內容設計合宜	68%
認為線上課程必須花費更多時間來學習大一英文內容	68%
認為修習線上課程有助於準備大一英文統一考試	68%
認為線上補救教學的補充教材難易度符合其需求	66%
希望大二英文課程繼續提供線上補救教學	66%

檢討補救教學與未來發展 (余綺芳)

補救教學成效

幫助同學培養自律及讀書習慣

95 年及 96 年四次線上問卷，分別有 73%、69%、68%、78% 同學同意每週線上測驗讓他們養成定期溫習課業的習慣，同時問卷裡有一題開放式問題，讓同學自由抒發

對補救教學的意見，每次問卷都有相當多數量同學表示「養成自發性的好習慣」、「有更多練習機會、幫助很大」、「有預習、複習效果」等。這樣的訓練對於低程度同學尤有助益，因為一年內課業的進步有限，但是如果能夠改變他們的學習英語的態度將會對他們終生的學習有助益。

提升教師教學專業成長

由於從事網路補救教學，必須瞭解如何幫助程度差的同學改變不正確的學習模式與方法，重建學習英文的興趣與信心。任課教師必須改變以往教室教學的模式，適應線上學習的環境，來幫助弱勢同學從事課餘的線上學習。由於學習使用新科技來輔助教學，大幅度增進了教師的教學知能。

落實課程革新與特色規劃

過去數十年內，東吳大學的大一英文課程對所有同學一視同仁，不管程度優劣，所有同學接受同樣時數的訓練，忽略程度差同學的特殊需求。94 年度開始實施的線上補救教學可說是課程的一大革新與特色。另一方面網路教學把重心轉移到學習者身上，訓練學生自主獨立學習。

充實教學電子化資源

由於所有文法、讀本教材皆是電子化教材，先用 PowerPoint 軟體製作，再用 webcam 錄製影音，及用 Stream Author 軟體串流而成，建置於網站上。另外還建置線上測試題庫，讓學生作線上作業測驗及參與討論，這些電子資源容易編輯複製，可以方便修訂，讓以後的學生更能有效利用。

加強學生學習輔導

以往每班大一英文班級平均有五、六十位同學修習，任課教師很難個別輔導有問題同學，助教實體輔導機制可以滿足有特殊需求的同學，加強補足他們以往學習的缺失，幫忙建立這些同學的學習自信心。

提升「教與學」環境品質

網路學園平台提共了一個教室外的虛擬教室，每週教材、作業、測驗皆妥善建置網站，學生查詢教材目錄即可清楚明瞭每週進度工作，安排時間上網使用教材自學，如果一次看不懂可重覆播放，在線上繳交的作業老師批改完後，完整存放網站，可供所有同學觀摩，所有測驗完成後，皆開放複習，學生可以瞭解自己所犯錯誤。老師、助教可隨時上網利用「學生學習統計」「教材使用統計」的功能督導學生學習，如果發現某些同學太久未上網學習，即可發簡訊提醒。藉此可充分掌握學生的學習情況督導他們學習。這樣的環境讓中、高程度班級同學都欣羨。

計畫需改進之處

如何提升同學的學習動機

由於低程度同學學習意願低落，所以利用幫助他們通過統一考試作餌，又利用線上成績佔學期成績 30% 來箝制他們從事線上學習，所以輔導及學習重心都是放在準備考試，像是 catch 22 互相牽制的情況，學習及輔導內容就難免枯燥。但是如果要想讓同學輕鬆愉快學習有趣的教材，就不能確保他們在考試上得分，結果學生會覺得這樣的學習對考試無助益而失去參與的意願。而且如果線上成績不計分，學生不當回事，如果要算分，就必須能檢驗他們的學習，確實很難設計。不過希望在新的學年裡還是要嘗試改變設計，強調同學的學習動機。

行政業務繁瑣

由於牽涉到數百位同學的補救教學，諸如規劃統籌計畫案、與課務組及電算中心協商使用網路學園平台、與老師協調商議製作教材、督導助理將教材上網、面試線上助教並定期培訓他們、與任課教師定期集會討論線上教學問題、幫忙解決各班同學線上學

習問題、收集整理參與補救教學同學線上問卷結果分析、以及其他許多諸如經費核銷、定期撰寫計畫案報告等等行政繁雜事務，所以目前由語言教學中心及教學資源中心通力合作，不過要整合兩個單位也有實際困難。

線上學習的困擾

許多同學反應由於家中與學校連線的問題，經常做作業或測驗一半就當機，再重新開機就已經測驗結束，徒生許多困擾。有幾回聽力考古題測驗多達三十題，有時傳輸太慢更容易當機，所以有許多學生都無法完成測驗。還有少數學生抱怨聽力測驗聽不到聲音，由於事關個人電腦設定及設備等問題，實在無法解決。

學生使用率及輔導學習有困難同學

學期初許多學生採觀望態度，不願意上線練習，許多教師及助教花費很多心血，循循善誘同學上網練習，不斷延後截止時間，給他們更多機會，通常要上課一個月後才會逐漸穩定，每班平均一半以上學生會固定做作業測驗，但是使用率仍待提升。尤其低程度班級同學成績稍有進步覺得不會被當就鬆馳下來，老師及助教必須不斷耳提面命讓他們持續學習。

Developing Statistic-based and Rule-based Grammar Checkers for Chinese ESL Learners

Howard Hao-jan Chen (陳浩然)
National Taiwan Normal University
hjchen@ntnu.edu.tw

Many ESL students need to improve writing skills to pass various language tests. Thus, writing teachers would need to grade many compositions and provide feedback. To help teachers reduce their teaching loads and to give students faster feedback, there is a pressing need to develop better grammar checkers for ESL writers. In this paper, we first introduce two different types of grammar checkers (rule-based grammar checker and statistic-based grammar checker) we developed. Then we examined the strengths and weaknesses of these two types of checkers. About 400 sentences (about eleven major types of errors) randomly selected from a large Chinese EFL learner corpus were used to test these two checkers. The results showed that statistic-based checker performed better in these seven categories, and the rule-based checker only performed better in two categories (conjuncts and sentence structures). Based on the preliminary findings, statistic-based grammar checker could detect more errors, but it did not provide adequate answers and explanations. On the other hand, rule-based checker found fewer errors but it often could offer models and explanations. Each type of checker has its own strengths and limitations, and the proper integration of these two types of grammar checkers might produce better results.

INTRODUCTION

It is challenging for language learners to become proficient writers of the target language. Second language learners would need many writing practices before they can write fluently. In addition to writing practices, many teachers and researchers believe that learners need to receive proper feedback on their writings (Ferris, 2003). If learners only keep on writing and do not receive any feedback, they will not be able to make progress quickly. Even though the role of corrective feedback in second language learning remains controversial, many teachers and learners firmly believe that feedback plays an important role in second language writing.

However, if the class size is too large, it is often difficult for writing teachers to provide detailed feedback. In many ESL/EFL settings, teachers would need to work with 40 or 50 students in their classes. Reading and commenting on students' essays become a great burden for many writing teachers. Many senior high school and college teachers are searching for a useful grammar checker but they often cannot find suitable tools to use.

Different Approaches to Developing Grammar Checkers

Given that there is an urgent need for a grammar checker, we decided to build a grammar checker by using the NLP (Natural Language Processing) technologies and corpus resources. Researchers have been interested in developing grammar checkers for various languages. There are several possible designs for a grammar checker; we have to be careful in choosing the right approaches to implement a new grammar checker based on new NLP tools and corpus resources.

According to Naber (2003), there are basically three different ways to implement a grammar checker. Each approach has its strengths and weakness.

1. Syntax-based checking, as described in Jensen et al. (1993). In this approach, a text is

completely parsed, i.e. the sentences are first analyzed and each sentence is assigned a tree structure. The text is considered incorrect if the parsing does not succeed. A robust syntactic parser plays a very important role in this approach.

2. Statistics-based checking, as described in Atwell and Elliott (1987) and Chodorow and Leacock (2000). In this approach, a language model is trained from a large training corpus (e.g., a native corpus), which contains many short phrases. It can be used for detecting and correcting certain types of grammar errors, where local information is sufficient to make decision. Sequences which occur often in the corpus can be considered correct in other texts, and uncommon sequences might be errors. Some recent developments of grammar checker based on corpus linguistics and statistics seem to be quite useful (Chodorow & Leacock, 2000; Sjobergh, 2005; Wu, Su, Jiang, & Hsu, 2006). A popular example is the grammar and style checker developed by ETS.

3. Rule-based checking, as described in Naber (2003). In this approach, a set of rules is matched against a text which has at least been POS tagged. This approach is similar to the statistics-based approach, but all the grammar rules are developed manually. Park et al. (1997) is a study using this approach.

Statistic-based Grammar Checker

In this paper, we discuss the statistics-based approach and rule-based approach and discuss how they can be used in dealing with various learner errors.

Statistics-based approach is the approach which is currently used by ETS and several other research teams (e.g., Chodorow & Leacock, 2000; Sjobergh, 2002). The advantages of using statistics-based approach are the followings: first, there is no need to write grammar rules manually. Second, it can detect many errors which cannot be detected by traditional grammar checker and parsers. The bigram and trigram can detect more errors. Third, the lexical errors are more likely to be detected by this type of grammar checker.

The procedures of developing a statistics-based grammar checker are introduced briefly. The procedures used to detect the violations of English grammar in a statistics-based grammar checker are not complicated. The grammar checker system is first trained on a large corpus of edited text, from which it extracts and counts *bigrams* that consist of sequences of adjacent word.¹ We in this study used the British National Corpus as the reference corpus. The British National Corpus (or just BNC) is a 100-million-word collection of samples of written and spoken English from a wide range of sources. It was compiled as a general corpus (text collection) in the field of corpus linguistics. The corpus covers British English of the late twentieth century from a wide variety of genres with the intention that it be a representative sample of spoken and written British English of that time. We first used SRI (Stanford Research Institute) Language Modeling Toolkit (SRILM) to develop an ngram language model based on the British National Corpus (BNC). Various bigram and trigrams were extracted from the BNC corpus and stored in a large database. A web-based grammar checker interface linked to the ngram database was also developed.

ESL students can submit their writings to the web-based grammar checker, and the system then searches students' writings for bigrams and trigrams. The bigrams and trigrams which never or rarely show up in BNC corpus were highlighted by the system. Within a few seconds, this checker can quickly check the article and highlight the problematic word strings (clusters) not found in the native corpora. With the help of the bigram and trigram

¹ Bigrams are groups of two written letters, two syllables, or two words, and are very commonly used as the basis for simple statistical analysis of text; one of the most successful language models for speech recognition (Collins, 1996). An n-gram is a sub-sequence of n items from a given sequence. n-grams are used in various areas of statistical natural language processing and genetic sequence analysis. An n-gram of size 1 is a "unigram"; size 2 is a "bigram"; size 3 is a "trigram".

information from the very large native corpora, the system can efficiently detect and highlight the problematic words in students' essays.

The web interface of the statistic-based grammar checker is shown below in figure 1. The highlighted words, shown below in Figure 2, are possible errors identified by the statistic-based checker.

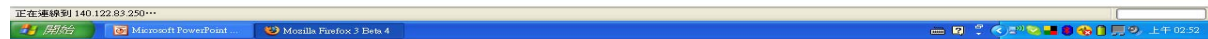
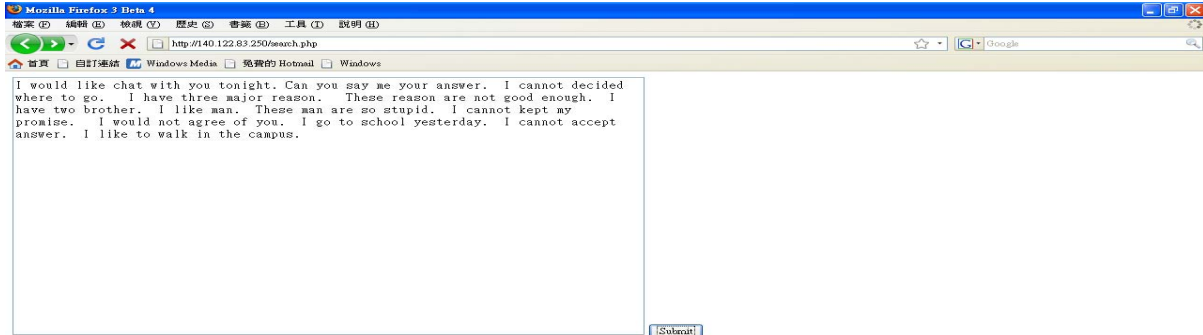


Figure 1. The Web Interface of NTNU Statistic-based Grammar Checker

I would like chat with you tonight.
 Can you say me your answer.
 I cannot decided where to go.
 I have three major reason.
 These reason are not good enough.
 I have two brother.
 I like man.
 These man are so stupid.
 I cannot kept my promise.
 I would not agree of you.
 I go to school yesterday.
 I cannot accept answer.
 I like to walk in the campus.

Figure 2. The Errors Highlighted by the Grammar Checker

Rule-based Grammar Checker

In addition to the statistics-based grammar checker, we developed a rule-based grammar checker for ESL learners. Naber (2003) argued that the rule-based approach has some advantages: Unlike the parser-based checkers, a sentence does not have to be complete to be checked. The rule-based grammar checker only uses POS tags and grammar rules to detect learners' errors. Rule-based checker is not difficult to configure, as each grammar rule has a specific description and can be turned on and off individually. In addition, detailed error messages with helpful comments and rule explanations can be added in. It is possible to extend the rule system to cover various common error cases.

It should be pointed out that the development of rule-based grammar checker engine is often time-consuming. However, as more and more researchers are interested in this application. Some rule-based grammar authoring tools are available. After reviewing several different grammar development tools, we found the WGrammar Software Development Kit is quite suitable for grammar checker development. It can be used as the core grammar engine for checker development and we can gradually add in grammar rules to

locate and identify problems committed by writers of English in dozens of problem categories.

WGrammar can recognize words according to the following parts-of-speech categories: Adjectives; Adverbs; Conjunctions; Singular and plural nouns; Prepositions; Pronouns in nominative, objective, and possessive case. WGrammar comes with a database of over 78,000 words tagged by parts of speech.

WGrammar uses the pattern matching capabilities to detect errors. We can supplement this set by adding new patterns. Moreover, Word patterns can match words using wildcards and words as parts-of-speech categories (e.g., plural nouns or third-person-singular verbs). WGrammar's search engine matches patterns containing sequences of words, using wildcards and parts-of-speech categories (e.g., match words as verbs, nouns, adjectives, etc.). Each pattern has an associated replacement, and the replacement can refer to specific words matched by the pattern.

This grammar authoring tool helps us save time in developing a completely new rule-based system. In addition to the rules included with the prototype, we also add on many rules based on the common errors found in Chinese ESL learner corpora. These new rules should help to improve the performances of this rule-based checker to deal with the errors in Chinese ESL learners' writing.

The rule-base grammar checker, as shown below in Figure 3, takes a learner's text and returns a list of possible errors. To detect errors in a text, each word of the text is assigned its part-of-speech tag and each sentence is split into chunks (e.g. noun phrases and verb phrases). Then the text is matched against all the checker's pre-defined error rules. If a rule matches, the text is supposed to contain an error at the position of the match. The errors were highlighted and shown to the learner. When a learner moves the cursor to a specific error, she then found detailed message explain why this is an error.

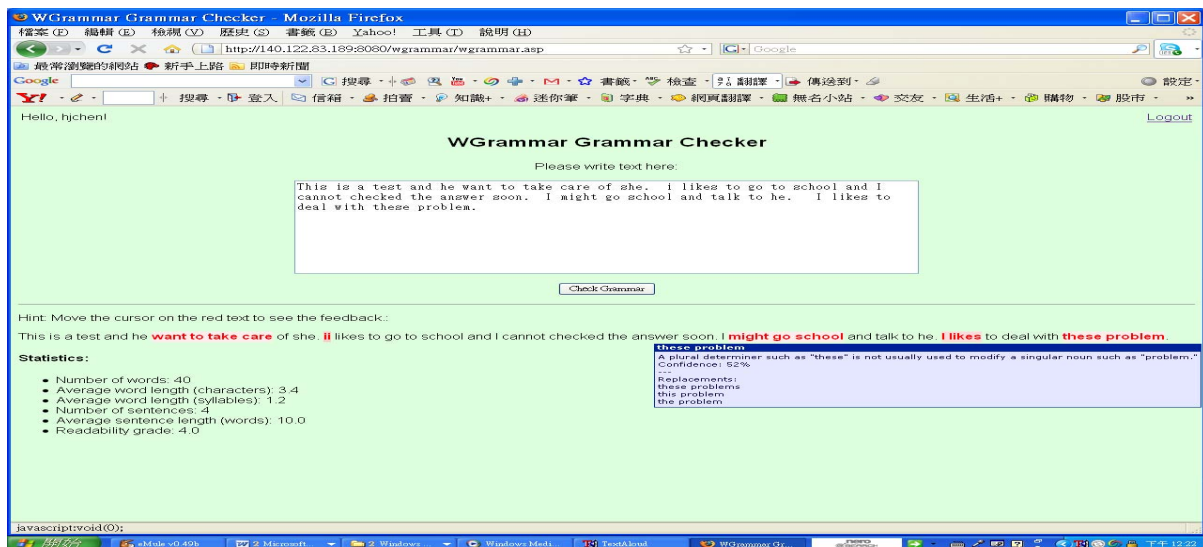


Figure 3. The Highlighted Errors and Explanations Provided by the Rule-based Checker

A PRELIMINARY EVALUATION OF TWO TYPES OF GRAMMAR CHECKERS

After building the prototypes of these two different grammar checkers, we need to further test their capacities with real learner errors. We found a very useful resource for testing these checkers is a learner corpus called CLEC (Chinese Learner English Corpus). The corpus contains 1 million words of English compositions collected from Chinese learners of English with differing levels of proficiency, covering senior secondary school students, English-major, and non-English-major university students in China. (cf. Gui & Yang, 2002;

Yang & Wei, 2005) The corpus is also error tagged according to an error marking scheme of 61 types of error, including various lexical, grammatical, semantic, and sentence level errors. The CLEC project was directed by Professor Gui Shichun of Guangdong University of Foreign Studies, and Professor Yang Huizhong of Shanghai Jiaotong University, both in China. This learner corpus is precious because the errors were tagged manually by language teachers.

To examine the performances of these two different types of grammar checkers, we randomly selected 400 sentences (5-10 sentences from each of these 61 categories) from the CLEC corpus. Given that the errors in these sentences were all marked by teachers, we then could check if these two different grammar checkers could find the errors marked by teachers. We could also calculate the percentages of the errors found by these two checkers. Table 1 summarizes the results of our tests.

Table 1. *The Error Detected by Statistic-based and Rule-based Grammar Checker*

1. Word Forms

Codes	Errors	Statistic-based	Rule-based
Fm1	Spelling	80%	0%
Fm2	Word building	80%	20%
Fm3	Capitalization	40%	20%

2. Verb Phrases

Codes	Errors	Statistic-based	Rule-based
Vp1	Pattern	70%	10%
Vp2	Set phrase	60%	40%
Vp3	Agreement	70%	50%
Vp4	Finite/non-finite	70%	50%
Vp5	Non-finite	70%	80%
Vp6	Tense	0 %	0%
Vp7	Voice	60%	20%
Vp8	Mood	40%	20%
Vp9	Modal/ auxiliary	70%	20%

3. Noun Phrases

Codes	Errors	Statistic-based	Rule-based
Np1	Pattern	50%	50%
Np2	Set phrase	60%	10%
Np3	Agreement	70%	30%
Np4	Case	70%	40%
Np5	Countability	70%	70%
Np6	Number	50%	30%
Np7	Article	60%	50%
Np8	Quantifier	70%	60%
Np9	Other determiners	20%	20%

4. Pronouns

Codes	Errors	Statistic-based	Rule-based
Pr1	Reference (difficult to detect)	20%	40%
Pr2	Anticipatory it	0%	0%
Pr3	Agreement	0%	10%
Pr4	Case	50%	20%
Pr5	Wh-	40%	40%

Pr6	Indefinite	40%	20%
-----	------------	-----	-----

5. Adjectives

Codes	Errors	Statistic-based	Rule-based
Aj1	Pattern	60%	20%
Aj2	Set phrase	40%	0%
Aj3	Degree	70%	40%
Aj4	-ed/ -ing confusion	70%	40%
Aj5	Predicative/ attributive	60%	40%

6. Adverbs

Codes	Errors	Statistic-based	Rule-based
Ad1	Order	70%	60%
Ad2	Modification	60%	60%
Ad3	degree	60%	20%

7. Prepositions

Codes	Errors	Statistic-based	Rule-based
Pp1	Pattern	30%	30%
Pp2	Set phrase	30%	20%

8. Conjunctions

Codes	Errors	Statistic-based	Rule-based
Cj1	Pattern	20%	60%
Cj2	Set phrase	20%	20%

9. Word Usage

Codes	Errors	Statistic-based	Rule-based
Wd1	Order	60%	50%
Wd2	Part of speech	70%	50%
Wd3	Substitution	60%	40%
Wd4	Absence	50%	20%

10. Collocations

Codes	Errors	Statistic-based	Rule-based
Cc1	Noun/ noun	60%	10%
Cc2	Noun/ verb	60%	20%
Cc3	Verb/ noun	40%	10%
Cc4	Adj/ noun	60%	10%

11. Sentence Structures

Codes	Errors	Statistic-based	Rule-based
Sn1	Run-on sentence	0%	0%
Sn2	Sentence fragment	0%	60%
Sn3	Dangling modifier	20%	40%
Sn4	Illogical comparison	30%	20%

The Strengths and Limitations of the Statistic-based Grammar Checkers

Based on the test results, we found that the statistic-based grammar checker was good at detecting errors in the “local” domains or “narrow” domains. It can indeed find many errors in the following categories.

- a. Word Forms (Spelling, Word building)
- b. Verb Phrases (Pattern, Set phrases, Agreement, Finite/non-finite, Non-finite, Voice, Modal/ auxiliary)
- c. Noun Phrases (Set phrase, Agreement, Case, Countability, Article, Quantifier)

- d. Adjectives (Pattern, Degree, -ed/ -ing confusion, Predicative/attributive)
- e. Adverbs (Order, Modification, degree)
- f. Word Usage (Order, Part of speech, Substitution)
- g. Collocations (Noun/ noun, Noun/ verb, and Adj/ noun)

However, the statistic-based grammar checker had great difficulties in detecting errors in the following five categories.

- a. Pronouns
- b. Prepositions
- c. Conjuncts
- d. Sentence Structures
- e. Tenses

It seemed clear that the statistic-based checker failed to catch the errors in the non-local domains or “broad” domains. The results are not difficult to interpret because the language models used by statistic-based grammar checker were bigram and trigrams. The statistic-based checker thus was very efficient in detecting more “local” errors like various improper word combinations. It performed very well in dealing with errors like spelling errors and agreement errors.

Nevertheless, the statistic-base grammar checker failed to capture errors if the errors were not word combination problems or the errors involve non-adjacent word strings or some conflicts across different sentence boundaries.

The Strengths and Limitations of the Rule-based Checker

The rule-based grammar checker seems more conservative in detecting learner errors. In general, it highlighted fewer problems. Based on Table 2, it can indeed find errors in verb phrases (non-finite sentences), noun phrases (countability and quantifier), adverb (order and modifications), conjuncts (patterns), and sentence structures (fragments).

However, the rule-based grammar checker performed poorly in the following categories.

- A. word forms
- B. prepositions
- C. pronouns
- D. collocations

It seems clear that both types of checkers had difficulties in dealing with the preposition and pronoun errors. For the other two categories like adjectives and word usage, the rule-based checker could detect a few errors but its performance was not satisfactory.

The errors found by the rule-based grammar checker were not as many as those found by statistic-based grammar checker. It was clear that the statistic-based checker was far more sensitive in detecting ESL learner errors because it was based on bigrams and trigrams. The statistic-based checker runs through the sentences by checking each and every word combination. The rule-based checker work differently by checking the text with reference to the grammar rules stored in the rule database. Although quite a few grammar rules were added into the rule-based grammar checkers, there are still many ungrammatical patterns which were not covered by the rule database. Even though we can continue to add in more rules to capture some ungrammatical sentences, it should be noted that the rule-based grammar checker will still have great difficulties in detecting collocation errors and various lexical problems. It is very difficult to write various rules for individual words. Even if we try to write rules for individual words, there will be too many rules and the error processing mechanism of the whole system will be very slow.

POSSIBLE DIRECTIONS TO IMPROVE GRAMMAR CHECKERS

The Statistic-based Grammar Checker

Based on the test results, the statistic-based grammar checker seems to be useful for second language learners. However, the checker so far can only effectively highlight problems in learners' writing; it fails to provide corrective feedback on the errors it identifies. Highlighting errors is a way of implicit corrective feedback, and it can be useful for some learners. However, most learners might need more explicit feedback on their writing errors. If possible, they need some metalinguistic feedback or correct model to follow. It is important to figure out an effective way of adding in proper feedback for the statistic-based checker.

For the statistic-based grammar checker, it is difficult to give explicit feedback since the system does not have explicit rules like the rule-based grammar checkers. One possible way to provide suggestions is to search the native corpus and find out the similar expressions and show them to learners. However, making the right recommendations for learner errors is not easy in many cases. For instance, some "confusing words" errors are easier. The affect of this medicine was so obvious. It would not be difficult to recommend the word "effect" if we have built a list of "confusing words" in the database. However, it would be difficult if these errors are collocation errors. For instance, if we need to provide automatic feedback on a sentence like "I would like to increase my life". It would be more difficult since there are so many verbs can co-occur with the noun "life". If we simply try to locate the possible verbs, we cannot make right suggestions.

A possible solution is to allow learners to search for the powerful collocation retrieval tools like "just the word" (<http://193.133.140.102/JustTheWord/>) and decide which verb is better for their intended meaning. Google can also be used as learning tool for checking the usage. For a certain string highlighted by the statistic-based grammar checker, students can search the strings and they can find some correct usage from Google. If a student write "...prepare the exam", she can easily find the correct answer "prepare for the exam" from Google.

Another two possible ways of improving the statistic-based grammar checker would involve a large reference corpus, proper use of POS-tagged reference corpus. The new data set, Web 1T 5-gram Version 1 contributed by Google Inc., contains English word n-grams and their observed frequency counts. The length of the n-grams ranges from unigrams (single words) to five-grams. The new huge corpus will be useful for statistical language modeling. We will soon develop another ngram checker based on Google data and hope this version with a much larger corpus will reduce some false alarms.

Incorporate information from tagged corpora would be another way of improving the statistic-based checker. A better checker can use the tag information to detect errors. Similar to raw text data, a part of speech (POS)-annotated corpus can be used to build a list of POS tag sequences. Some sequences will be very common (for example determiner, adjective, noun as in the old man), others will probably not occur at all (for example determiner, determiner, adjective). The powerful tagger can help us to tag large text files. The clusters of various POS tags can be further extracted and used to help to identify learners' errors. The tagged corpus also helps to provide better grammatical explanations.

The Rule-based Grammar Checker

Although the rule-based grammar checker has some serious limitations in dealing with the lexical errors generated by second language learners, it can help to resolve a few thorny grammatical problems. The special feature of rule-based checker is that it can often provide learners with models and explanations. Its feedback is more effective for second language learners.

In the future, we can continue to add in grammar rules for dealing with more structural problems. The more efficient way of finding out the real errors made by Chinese ESL

learners is to use learner corpora. More and more Chinese ESL learner corpora are available and reports on common errors are published (Wen, Wang & Liang, 2005). These various error analysis reports can be used to prepare various rules. The data come from the learners who share the first language, the L1 transfer phenomena will be systematically extracted from the corpus. More L1 transfer errors can be encoded into the rule-based grammar checker. The information from various learner corpora should help researchers to prepare grammar rules and feedback.

The Integration of Two Types of Grammar Checkers

In previous sections, various ways of improving each type of grammar checker were proposed. Given that each type of checker has its own strengths and limitations, it seems reasonable to integrate these two different systems. The strengths and weaknesses of these two types of checkers were summarized in the Table 2 below.

Table 2. *The Strengths and Weaknesses of Different Grammar Checkers*

Error categories	Statistic-based checker	Rule-based checker
Word Forms	Acceptable	Poor
Verb phrases	Acceptable	Acceptable
Noun phrases	Acceptable	Acceptable
Pronouns	Poor	Poor
Adjectives	Acceptable	poor
Adverbs	Acceptable	Acceptable
Prepositions	Poor	Poor
Conjuncts	Poor	Acceptable
Word usage	Acceptable	Poor
Collocations	Acceptable	Poor
Sentence structures	Poor	Acceptable

For the eleven major types of learner errors, statistic-based grammar checker can help with 7 categories and rule-based grammar checker can help with at least 5 categories. For some error categories (verb phrases, noun phrase, and adverbs), both types of grammar checkers can help. For other categories (word forms, adjectives, word usage, and collocations) statistic-based checkers can provide more help. Rule-based checker can help more with conjuncts and sentence structures. If the two different grammar checkers can be integrated into one system, then the integrated system should be able to detect more learner errors. However, the two checkers still failed to deal with some structures like prepositions, pronouns, and tenses adequately.

It might be very difficult for the statistic-based checker to deal with these three categories of errors. However, a long list of preposition errors might be prepared and added into the rule-based grammar checker to detect Chinese ESL learners' preposition errors. For pronoun errors and tense errors, it is also possible to write more individual grammar rules to deal with these errors commonly found in learner corpus. It seemed obvious that the integration of these two different types of grammar checker can help to capture more learner errors. It might be necessary to further integrate these two different checkers into one single system.

REFERENCES

- Attali, Y. (2004). Exploring the feedback and revision features of Criterion. Paper presented at the National Council on Measurement in Education (NCME) in San Diego, CA.
- Atwell, E. and Elliott, S. (1987). Dealing with Ill-formed English Text. In: R. Garside, G. Leech and G. Sampson (eds), *The Computational Analysis of English: A Corpus-based*

- Approach. London: Longman.
- Burstein, J., Chodorow, M., & Leacock, C. (2003). Criterion: Online essay evaluation: An application for automated evaluation of student essays. *Proceeding of the Fifteenth Annual conference on Innovative Application of Artificially Intelligence*, Acapulco, Mexico.
- Burstein, J., Chodorow, M., & Leacock, C. (2004). Automated essay evaluation: The Criterion online writing service. *AI Magazine*, 25(3), 27-36
- Chen, H. J. (2006). Examining the Scoring Mechanism and Feedback Quality of My Access. *Proceedings of Tamkang University Conference on Second Language Writing*.
- Chen, C-F. & Cheng, W-Y. (2006). The Use of a Computer-based Writing Program: Facilitation or Frustration? Paper presented at the 23rd International Conference on English Teaching and Learning in the R.O.C. May 27, 2006. Kaohsiung: Wenzao Ursuline College of Languages.
- Chodorow, M., & Leacock, C. (2000). An unsupervised method for detecting grammatical errors. In *Proceedings of the 1st Annual Meeting of the North American Chapter of the Association for Computational Linguistics*, 140-147.
- Elliot, S. (2001). IntelliMetric: From here to validity. Paper presented at the Annual Meeting of the American Educational Research Association (AERA), Seattle, WA. April, 2001.
- Ferris, D. (1999). The Case for Grammar Correction in L2 Writing Classes: A Response to Truscott. *Journal of Second Language Writing*, 8(1), 1-11.
- Ferris, D. (2003). Responding to student errors: Issues and strategies. *Treatment of error in second language student writing*. Ann Arbor, MI: University of Michigan Press.
- Granger S. (2002). A Bird's-eye View of Computer Learner Corpus Research. In Granger S., Hung J. and Petch-Tyson S. (eds) *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Language Learning and Language Teaching 6. Amsterdam & Philadelphia: Benjamins pp. 3-33.
- Granger, S. (2003). Error-tagged learner corpora and CALL: a promising synergy. *CALICO (special issue on Error Analysis and Error Correction in Computer-Assisted Language Learning)* 20(3), 465-480.
- Granger, S. (2004). Computer learner corpus research: current status and future prospects. In Connor U. and T.A. Upton (eds) *Applied Corpus Linguistics: A Multidimensional Perspective*. Amsterdam & Atlanta: Rodopi. 123-145.
- Gui, S. and Yang, H. (2002). *Chinese Learner English Corpus (CLEC)*. Shanghai: Shanghai Foreign. Language Education Press.
- Higgins, D; J. Burstein; D. Marcu; C. Gentile, (2004). Evaluating multiple aspects of coherence in student essays. Retrieved May, 20, 2006, from: http://www.ets.org/Media/Research/pdf/erater_higgins_dis_coh.pdf
- Holland, M. V., Kaplan, J.D., and Sama, M.R., eds. (1995). *Intelligent Language Tutors: Theory Shaping Technology*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.
- Jensen, K. (1993) "PEG: the PLNLP English grammar", in Jensen, K., Heidorn, G. and S. Richardson (eds) *Natural Language Processing: the PLNLP Approach*
- Liou, H. (1992). An automatic text-analysis project for EFL writing revision. *System*, 20(4) 481-492.
- Naber, D. (2003) "A Rule-Based Style and Grammar. Checker", *Diplomarbeit Technis Fakultät, Universität. Bielefeld, Germany*
- Nagata, N. (1995). An Effective Application of Natural Language Processing in Second Language Instruction. *CALICO Journal*, 13 (1), 47-67.
- Nagata, N. (2002). BANZAI: An Application of Natural Language Processing to Web based Language Learning. *CALICO Journal*, 19 (3), 583-599.
- Park, J. C., Palmer, M., and Washburn, G. 1997. An English grammarchecker as a writing aid

- for students of English as a second language. In Proc Conf. on Applied Natural Language Process.
- Truscott, J. (1999). The case for “The case against grammar correction in L2 writing classes”: A response to Ferris. *Journal of Second Language Writing*, 8(2), 111-122.
- Sjöbergh, Jonas. (2005). Chunking: an unsupervised method to find errors in text. Proc. Nodalida 2005, Joensuu, Finland
- Yang, N. D. (2004). Using My Access in EFL writing. The proceedings of 2004 International Conference and Workshop on TEFL & Applied Linguistics (pp. 550-564). Taipei, Ming Chuan University.
- Yang, J. & Akahori, K. (1999). An Evaluation of Japanese CALL Systems on the WWW. Comparing a Freely Input Approach with Multiple Selection. *Computer Assisted Language Learning*, 12 (1), 59-79.
- Yang, H. & N. Wei (2005) College English Learners’ Spoken English Corpus (COLSEC). Shanghai: Shanghai Foreign. Language Education Press.
- Yu, Y. T., & Yeh, Y. L. (2003). Computerized feedback and bilingual concordancer for EFL college students’ writing. Proceedings of 2003 International Conference on English Teaching and Learning in Republic of China (pp. 35-48). Taipei, Crane.
- Wen, Q, L. Wang & M. Liang (2005) Spoken and Written English Copus of Chinese. Learners (SWECCCL 1.0). Beijing: Foreign Language Teaching and Research Press.

Developing an English Tense and Aspect Checker Using Natural Language Processing Techniques

Zhao-ming Gao (高照明)
National Taiwan University
zmgao@ntu.edu.tw

The applications of natural language processing (NLP) techniques in language learning and assessment have drawn much attention in recent years. Several successful NLP applications have been reported, including automated essay scoring (Burstein et al. (2004)), collocation checkers (Futagi et al. (2008), Chang et al. (2008)), among others. In this talk, an NLP-based English tense and aspect checker developed by the author's lab will be introduced. The system draws on the state-of-art NLP tools and techniques in identifying the tense and aspect information encoded in a verb and the temporal expression in the sentence. If contradictions occur between the tense and aspect in a verb and the temporal expression, the system can suggest the corrected form. A prototype of the system has been developed which can identify many of the common mistakes in English tense and aspect made by Chinese learners. To evaluate the performance of the system, both learner corpus and native speaker corpus are used. The fine-grained error tagging of more than one million words in the Chinese learner English Corpus (CLEC) facilitates the objective evaluation of the system. Preliminary experiments show that about 77.46% clauses with temporal expressions which are tagged with of tense and aspect errors in the CLEC can be correctly identified by the system. However, the program also misanalyses 9.74% of clauses with temporal expressions in texts by native speakers of English. While the present system cannot handle some cases, the problems might be solved, as more detailed rules are added to the system via our interactive web-based interface. More sophisticated implementation and experiments are under way. Meanwhile, the pedagogical impact of this tool in composition will be explored in future study.

Development of an Online Concordancer for Primary School Students

Hsien-chin Liou (劉顯親), Yu-ru Chen (陳郁儒), Li-tang Yu (余立棠),
Jason S. Chang (張俊盛)

National Tsing Hua University

hcliu@mx.nthu.edu.tw
g9662559@oz.nthu.edu.tw
cdpsttarn@gmail.com
jason.jschang@gmail.com

Due to the advancement of technologies nowadays, quite a few scholars either endorsed, or investigated the use of web-based concordancing and found supportive evidence. Previous literature shows concordancing is generally helpful on learners in tertiary levels or above, but whether such a tool is helpful or even feasible for primary school students remains to be verified. In this presentation, we describe the development process of a Chinese-English concordancer designed specifically for Taiwanese primary school students with a corpus based on English textbooks. The textbooks were collected from various publishers; English sentences with their Chinese translations were made by two researchers, based on the patterns of the textbooks. The corpus includes 46730 words in 10078 sentences. Besides, the concordancer is augmented with links to online dictionaries (*Yahoo* and *Google*) in case the user needs help with word senses. It also has key words highlights and two arrangements of the search output displays of illustrated sentences. In order to investigate the effectiveness of concordancing on vocabulary learning at primary school level, we have experimented how the concordancer was infused into an elementary school English class with 22 EFL students in a typical school in Taiwan. The assessment phase involved a time-series design (Ishida, 2004) to investigate the effectiveness of concordancing on students' learning of verb phrases. Seven fifth graders were assessed with 2 pretests to identify their unknown verb phrases, followed by multi-session treatment and posttests. The findings will be reported in another presentation of this conference (Yu & Liou, 2009).

INTRODUCTION

Concordancing is a language lookup process based on large corpora data which are authentic materials about languages. Concordancing provides the learner with direct access to corpora; the process is known as “data driven learning” (Chambers, 2005). To use concordancing in the English classroom changes the roles of both the teacher and the learner. Instead of being a knowledge disseminator, the teacher becomes a facilitator. The learner is encouraged to play an active role similar to a researcher and an explorer, who does not receive rules given by the teacher but finds patterns through examples generated from a concordance program. With abundant information of language, the learner inductively discovers the pattern and exception of language use. The learner could take responsibility with the control of their learning pace, becoming an autonomous learner. A computer program that can generate concordance output is called a concordancer. In this paper, development of a Web-based concordancer for primary school students in Taiwan is described. This report provides a short literature review of concordancer-assisted English learning, and the steps in building such a program specifically designed for young learners.

Several studies have been conducted to examine the effectiveness of using concordancing to help learners acquire vocabulary knowledge (e.g. Cobb, 1997, 1999; Lee & Liou, 2003; Sun & Wang, 2003). This line of literature addresses the improvement of vocabulary growth, productive lexical ability, and collocation. To help learners to acquire as much vocabulary necessary for their academic studies as possible in a short time, Cobb (1997) designed two versions of a computer program with either a concordance function or not. 107 freshmen of a university took part in the study for 12 weeks. Each week, subjects took 5 activities, which were choosing a definition, finding words, spelling words, choosing words for new texts, and writing words for new texts. Randomly selected 11 students were examined on their improvement of vocabulary learning with regard to the effect of using concordancing. The improvement from a pretest to a posttest was significant. Besides, from the results of a questionnaire, students held positive attitudes towards using concordancing. Likewise, Cobb (1999) designed a new concordancing system with a database of freshmen reading materials, *PET 200*. Through the concordancing system, fragmented texts were generated helping students to notice new words they were encountering, and many contexts and situations where words occurred were provided for learners as well. The study supported that via direct corpus-based vocabulary learning, learners could gain lasting and transferable word knowledge.

St. John (2001) examined how one German beginner with linguistic background accomplished 10 out of 17 concordancing tasks designed by the researcher. The tasks consisted of word and phrase searches, requiring the learner to find the correspondence of native language or vice versa. From the answers to each task and feedback regarding corpus and concordancer use, the learner not only developed useful strategies of using both tools but also observed the contexts of the searching words or phrases with high motivation. Horst, Cobb, and Nicolae (2005) designed 5 vocabulary learning activities through creating interactive online database called *Compleat Lexical Tutor* (<http://www.lex Tutor.ca>). The activities included online dictionary, concordance consultation, hypertext reading, word bank quiz and cloze maker. The design met three qualifications: focusing on academic words, offering learners a range of opportunities for vocabulary rehearsal, and engaging them into deep vocabulary learning. Sun and Wang (2003) conducted an experiment concerning 81 senior high school students' collocation learning via using concordancing. The results demonstrated that the inductive group outperformed the deductive group, especially in the easy level of collocation patterns, while both groups' performances show no significant difference in the collocation patterns of the difficult level.

Chan and Liou (2005) designed five web-based practice units for a regular Freshmen English course. 32 freshmen participated in the study. The students received 3 units with the use of a bilingual concordancer (*TOTALrecall*, <http://candle.fl.nthu.edu.tw/totalrecall/totalrecall/totalrecall.aspx?funcID=1>) to consult appropriate verb collocates and to induce language patterns from a large number of examples with key words. 2 units without the incorporation of the concordancer were designed by providing semantic grid analysis and language pattern explanation. Results revealed that, overall, learners made a significant progress of collocation knowledge in the immediate posttest despite that they regressed in the delayed posttest. Rich input from the corpus and Chinese-English mutual translation in the parallel corpus may help learners to notice, read, and assimilate the expression of collocation in the two languages. The concordancing method could effectively scaffold learners' collocation learning and to raise their awareness of collocations.

While some scholars focused receptive lexical ability, other researchers are concerned with how learners can gain productive vocabulary ability by means of concordancer application. Lee and Liou (2003) examined the feasibility of integrating web concordancing

(<http://vlc.polyu.edu.hk/> in Virtual Language Centre) on receptive and productive vocabulary ability in a required senior high English curriculum. 46 second-year senior high school students were involved in the study. The results showed that the gap between the high vocabulary level group and the low level one was minimized after using concordancing, although, compared with the other two groups, the high vocabulary group still performed the best in the pretest and posttest. Besides, while the participants preferring deductive learning performed better than those preferring inductive learning in the pretest, after using concordancing, the inductive learning group outperformed the deductive learning group in the posttest. Lee and Liou mentioned that the access of the computer lab and the training for teachers and students about using concordancing were two preconditions in order to incorporate concordancing at high schools. Teacher's intervention and students' cooperation were critical.

Focusing on academic vocabulary or the transfer of the word knowledge to an authentic writing task, Kaur and Hegelheimer (2005) examined the effectiveness of using an online concordancer, *Compleat Lexical Tutor*, and an online dictionary, *Dictionary.com*, in comparison with the use of an online dictionary only in acquiring vocabulary from the Academic Word List by 18 undergraduate students. Learning vocabulary with an online concordancer and dictionary did help ESL learners transfer the word knowledge correctly in their writing task. Kaur and Hegelheimer suggested that optimum opportunities to familiarize with the concordancer and ample time to complete the activities are important. In addition, instructions for learners to realize the importance of words and how to use them productively are essential.

The studies reviewed indicate that the concordancing application does exert a positive influence on vocabulary learning and teaching. To add our understanding of students' perspectives on the use of a concordancer, Yoon and Hirvela (2004) examined students' actual corpus use, their views on the corpus activity and overall evaluation of corpus use in their academic writing instruction. Two classes with 8 intermediate students and 15 advanced students in an ESL writing course at a large American university participated in the study for ten weeks. Both classes, which were taught by the same teacher, worked with the Collins COBUILD corpus. The students reported that corpus use was helpful for learning the usage of vocabulary and phrases, and developing writing skills and confidence increase. Corpora were neither very easy nor very difficult to use. But the time-consuming data analysis of corpus was difficult to them. Both classes viewed the description of language patterns as a useful experience for L2 writing. As for the preference on corpus-based activities, the intermediate students showed higher favor than the advanced students did. Yoon and Hirvela suggested that the level of the students proficiency may be worth special consideration.

In this section, a brief review of the empirical studies concerning concordancer application on vocabulary learning evidenced that a concordancer is a useful tool for assisting learners' vocabulary size growth, receptive and productive lexical abilities, and collocation knowledge. Furthermore, learners generally help a positive attitude towards adopting the concordancing system in their vocabulary learning. However, little is known about how children, the beginning English learners, can benefit from consulting the concordancer on their vocabulary development. Further research on this issue is worth consideration.

DEVELOPMENT OF *Concordancer for Children*

For elementary school students as beginning English learners, their vocabulary knowledge and proficiency levels for comprehending the content of the available on-line concordancers, such as the *VLC* web concordancer and *TOTALRecall* (Liou, et al., 2006), were insufficient. Thus, it is crucial to develop a concordancer specifically for elementary school students.

Before developing the concordancing program, it was crucial to select texts in order to

establish a corpus suitable for elementary school students. The corpus aimed at helping beginning learners in elementary school students. We took elementary school students English textbooks as the sentence pattern models of the corpus. According to the announcement published by National Institute for Compilation and Translation in Taiwan, nine widely-used English textbooks² censored by Ministry of Education were adopted as models for the researcher and a student worker to make sentences for the corpus. The corpus had a total of 46730 words in 10078 sentences taken from the English textbooks.

For slow learners of primary schools, it could be a problem for them to understand the word sense of basic words. In order to provide scaffolding for comprehension of the sentence meaning, Chinese texts were integrated in the concordancing program (see Figure 1). They allow users to compare structures in target language (TL), English, and in their first language (L1), Chinese, and understand the differences between them (Wang, 2001). There are several options for users in *Concordancer for Children*. Users can set the program to present 10, 20, 50, or 100 sentences with searched words at one time. In addition, the concordancer was augmented with links to two web-based dictionaries, *Yahoo* and *Google*, in case the user needs help with word sense. Users could choose searched words to align to left or in the middle and sort output sentences from the shortest to the longest and vice versa. The URL (Uniform Resource Locator) of *Concordancer for Children* is <http://140.114.75.17>.

The screenshot shows the 'Concordancer for Children' web application. The search term is 'flower'. The results include:

- Word: flower
- Dictionary for 'flower'
- flower [美國][ˈfl a u ə][英國][ˈfl a u ə]
- 名詞 (noun)
- (pl. flowers)
- 花, 花卉, 開花植物
- Bees gather nectar from flowers. 蜜蜂採集花蜜。
- 精華, 精英
- They are the flower of our nation. 他們是我們民族的精英。
- 開花, 盛時
- The narcissuses were in flower. 水仙花盛開著。
- 不及物動詞 (intransitive verb)
- (flowered; flowered; flowering)
- 開花

example sentences for 'flower'		
No.	English Sentences	Chinese Sentences
1	A flower ?	一朵花
2	Is this a flower ?	這是朵花嗎
3	Look at the flower .	看這朵花
4	It's a flower for Mom.	這是要給媽媽的花
5	It's next to the flower shop.	它在花店的旁邊
6	There's a butterfly on the flower .	有隻蝴蝶在花上。

Figure 1 Concordancer for Children

For technological specifications, we used PHP to program. It is running on an Apache HTTP server set up on Windows with a Visual Foxpro database. The output is sorted based on sentence length. The technical team is similar to those in Liou, et al. (2006).

² The nine censored English textbooks are from Pearson (台灣培生), Kid Castle (吉的堡), Hess (何嘉仁), Joy (佳音), Nan I (南一), Melody (美樂蒂), Rainbow (彩虹), Kang Hsuan (康軒), and Kang Hsuan B (康軒 B 版).

CONCLUSION

In this paper, the development of a Web-based concordancer for primary school learners is described. In order to investigate the effectiveness of concordancing on vocabulary learning at primary school level, we have experimented how the concordancer was infused into an elementary school English class with 22 EFL students in a typical school in Taiwan. The assessment phase involved a time-series design (Ishida, 2004) to investigate the effectiveness of concordancing on students' learning of verb phrases. Seven fifth graders were assessed with 2 pretests to identify their unknown verb phrases. This was followed by six-session treatment, two posttests, and a delayed posttest. The findings will be reported in another presentation of this conference (Yu & Liou, 2009).

Acknowledgements: The paper is supported by grants from two national science council projects under the grant numbers of NSC96-2411-H-007-033-MY3 and NSC95-2520-S-007-002-MY3.

REFERENCES

- Chan, T. P., & Liou, H. C. (2005). Effects of web-based concordancing instruction on EFL students' learning of verb-noun collocations. *Computer Assisted Language Learning*, 18(3), 231-250.
- Chambers, A. (2005). Integrating corpus consultation in language studies [Electronic Version]. *Language Learning & Technology*, 9, 111-125. Retrieved July 20, 2008, from <http://llt.msu.edu/vol9num2/chambers/default.html>
- Cobb, T. (1997). Is there any measurable learning from hands-on concordancing? *System*, 25(3), 301-315.
- Cobb, T. (1999). Breadth and depth of lexical acquisition with hands-on concordancing. *Computer Assisted Language Learning*, 12(4), 345-360.
- Horst, M., Cobb, T., & Nicolae, I. (2005). Expanding academic vocabulary with an interactive on-line database [Electronic Version]. *Language Learning & Technology*, 9, 90-110. Retrieved July 31, 2008, from <http://llt.msu.edu/vol9num2/horst>
- Ishida, M. (2004). Effects of recasts on the acquisition of the aspectual form -te i-(ru) by learners of Japanese as a foreign language. *Language Learning*, 54(2), 331-394.
- Kaur, J., & Hegelheimer, V. (2005). ESL students' use of concordance in the transfer of academic word knowledge: an exploratory study. *Computer Assisted Language Learning*, 18(4), 287-310.
- Lee, C. Y. & Liou, H. C. (2003). A study of using web concordancing for English vocabulary learning in a Taiwanese high school context. *English Teaching & Learning*, 27(3), 35-56.
- Liou, H. C., Chang, J. S., Chen, H. J., Lin, C. C., Liaw, M. L., Gao, Z. M., Jang, J. S., Yeh, Y. L., Chuang, T. S., & You, G. N. (2006). Corpora processing and computational scaffolding for an innovative web-based English learning environment: The CANDLER project. *CALICO Journal*, 24(1), 77-95.
- St. John, E. (2001). A case for using a parallel corpus and concordancer for beginners of a foreign language. *Language Learning & Technology*, 5(3), 185-203. Retrieved April 1, 2009, from <http://llt.msu.edu/vol5num3/stjohn>
- Sun, Y.-C., & Wang, L.-Y. (2003). Concordancers in the EFL classroom: cognitive approaches and collocation difficulty. *Computer Assisted Language Learning*, 16(1), 83-94.
- Wang, L. (2001). Exploring parallel concordancing in English and Chinese [Electronic Version]. *Language Learning & Technology*, 5, 174-184. Retrieved July, 18, 2008, from <http://llt.msu.edu/vol5num3/wang>

- Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing, 13*(4), 257-283.
- Yu, L. T., & Liou, H. C. (2009). Integrating web-based concordancing in EFL primary school students' learning of verb phrases. Paper presented at CALICO 2009, March 10-14, Arizona State University. Also in 26th Conference on English Teaching and Learning in ROC with more findings.

Assessing the Learning Management System *Ecto* for EFL Professors

Christopher J. O'Brien (歐書華), Mei-hsiang Chen (陳美祥)

National Yunlin University of Science and Technology
Transworld Institute of Technology

oshuhua@yuntech.edu.tw

This workshop introduces educators in Taiwan to a free Learning Management System (LMS), *Ecto*, an online Web service for teachers and their students to which the presenters have made modest contributions. Educators in Taiwan are becoming increasingly aware of the value and the benefits of using an LMS for their classes, but little local attention has yet been given to this new service, started early in 2008, which offers a variety of benefits that others, such as *Blackboard*, *Sakai*, or *Moodle*, may lack. The workshop begins with an exposition and demonstration of the ways that *Ecto* can benefit EFL college teachers in Taiwan. The presenters will provide insights, suggestions, and stories to illustrate various ways in which *Ecto* has already been used in Taiwan's English classes. Next, participants will get a chance for hands-on interaction with some of the basic *Ecto* features, led by the presenters, including setting up classes, inviting students to join their classes, and constructing an elementary online test. A group discussion of the service will allow participants to compare notes on various other LMSs and on *Ecto*, and share their ideas on further ways that *Ecto* may grow to be more useful to educators in Taiwan. The presenters will indicate how *Ecto* can be used as the central element of an array of online teaching tools that can include more free Web services to complement the functions of *Ecto*, including *Google Docs*, *Jotform*, and others.

In this conference workshop, we will introduce Ectolearning.com's features and uses. After that, we will demonstrate how these can be used by teachers, with cases from real life educational situations. Further, we will encourage workshop participants to try the service for themselves, using the Internet-connected computer lab so that all participants can get hands-on experience using *Ecto*. We will walk the teachers in attendance through the basic steps of signing up for *Ecto* and setting up their own classes, as well as inviting learners to enroll in the classes and other higher-level activities that may not be as obvious to beginners. Other services with similar features will also be reviewed and introduced so that those in attendance may see what else is available, for clearly different teachers may have different features that matter more to them personally, and with such variety at hand another service may serve them better. A question and answer period will follow, in which all may participate.

There are various services available for teacher managing their classes and most of their features are comparable, but a majority of the popular services are commercial packages. Even *Moodle*, which is an open-source program free to download, requires users to upload their software onto the user's own server—or onto a professional *Moodle* host provider for a considerable sum. If a school's financial officers decide the system is right for the school, the school obviously pays the bills for it and makes it available to the whole school. However, many times a teacher would like a way to use the same services without requiring the whole school's cooperation or lots of money. *Ecto* is a new service—barely a year has passed since it

went online—that individual teachers and their students can use for free, though like Blackboard, Moodle, and the rest, Ecto can be used at the whole-school level for a fee. This fee is lower than the leading Class Management Software packages, and doubtless would serve quite adequately for a school system, but in this case we will focus on the individual drama class independent from other teachers and their students.

By signing up with Ectolearning.com, a teacher can create an account, set up classes, and invite his or her first-time-user students to enroll in the classes, creating their own student accounts in the process. These classes each have their own home pages at Ecto, which serve as the hub for teacher-student communication and interaction. Teachers can apprise their students of the latest homework and due dates for each class, invite comments and discussion, keep a blog and make podcasts specific to the class. Students can submit their homework to a drop box the teacher creates on the homepage. Teachers can keep track of students' attendance and both input and figure students' grades. Teachers can also send private message to individual students on Ecto which are likely to be picked up even if the students are not in the habit of logging onto Ecto daily, as a notification is also sent to the student's email box as well.

In our opinion, the most useful features that Ecto currently offers are sending messages to a whole class at once, and giving students online tests. Messaging the entire class is no more time-consuming than messaging a single pupil, and all those students enrolled in the class on Ecto get the message in their inbox on Ecto as well as a notification in their ordinary email boxes. This is far more efficient, convenient, and rapid than cutting and pasting thirty to sixty student email addresses from a database or spreadsheet into a small "TO:" box in an email program—a tedious action that need never be done when one uses Ecto. One drawback of this is that, unlike the specific-student messages, the whole-class messages seem not to be saved among the other private messages, so once it is sent, a message cannot be viewed again, at this point. The online testing feature, while currently rather simple, is still a great boon, and it will only improve. True-False, multiple choice, and single-word answer items are graded automatically (more than one single-word answer can be counted as acceptable if the teacher wishes), and longer answers can be graded individually by the teacher. Any answer's point value may be adjusted on a case-by-case basis at need. After the exams' deadline has passed, students may review their answers and see what their errors were, and also see any comments or extra information the teacher has provided them.

Another positive quality of this relatively new software-as-service is that its creators and maintainers are very responsive to user input and are frequently adding new features in response to teachers' suggestions; several of my suggestions have already been put in and more are coming. For example, soon pictures and sound will be available for use with tests, so that, for instance, listening quizzes or writing tests in which the student must describe a photograph will be possible. Also, the creators report, a teacher's past classes' pages and other material will be archivable: they will no longer show up on the teacher's home page and will only be visible and editable on demand. Further, the creators report, quizzes made for previous classes will soon be able to be duplicated for revision or reuse in the same class or in another, saving much time and frustration.

Overall, we hope in our workshop to encourage educators who rely entirely on paper-and-pencil means of record keeping to consider employing newer media for their daily teaching and record-keeping tasks, allowing them to save much time and effort and to concentrate their attention more closely on teaching. The shift in effort will also allow these teachers to have more time to get to know students more personally and to assess—and to keep records of—their individual proclivities and abilities, as well as identifying their personal educational needs. In the modern world of electronic lifestyles, a teacher who is up to date with today's trends will also be able to relate to students' expectations and changing means of

spending their free time. Further, by interacting with students online, educators may even be able to reclaim part of that free time for their own educational purposes, and engage students' attention and interest more fully.

資訊融入大專進階英文教學實務工作坊：

以「網路英文」為例

Chin-lung Wei (韋金龍)

National Chung Hsing University

clwei@dragon.nchu.edu.tw

國內大專院校開設「進階英文」課程，對強化學生英語文能力有必要而且重要，然而，「進階英文」的教學目標及選修學生的程度、動機、態度等可能都和「大一英文」大不相同，教材的選擇及教法的運用自然得作適度地調整，教學成效才不會打折扣，而值此資訊科技時代，將資訊融入「進階英文」教學是最能符合學生需求的策略。筆者抱持野人獻曝態度，以自己開設的「網路英文」為例，演示五個結合網路資源及融入教學的實作活動經驗：(1) 結合聽解/聽寫練習的精讀教學，(2) 口語意譯(paraphrasing)錄音，(3) 結合部落格(Blogs)的電子檔案學習評量(e-Portfolio)，(4) 非即時(non-real-time)文字線上討論，(5) 聽力學習日誌(learning log)作業，期能達到拋磚引玉、切磋交流、學習成長的目的。

前言

傳統上，國內大專院校的英文課程設計主要以「大一英文」為主軸，定位為通識或基礎課程，通常為必修科目，近年來，多數校系鑑於持續與規律學習是維持及提升英語能力的重要關鍵，除了建構英文學習資源中心引導學生自主自學外，也紛紛開設大二(含)以上之各年級英文或不同領域的專業英文課程，雖然科目名稱及修習規定各有不同，但除了有些學校規定為必修外，多數規定學生需先修完或免修「大一英文」才能選修，期盼經由開設這類所謂的「進階英文」課程，改善學生的英語文能力。

「進階英文」課程的增設，對強化學生英語文能力是有必要而且是重要的，然而，「進階英文」的教學目標及選修學生的程度、動機、態度等可能都和「大一英文」大不相同，教材的選擇及教法的應用自然得作適度地調整，教學成效才不會打折扣。而值此資訊科技時代，將資訊融入「進階英文」教學是最能符合學生需求的策略，筆者抱持野人獻曝態度，以自己開設的「網路英文」的教學設計為例，演示及說明結合網路資源及融入教學的策略，以期能達到切磋交流、學習成長的目的。

實作步驟

「網路英文」是門選修的通識課程，上、下學期各 2 學分，但非連續性課程，也沒有前後順序的限制，學生可以單修任一學期，也可以兩學期都修，除了有先修過或免修「大一英文」的規定外，沒有其他的擋修限制，目的是希望發揮「進階英文」的精神，引導有意願的同學持續修習以增強英語文能力。

「網路英文」課程設計的內涵主要分為兩個層面：(1)充分利用網路資源，以「道地」的(authentic)多媒體(含文字、影像、聲音等)語料作為上課教材，(2)透過網際網路

輔助教學，有效引導學生參與聽、說、讀、寫的語技整合(4-skill integration)教學活動，進而增強其口語及文字溝通的能力。本工作坊將演示的教學策略及設計共有以下五項(部分內容參見 Wei, 2007 & 2008)：(1)結合聽解及聽寫練習的精讀教學、(2)口語意譯(paraphrasing)錄音、(3)結合部落格(Blogs)的電子檔案學習評量(e-Portfolio)、(4)非即時(non-real-time)文字線上討論、(5)聽力練習學習日誌(learning log)作業。各個教學設計分別依教學目標、語言技能、教學流程(含課前準備、課中活動、及課後作為)不同項目條列說明，這些教學活動都需適度結合使用網路多媒體設備，學生亦需在家裡或宿舍自備多媒體網路電腦配合教學設計，完成交代之作業。

一、結合聽解及聽寫練習的精讀教學活動

教學目標	引導學生結合聽解及聽寫練習進行網路文章精讀，培養並增強學生的聽解及閱讀理解能力。
語言技能	聽、讀。
教學流程	
課前	設計教材並印製講義(參見 Appendix A)。
課中	<ol style="list-style-type: none"> 1. 播放課文語音檔 1 遍，學生完成聽解練習(pre-test)。 2. 播放課文語音檔 3 遍，學生完成課文聽寫克漏字練習。 3. 再播放課文語音檔 1 遍，學生再作 1 次聽解練習(post-test)，比對前、後測的成績。 4. 教師逐段講解章節大意、艱澀字詞、疑難解答。 5. 學生再聽課文語音檔 1 遍(不看講義)。 6. 學生配對或分組口語討論相關問題。
課後	學生參與線上討論活動(參見第四個教學活動)。

二、口語意譯(paraphrasing)錄音活動

教學目標	引導學生利用數位錄音筆，進行口語心得或摘要報告練習，培養並增強學生的口語表達能力。
語言技能	口語表達及溝通
教學流程	
課前	<p>(家庭作業)</p> <p>複習上過的單元，並撰寫內容摘要或學習心得</p>
課中	<ol style="list-style-type: none"> 1. 學生配對口語練習心得或摘要報告。 2. 學生輪流上台口述心得或摘要並以錄音筆錄音。
課後	上傳錄音檔、評分並公告成績。

三、結合部落格(Blogs)的電子檔案學習評量(e-Portfolio)活動

教學目標	引導學生利用部落格，建構個人學習檔案，培養並增強學生的探索、分析、整合、及英文口語與文字表達溝通能力。
語言技能	讀、寫。
教學流程	
課前	(家庭作業) 1. 學生自行決定主題。 2. 申請免費部落格。 3. 蒐集相關影音、文字資料。 4. 撰寫與主題相關的介紹或心得(每月至少 1-2 則)。 5. 造訪同學的部落格並適度撰寫回應文章。
課中	1. 學生依學號順序上台分享成果。 2. 其他同學台下聆聽、評鑑(評量表參見 Appendix B)、老師評分。
課後	收回評鑑表並檢視學生出席及填寫內容。

四、非即時(non-real-time)文字線上討論

教學目標	引導學生利用網路教學平台，進行文字溝通，培養並增強學生的英文文字表達溝通能力。
語言技能	聽、讀。
教學流程	
課前	無
課中	無
課後	(家庭作業) 1. 教師於課程單元結束後，在課目網路教學平台討論區公佈 1-2 個問題。 2. 學生在 2 週內進入教學平台至少 1 次，以文字輸入個人看法及意見。 3. 鼓勵學生回應同學的看法及意見。 4. 評分、登錄並公告成績。

五、聽力練習學習日誌(learning log)

教學目標	引導學生利用現成網站之聽力訓練教材及網路教學平台，自行定期做聽力練習，培養並增強學生的英語聽解能力。
語言技能	聽、寫。
教學流程	
課前	(家庭作業) 1. 學生每周連線 http://www.esl-lab.com/ 聽力訓練網站的“Listening Quizzes for Academic Purposes”單元。 2. 自行挑選至少一單元，進行聽解練習。 3. 依規定格式編寫學習日誌(Appendix C)。 4. 上傳至網路教學平台。
課中	隨機展示作品分享、切磋。
課後	評分、登錄並公告成績。

結語

國內各級學校英語教學長期以來受到教科書及教學進度的箝制，加上考試領導教學的影響，文法翻譯(grammar translation)或聽講(audio-lingual)教學法、教師為中心的「講光抄」教學模式總成為英語教室內常見的主要活動，有心變革的老師，雖然絞盡腦汁、苦心經營，採行一些互動溝通、以學生為中心的教學活動，如分組/配對討論練習、專案演示等，但又多受到班級人數太多或學生語文能力參差不齊的限制，效果不如預期又走回老路，因循苟且、習以為常，即便進入大學，雖然考試及分數的壓力不如中學階段，多數老師與學生還是比較習慣這種「老師講、學生聽」的被動教學模式，學生上課期間主動使用及操弄(manipulate)英語的機會相對較少，以致學習動機低落，教學成效也不是很理想，大學英語教育實有必要進行改革。

時序進入網路科技時代，國人的電腦網路素養已不可同日而語，大專英語教師應與時俱進、順勢而為，教學設計可適度結合網路多媒體，以活化課室教學氣氛，進而提升英語教學成效。藉由網路電腦的結合應用，英語教師要引導學生活用語文知能已非難事，網路多媒體已成為落實「溝通式英語教學」值得期待的利器。本工作坊演示說明的五個教學技巧植基於溝通式語言教學理念，強調學生為中心、興趣為考量、能力為導向、聽說讀寫整合、活潑多樣、激勵參與、誘導持續練習、發揮教師專業自主精神、提昇教學管理成效，旨在供英語教學同業參考，期盼能發揮拋磚引玉之效。

參考文獻

- Wei, C. L. (2008). 資訊融入大一英文教學實務工作坊. The 25th Conference of English Teaching and Learning in R.O.C. Chayi, Taiwan: National Chung Cheng University.
- Wei, C. L. (2007). e- TESOL: A Comprehensive Guide to Internet-Based TESL/TEFL. Taipei, Taiwan: Kuan Tang International Publications.

Appendix A

A Sample Lesson

Lesson One
Foreign Students Learn ABCs of USA

Warm-Up

(For the pre-test, listen carefully and answer the comprehension questions.)

Pre-Test		Post-Test
()	1. (A) A math course. (B) An electronic engineering course. (C) A culture training course. (D) An international relationship course.	()
()	2. (A) More than 50%. (B) Less than 50%. (C) All of them. (D) Less than 30%.	()
()	3. (A) One semester. (B) Ten weeks. (C) One year. (D) One quarter.	()
()	4. (A) A bad thing. (B) Impolite. (C) Unacceptable. (D) Absolutely necessary.	()
()	5. (A) American culture is diverse. (B) Japanese would be more modest. (C) Chinese would be more serious. (D) Brazilian students might open up a conversation with inviting you to come to their house.	()

Reading while Listening

I. Listening Cloze

Performance Check → _____/20

Foreign Students Learn ABCs of USA

Students from every corner of the world flock to American universities. They may be at ease with the most complex mathematical theories, but are often by Americans. So the University of California at Berkeley initiated a class for foreign students, who _____ up 60 percent of all post-doctorates on campus. Called America 101, it's a basic _____ course explaining everyday life in the United States.

"Americans (they) have to do everything by themselves." "You can get along here easier by yourself than any other countries." "They have this, you know, smiling face." "Well, to me, most of these men are not genuine. It's just something plastic." The _____ students in the America 101 class are perplexed.

A German asks what it meant when an American friend invited him to dinner, saying, 'Let's go pig out on buffalo wings.' Others want to know when to use a _____, a hug or a kiss on the cheek. "Culture shock means you come to a new country and everything is _____." English study institute director, Bonu Ghosh, teaches this 10-week course, intended to ease foreign students' feelings of _____ and alienation. She explains that if culture shock gets the best of them, the result is "an _____ roller coaster." "Everybody knows what's a roller coaster?" The students answer with silence. Ghosh can't tell if they've her or not. But she forges on.

"Is it a bad thing; is it impolite to say no?" "No! It's necessary. So if I say, 'would you like to have lunch with me today?' and you say, 'mmmm...!' what are you really saying? No. In an _____ way. In this culture, American English, it is very important to say no."

Students worldwide, from Israel to Kuwait to Uganda, agree that Americans' comfort with a cool 'no' is typical of our abrupt and direct style. But such directness makes many _____. Sâfa Gasser, an electrical engineer from Egypt [admits], "Sometimes I just get shocked when the answer is like, strict and, 'no.' This is typically American, you know. They don't get _____."

America 101 course topics include restaurant protocol, telephone styles, and the how-to of a firm handshake. Instructor Ghosh says cultural training is as crucial as the students' _____ training. Today, she focuses on the American job interview. A student from Japan, she says, would be _____. "It would be bragging and it would be completely unacceptable to sell yourself. So they say, 'ah, I try hard.' And to an American _____ this is completely inadequate."

Or, a Brazilian student might open the interview with, 'Come to my house.' "And this doesn't mean you're invited to come to their house. It just simply means I'm communicating with you. It's kind of like saying, 'how's it going?'" And a Chinese student might _____ entirely on how serious he is about his work. "Maybe another candidate will come across as easier to work with because they're casual, they're friendly, and they're _____."

But before getting to the job interview, the foreign students have to deal with Americans on campus and in the community. And for many, *that* is the most significant _____: who are the Americans? Egyptian Safa Gasser says much of the U.S. population -- filled with Asians, Latinos, blacks, Middle-Easterners -- looks foreign to her. "I don't know. The figures of Americans to me are the white, blond people that I see on TV. Then, where are the Americans? I can't find a lot of Americans!" For Gasser, the _____ is that the America she's trying to learn about is as diverse and multi-cultured as the students in America 101.

--Adopted from <http://www.voanews.com/>

II. Comprehension Check (for the post-test only)

- 1. Turn to the first page of the lesson.**
- 2. Listen to the talk and take notes once again.**
- 3. Answer the questions for the post-test.**
- 4. Check the answer and compare the results of the pre- and post-tests.**

Post-Reading

I. Comprehension questions:

1. What is America 101?
2. Who initiated the course?
3. How many post-doctorates on campus took the course?
4. How long does this course take?
5. What is culture shock?
6. What is the main idea of the passage?

II. Post-Class Activities/Exercises

1. Use a search engine (such as Google, Yahoo!, etc.) to look for an article on similar theme for outside reading/listening activities.
2. Send an email to the instructor/a classmate/a friend to share your reflections on the topic.
3. Log on the Blog/website (<http://webenglish.wikispaces.com/>) created for this course to share your reflections on this lesson.

Appendix B

A Sample Evaluation Sheet

Guidelines for Evaluation: 5: 90-100 4: 80-89 3: 70-79 2: 60-69 1: Fail

Presenter	Student #	Date	Title	Evaluation	Notes
James Wang	9408001	09/27/2007	Tom Cruise	5 ④ 3 2 1	1. Good Pronunciation! 2. Clear Presentation!
				5 4 3 2 1	
				5 4 3 2 1	
				5 4 3 2 1	
				5 4 3 2 1	

Appendix C

A Sample Learning Log

Name: Tony Chen

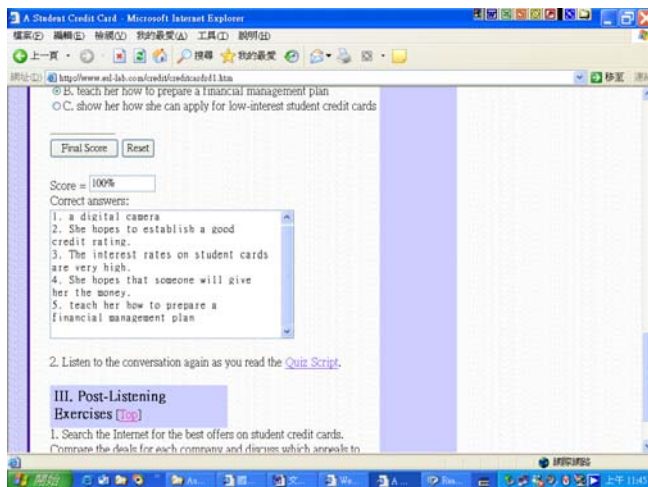
Student ID: 45678911

Date: 09/14/2007

Source: Randall's ESL Cyber Listening Lab

Topic: Student Credit Cards

Level: Medium:



Reflection:

I tried the second time to get all the items correct. It's been a long time since I took a listening test for school admission last year. I almost forgot how to focus my attention on the dialogue when it started. I was not really ready for the talk in the first place because I thought the medium level might be a piece of cake. But I was wrong. Both speakers spoke so fast.

I followed the sequence of the items while listening the second time. I found I could follow the content of the dialogue more. And I read the notes for pre-listening exercises to get myself familiar with the topic. I did well on the second try.

Notes on the topic:

Words and expressions learned: dough (money), credit rating

I don't think banks should issue credit cards to those who can not provide financial statements of their income. Students are likely to spend more than they can afford and of course create endless debt for themselves and their parents.

Beyond Classroom: Exploring Activity Theory to Informal English Speaking Sessions

Ching-fen Chang (張靜芬), Yu-ting Lin (林郁婷), Ming-chien Lee (李明倩),
Yi-wen Liu (劉宜汶)

Institute of TESOL, National Chiao Tung University

cfchang@mail.nctu.edu.tw
dantee12@yahoo.com.tw
tifflee1113@yahoo.com.tw
smilingxd@msn.com

To meet the goal of improving students' English proficiency, many universities have designed a series of non-credit, out-of-classroom English learning activities. Research in second language (L2) or English-as-a-foreign-language (EFL) research has mostly been conducted in formal and credited language learning context (e.g., classroom-based learning). Non-credited language learning activities, however, have drawn little attention to issues generated in the specific contexts. Drawing from Activity theory (Engeström, 1987, 1999) as a theoretical framework, the presentation aimed to report four case studies to map out the complexities of the factors that influenced two teachers and two students' teaching and learning in a series of informal, out-of-classroom English speaking sessions held by a public university in northern Taiwan

The case studies adopted a qualitative inquiry. Two major data sources were collected: field notes taken from intensive observations and interviews. Data were analyzed based on the factors underlying individuals' activity systems (Engeström, 1987, 1999). The findings showed that the two English teachers' teaching was highly shaped by the unique non-credited nature and contexts of English sessions, which were different from their regular credited English courses. The two student participants' intensive participation in the non-credited English practice sessions highly resulted from their strong motives, their cultural background, and the contextual factors situated in the community. The study findings reveal that activity theory serves as a theoretical and analytical framework to explain and interpret the complex relationship of individuals' actions and other factors in particular teaching and learning contexts

INTRODUCTION

In the past few years, to meet the goals of enhancing students' English proficiency set by the Taiwan Ministry of Education (MOE), many universities in Taiwan have been holding a wide range of English learning activities to help students gain rich language practice activities out of regular credited English courses (Chiang, 2007). Compared to regular credited English courses which set clear learning goals and objects, last for a certain amount of time (e.g., a semester, an academic year), and are usually taught by a same teacher, these non-credit English learning activities, led by English teachers or graduate students, are designed to offer additional practice opportunities in a flexible and relax atmosphere. Students, with full autonomy, may decide what and when they attend these non-credit informal learning activities.

Several studies undertaken in Taiwan have explored what types of out-of-classroom

activities have been offered by school institutes to facilitate out-of-classroom English learning opportunities for students (Chiang, 2007) or what kinds of out-of-class, leisure English learning activities students choose for self-study (Shen, et al, 2005). These studies have provided general information about the activities institutes have offered or students locate by themselves. However, little is known how individuals teach and learn in these non-credit learning activities or how the contexts shape their performance.

Activity theory

From a sociocultural perspective, human learning is a situated and mediated process interweaving between individual mind and social milieu (Lantolf, 2000). When individuals are situated in a teaching or learning context which is quite different from a regular context they usually practice, it is valuable to explore what changes have been generated from the process and how the changes are transformed into individuals' social practice. Activity theory, a key theory rooted from Vygotsky's sociocultural theory, seeks to explain human behavior by examining the dialectic relationship between individuals and the social environment. In the Activity theory model proposed by Engeström (1987; 1999), as shown in Figure 1, the complexities of individuals social practice are mapped out with the elements from the objects set by individual agency, mediated artifacts, the rules, the community, and the division of labor.

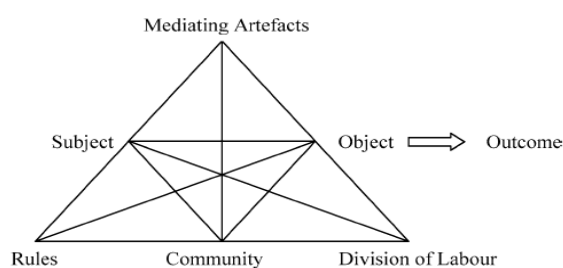


Figure 1. The expanded activity system (Engeström, 1987, 1999)

According to Nelson, et al (2001) and Jurdak (2006), subject is the human agent undertaking the activity; object represents the targeted point the subject is working towards, the process of which is mediated by different artifacts, tangible and intangible alike, to finally reach the desired outcome. Subjects who share the same object belong to one distinct community, and the rules refer to the norms, conventions, or regulations that regulate the actions and interactions within that particular community. The division of labor refers to how tasks and power status are divided among members of that community. The components described above are by no means fixed, but undergo constant changes as a result of the interaction among these elements.

Activity theory in second language teaching and learning

In the past two decades, the adoption of Activity Theory in L2 research is gradually increasing. Researchers in the field of second or foreign language learning have started to emphasize on the influence of context on teaching and learning from a sociocultural perspective. Activity theory, thus, is adopted as an analytical framework to address the complexity of language learning process (e.g., Nelson et al., 2001; Lantolf et al., 2002; Lantolf & Thorn, 2005; Storch, 2004; Jurdak, 2006; Haneda, 2007; Norton, 2007; Gao, 2008). Storch (2004), for example, investigated the dynamic interactions among learners through the lens of activity theory. He found various types of students' interaction patterns. Such difference resulted from learners' own interpretation of the situation, the goals they set, and the role they played. The results reveal that individuals undergo different activity systems depending on their learning experiences and their own activity systems embedded in a specific context.

Apart from explaining learners' behaviors, activity theory also contributes to language teaching. After a careful examination of factors underlying the activity systems, Jonassen and Rohrer-Murph (1999) argued that activity theory was with the potential to yield different perspectives for analyzing learning process and outcomes for the purpose of designing instruction. In addition, Grossman, Smagorinsky, and Valencia (1999) elaborated the key concepts in activity theory and proposed that activity theory provided a useful framework for studying the conceptual development of teachers. Therefore, the theory is beneficial for understanding the process of teaching, particularly in illuminating how teachers choose conceptual and pedagogical tool to conduct teaching.

These studies adopting activity theory as the theoretical framework highlight the impact of social, institutional, and historical factors upon individuals' choice of actions. Thus, activity theory provides a useful framework for analyzing teachers and students' behaviors.

The current studies

Drawing on the heuristics and tools provided by Activity theory as the theoretical framework, this presentation reported four cases which aimed to gain an in-depth understanding of how two English teachers hosted a series of non-credited English activities and how two students voluntarily and intensively participated in different *Table* sessions or attended two types of informal learning sessions—the *Table* sessions and *GEPT* speaking sessions. Two research questions were generated from the case studies.

1. How do English teachers and students operate in a series of non-credit English learning sessions?
2. What underlying factors affect their operation in the series of non-credit English learning contexts?

METHOD

The non-credited informal English learning sessions

Case study methodology with qualitative approaches were adopted to explore how two teachers and two students taught and learned in a series of informal English teaching and learning activities offered by a public university in northern Taiwan in Fall, 2008. To create rich English learning opportunities for campus-wide students out of regular classrooms, the language center of the university where the studies were conducted designed a series of out-of-classroom English learning sessions, the *English Tables* and *GEPT* sessions. The *English Table* sessions were daily-based, one-hour activities. They were held in two fixed time slots: noon (12:30-13:20) and afternoon (16:40-17:30), on each weekday in the Self-Study Center located in the library of the university. The sessions were led by the eight full-time contact-based English lecturers; each teacher was in charge of one noon-session and one afternoon-session per week. Except the mid-term and final weeks in the semester, there were in total 288 *English Table* sessions held in the center and 18 sessions per week.

Due to the nature of non-credit and out-of-classroom learning activities, the language center did not specify any teaching goals or emphasize certain English skills for *English Table* sessions. The institutional flexibility gave the teachers a great power to decide the topic, content, materials, activities, and target language skills that they expected to carry out in the *English Tables*. To ensure the teaching and learning efficacy of the *Table* sessions within one hour, only 10 students were allowed to participate in each *Table* session. Those who expected to join specific *English Table* sessions should sign them up two weeks before the sessions.

Another series of out-of-classroom English learning activities was *GEPT* sessions. They were held once a week for two hours over 10 weeks, led by graduate students of the TESOL program in the university. The sessions were open to 20 students who were interested in preparation for the GEPT spoken test. Unlike the *English Tables*, students registered in the *GEPT* sessions were required to regularly join the 10-week sessions.

Participants

Before we targeted our participants, we had observed the *English Table* sessions twice a week for a month. From the observation, we found several interesting phenomena: First, each teacher held their unique teaching beliefs and styles running their own sessions. Second, few students enthusiastically participated in different *English Table* sessions every week. Third, few students not only joined the *English Table* sessions but also attended the *GEPT* sessions. To understand how these teachers and students teach and learn in these English learning events, we decided to recruit the lecturers hosting the activities, the students who regularly and intensively participated in the activities, and the students who participated in both *English Tables* sessions and *GEPT* sessions.

In the recruitment process, among the eight lecturers, two lecturers voluntarily participated in the case studies, including a female Taiwanese lecturer, Teacher A, and a male native-speaking lecturer of English, Teacher B. They were both full-time lecturers in the language center of this university. Aside the 10 required teaching hours per week, each contract-based lecturer was required to host two one-hour, out-of-classroom, *Tables* sessions and three one-hour writing consultation sessions per week.

Two student participants were targeted and recruited because they either enthusiastically participated in the *English Table* Sessions or regularly participated in both *English Table* sessions and the *GEPT* sessions. Tim, a master's student majoring in Electronic Communication, signed up almost every *English Table* session during the data collection semester. Compared to most students who participated in one or two *Table sessions* each week, his high and stable participation in the English learning activities made him an outstanding case to be explored for his activity systems in these out-of-classroom English learning activities.

Emmy was the other student participant who was an undergraduate majoring in Civil Engineering at her senior year in this university. She was recruited because of her regular participation in one *English Table* session and one informal English learning session, *GEPT* sessions¹, every week. Unlike most students who only participated in only one out-of-classroom English practice events, Emmy's participation in both out-of-classroom English learning sessions for different purposes made her a special case in this study.

Data collection and analysis

The data were collected from four-week intensive field observations in each participant's participation and interviews. A summary of data collection is shown in Table 1.

Table 1. *The participants' demographic information and the summary of data collection*

participants	gender	age	participation activities	field observation#	Interviews#
Teacher A	female	late 20s	<i>English Tables</i>	6	7
Teacher B	male	late 20s	<i>English Tables</i>	5	3
Tim	male	23	<i>English Tables</i>	8	7
Emmy	female	22	<i>English Tables</i>	4	5
			<i>GEPT classes</i>	4	5

As shown in Table 1, we observed the sessions that each participant attended. Each participant received three to seven interviews. We took field notes in each observation. Interviews were taped and later transcribed. Based on the model of Activity Theory developed by Engeström (1987, 1999), the six dimensions in the model (subject, object, tools, rules, community, and division of labor) served as the analytical framework to identify the participants' activity systems in the out-of-classroom English teaching and learning activities.

RESULTS AND DISCUSSION

Case One: Teacher A, a non-native English speaker

Teacher A, holding a master's degree of TESOL, was a non-native English speaker at her late 20s. She had been a full-time contract-based teacher in the language center of this university since 2007. In addition to 10-hour regular credit-based English courses, she hosted two *English Table* sessions: a noon session (12:30~1:20) on Tuesdays and an afternoon session (4:40~5:30) on Wednesdays. To bring students different learning experience, she intentionally adopted different themes in the two sessions each week: In the Tuesday-sessions, a popular American episode, *Friends*, was adapted as the session theme. In the Wednesday sessions, she used English songs or designed simple tasks as the themes. She explained that in so doing, students could select and experience different types of *Table* sessions.

Teacher A's operation in the *English Table* sessions

At the beginning of each session, Teacher A greeted the students and asked them to briefly introduce themselves. After the opening, Teacher A announced the topic, distributed handouts, and explained the content by asking the students to take turns and read the lines. Since the classroom where the sessions took place was equipped with a networked computer and a projector, she always played video clips either from the Internet or from a CD-ROM along with the themes in her *Table* sessions. While students were watching the videos, Teacher A constantly paused the video, checked students' understanding, and explained the plots where students had problem with. After watching the video, students usually engaged in 10-minute group discussion, while Teacher A approached each group to observe and check if they met any difficulties in discussion. While interacting with students, she often tried to be humorous in the process. She confirmed her humorous attitude in the process in a later interview, "Being humorous could make my *English Tables* more relaxing occasions. Thus students would feel more comfortable to talk in English, which is not their native language after all." (interview #7, Jan. 06, 2009)

Case Two: Teacher B, a native English speaker

Teacher B was a native English speaker who has been in Taiwan as an English teacher for four and half years. Although majoring in Engineering, Teacher B held a TESOL certificate before formally starting his career as a teacher in Taiwan. According to Teacher B, his English teaching styles were highly influenced by the courses offered by the certificate program. Both of his *Table* sessions were in the afternoons (4:40~5:30) and were conducted in a sofa area in the same self-study center. Different from a regular classroom, the sofa area was a relaxing corner in the center set with a table and 4 sets of two-seater couches. No computers or any technical equipment were available in the setting. Teacher B reflected in the first interview that he was happy to host his sessions in the sofa area in which people felt cozy and relaxed. Thus, the atmosphere lowered students' anxiety to speak out in a second language.

Teacher B's operation in the *English Table* sessions

Different from Teacher A's style, Teacher B did not give students the handout while introducing the topic. Instead, he held the handouts and explained the content by himself. Then, he distributed handouts to the students and asked them to read through the texts quietly for several minutes. According to Teacher B, to push students to concentrate on his explanation and train their listening comprehension, he did not distribute the handouts until he finished his explanation. Furthermore, he claimed that because he emphasized on students' reading comprehension rather than their pronunciation, he did not ask students to read aloud. A more significant difference from Teacher A's sessions was that Teacher B assigned a twenty- to thirty-minute group discussion, which dominated most of the time in a *Table* session. While students discussed in their groups, Teacher B did not purposely approach each group or interact with them only when a group became quiet. As he asserted, "The more

teachers talk, the less students talk.” (interview #3, Dec.29, 2008). Thus, he chose to talk less so that there would be more time for students to talk.

Factors influencing Teacher A and B’s operation in the *Table* sessions

As Teacher A and Teacher B’s cases revealed, how they held the *Table* sessions were influenced by their agency. It is found that although the institutional policy did not set any particular goals for the *Table* sessions except providing English learning opportunities, the way Teacher A and Teacher B held the *Table* sessions was interweaved by their beliefs, mediated artifacts, rules, and the division of labors in the *Table* sessions. While holding the sessions, both teachers’ beliefs and objects set for the *Table* sessions influenced their selection of teaching materials and strategies. Compared to regular semester-long, credited courses, Teacher A set the object of the “one-shot” *Table* sessions as a friendly but rewarding learning experience. As she explained, “I think since *English Tables* are free language learning sessions... In my sessions I hope students can make new friends, share ideas, and gain new information in addition to learning the language.”(Interview 7# Jan, 9, 2009). Building upon the object, she emphasized on idea exchange and communication in her sessions. According to her, session topics and students’ engagement determined if an *English Table* session was successful. She chose interesting and various topics for different sessions. To enhance students’ engagement, she tried to invite every student to speak up in the whole or small group discussion. As for teaching strategies, although neither evaluation was involved nor specific language objects were set in *Table* sessions, Teacher A expected students to learn new knowledge besides practicing speaking skills. Therefore, she constantly helped students pay attention to the language by recasting their errors and correcting students’ pronunciation.

Same as Teacher A, Teacher B perceived *English Tables* as informal English speaking practice instead of a course in the school curriculum, as he stated, “I have English sessions and *English Tables*. I don’t see an *English Table* as a class. I see it as an activity for students to enjoy, to speak, and to practice what they learned after classes...I want them to feel like it’s a comfortable opportunity to practice English.” (Interview #3 Dec. 29, 2008) Aiming to help students use the target language in a least stressful context and build up their confidence, Teacher B claimed that he focused on fluency rather than accuracy. He seldom corrected students’ errors nor did he actively participate in students’ discussion to give students as many opportunities to talk as possible in his sessions. Furthermore, Teacher B’s assumptions of teaching were also influenced by his personal and cultural background. He mentioned in the third interview that language education in his country put equal emphasis on speaking, listening, reading and writing while the situation seemed different in Taiwan. He, thus, tried to balance students’ learning in terms of the four language skills in particular in his *Table* sessions (Dec. 29, 2008).

Similar to Teacher A’s case, rules and division of labors were the other two influential factors in his *Table* sessions. Less constrained by rules observed in regular credited course, the two teachers indicated that they enjoyed more flexibility to create a relaxing environment and to select materials. Teacher A reflected that her teaching became more creative and flexible in the *Table* sessions than in her regular credited classes. She could choose interesting topics, add multimedia materials, and vary different tasks to make her sessions interesting. Similarly, time limitation and unfixed students also affected the teachers’ instruction. Furthermore, since there was no specific rules for assessment, Teacher B changed his teaching procedure (i.e., teacher introduction, teacher explanation, and student group discussions) in credited courses to a focus on student group discussion only in the *Table* sessions because explanation to linguistic forms were not the focus of the learning outcome. As for the division of labor, both of the two teachers confirmed that the power relationship between teachers and students was more equal in the informal learning sessions than in traditional classroom. Thus, they positioned themselves as a facilitator rather than as a teacher,

Teacher B even claimed that he totally gave the power to his students.

Case Three: Tim, a graduate student

Tim was a 23 year-old graduate student in the master's program of Electronic Communication. He had been regularly attending *English Tables* throughout the school semester, with only a couple of absences in between. Tim believed that excellent English speaking skills was one of the prerequisite skills to work in an international company, a goal he set for his career after graduation. However, he was also aware that English speaking was his weakest English competence in terms of the four language skills. Furthermore, in the data collection semester, since he did not register any credited English classes, he decided to intensively attend the *Table* sessions to get speaking practice. With such a strong motive, he even skipped regular credited classes in his master's program. He verified his strong motive for the *Table* attendance.

Because for other courses I might be able to understand it on my own even without a teacher present, but for English you've got to be face-to-face to listen or speak. For those [courses at his master's program] you probably can grasp the meanings as long as you do the reading, but for this [speaking] you can't just learn by sitting in front of the desk. (interview # 6, Dec. 25, 2008)

Tim's Participation in the *Table* sessions

In the two-week intensive data-collection time, Tim reserved for and attended eight out of 20 sessions offered by the center at the same time, the choice of which was based not on his preferences for the hosting teachers, but on his available time instead. These sessions that Tim attended were led by five different teachers, including one native English speaker and four Taiwanese teachers.

Although jumping from one *Table* session to another, Tim seemed to exhibit certain similar characteristics. For example, he persisted speaking to his partners in English throughout the sessions. When other students jotted down notes as they did in their regular classes, Tim never brought any pens with him nor did he jot down notes. He claimed that he preferred to remember what teachers illustrated, a choice of actions he had been doing throughout his school years. The two distinct behaviors that Tim performed in the *Table* sessions stood as a sharp contrast from the rest of the participating students.

Despite Tim's reported enthusiasm for the *Table* sessions, he did not always act as an active learner: At certain times he actively initiated questions and leading group discussion but in other times he was a passive listener, responding only to partners' questions rather than initiating one. Tim explained that his participation depended on how active other group members were in a small group or whole group discussion. He explained his distinct behaviors.

Actually it depends, so if someone speaks first, then I would just let him go on. But like today it was [sharing of] a group discussion, so if he didn't finish the entire thing I would add up more details. And normally if there's just one question and one answer only, I would let others do the job.... So basically I would give the priorities to others. (interview #5, Dec. 15. 2008)

However, he was cautious when speaking up in English in the two sessions hosted by a native English speaking teacher. He confirmed that facing an English speaker, he tended to keep silent and think for a few seconds before uttering any words, whereas in the face of Taiwanese teachers, he would directly speak out what he planned to express.

Factors influencing Tim's participation in the *Table* sessions

Examining Tim's case through the lens of activity theory, it is found that Tim's

personal agency played an important role in his activity system, making him a particular case at the *English Table* sessions. To begin with, Tim showed strong motivation as he would attend the *Table* sessions almost on a daily basis. Such regular and intensive attendance appears to result from the dynamic interrelationships among his agency, his object, and his goal. Driven by his ultimate goal to work in international firms, Tim was motivated to operate towards his object, in this case, practicing his English speaking ability to boost the weakest part of his English proficiency. Such factors in turn prompted him to utilize the *Table* sessions to practice his object of enhancing speaking, helping him to move further towards his desired goal of getting a job in an international company. Furthermore, Tim's object along with his goal altogether also contributed to the exhibition of his powerful agency, as he would attend *Tables* sessions even to the extent of skipping regular classes at his master's program.

There were also instances where Tim's agency shaped his participation at *English Tables*. Upon closer observation of Tim's behaviors, his choices of actions at *English Tables* were the outcomes of his own belief systems, which can be traced back to his earlier learning histories. For example, Tim's strong motivation in cutting regular classes but a distinct behavior of never taking any notes as everyone else did was based on the conclusions reached from his previous learning experiences. Additionally, cultural factors also shaped Tim's personal beliefs. When deciding to initiate a discussion, Tim would first allow others to take initiative steps. According to him, allowing others to take the first step could maintain the harmony among group members. Such tendency seems to be attributed to the Chinese culture, which, known for its collective orientation, highly values the harmonious relationship with one another. In this sense, Chinese are situated to be prioritizing the rapport with others. It thus becomes understandable why Tim tended to put others before himself during group interactions.

Through the lens of activity theory, Tim's case at the *Table* sessions proved that subjects, acting as agents of their own, would construct their operation and beliefs as a result of interactions among their objects, goals, individual histories, and cultural influences.

Case Four: Emmy, an undergraduate at her senior year

Emmy, a 22 year-old undergraduate majoring in Civil Engineering, had studied English since junior high school. In the data collection semester, she had passed the first stage of GEPT, at the high-intermediate level. According to Emmy, speaking competence was an important indicator to assess English proficiency. Therefore, she intended to spend extra time practicing her English speaking skills under the circumstance of having more available time as a senior. Furthermore, with the attempt to pass the GEPT high-intermediate spoken test which she had failed twice, Emmy decided to attend the two series of non-credited, out-of-class English learning events: *English Tables* and *GEPT* speaking class.

Emmy's participation in *English Table* and *GEPT* speaking sessions

Emmy showed a laid-back attitude toward her participation in the two types of non-credited sessions in the belief that English speaking practice should be in a relaxing and non-threatening atmosphere. Perceiving that both *Table* and *GEPT* sessions were optional and non-credited learning activities, she never prepared for the sessions nor did she review the learning materials. She set clear goals to pass GEPT spoken test that she had failed twice and tried to gain English speaking practice while she did not take any credited English course during the data collection semester. However, she was usually quiet and rarely voluntarily responded to the instructor's questions. According to her in the fifth interview, her being in a low profile simply followed a cultural norm of not standing out in a group or in a class. Meanwhile, due to different contexts and community formed in each activity, Emmy behaved differently in the two sessions. For example, since she had a sense of security working with acquainted peers in the *GEPT* sessions, she was more active in group discussions and chatted

with her partners even after completing required tasks. Comparatively, since she usually worked with different group members in “one-shot” *Table* sessions, she seldom interacted with others after finishing a required speaking task.

The *Table* sessions were created as an English only environment. Emmy naturally used English throughout a whole session whereas she code-switched to Chinese from time to time in *GEPT* sessions where the instructor did not create an “English only” environment nor did she forbade students to use Chinese. As Emmy verified her code-switching in the *GEPT* sessions, “The teacher in *GEPT* class uses Chinese to elaborate certain terms. Also, since she didn’t say anything when we spoke Chinese, speaking Chinese becomes an acceptable behavior.” (interview #1, Dec. 1, 2008)

Jumping from one type of non-credited activities to another during the semester, Emmy concluded that in the future she would rather participate in activities where she could not only practice her English but also earn credits to make her efforts and attendance pay off.

Factors influencing Emmy’s participation in both *Table* and *GEPT* sessions

Examining Emmy’s case through the lens of activity theory, similar to Tim, her personal agency played an important role in her activity system while she participated in the two types of non-credit English practice activities. Her strong motive to pass the *GEPT* spoken test led her to decide joining both *Table* and *GEPT* sessions to gain plenty opportunities to practice speaking. Her full attendance in the *GEPT* sessions and intensive participation in the *Table* sessions proved the strong connection between her agency and the outcome.

Held by two types of instructors (English teachers versus graduate students), the *Table* and *GEPT* sessions created two drastically different learning contexts where the *Table* sessions were “English only” environment but Chinese was allowed in the *GEPT* sessions when the instructor and students encountered difficulty in English. The language rules which were not explicitly declared by the instructors but grounded in the two types of contexts resulted in Emmy’s different attitudes toward the two learning contexts. Emmy appeared to feel more relaxed in the *GEPT* sessions than in the *Table* sessions. Her comfort in the *GEPT* sessions was also attributed to the division of labor in the *GEPT* sessions. Emmy felt safer and more comfortable working with the same group of peers in the *GEPT* sessions. As she stated, “In *GEPT* classes, I work with the same guy for two times so that I feel very comfortable. After discussion, we would try to talk more in English.” (Interview #3, Dec. 15, 2008). In contrast, working with different people in “one-shot” *Table* sessions failed to help Emmy develop a sense of community while seeking for a sense of security and confidence seem to play an important role driving her to engage in a learning context.

The model of activity theory maps out the dynamic relationship of Emmy’s agency and her goals as well as the rules and division of labor embedded in the two types of non-credit English learning contexts.

CONCLUSION

The presentation reported four case studies exploring how two English teachers and two students participated in a series of non-credited English practice sessions. Activity theory was adopted as the theoretical and analytical framework to map out the complex relationship of individuals’ agency and the situated contexts. The findings revealed that the actions (either teaching or learning) of each case occurred in the non-credited teaching were interweaved by multiple elements, including contextual factors and individual agency. The two English teachers’ teaching was highly shaped by the unique non-credited nature and contexts of English sessions which were different from their regular credited English courses. The two student participants’ intensive participation in the *Table* sessions or in both *Table* and *GEPT* sessions resulted from their strong motives, their past and cultural experience of English

learning, and the situated community. The study findings demonstrate that activity theory may serve as an applicable theoretical and analytical framework to help us map out the complexity of individuals' behaviors in teaching and learning contexts.

Two limitations were observed in the case studies. First, the short data collection time failed to track the changes of the participants' participation in the informal sessions. Second, the lack of observation in the participants' regular and credited courses failed to gain a holistic picture of how each participant's activity system functions differently in various contexts. Future research may consider tracking cases across different contexts and different timeframes.

NOTE

¹ To help undergraduate students to prepare General English Proficiency Test (GEPT), the language center offered weekly-based *GEPT* sessions, including grammar, speaking, and in 1999 by Language Training & Testing Center. GEPT is "a skills-based battery, assessing test takers' listening, reading, writing, and speaking ability at five levels: Elementary, Intermediate, High-Intermediate, Advanced, and Superior" (Rover & Pan, 2008, p. 404)

REFERENCES

- Chiang, H. L. (2007). *Creating a rich English environment in college of Taiwan*. Paper presented in 2007 International Conference on English Education. Taipei: Shih Chien University.
- Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research*. Helsinki, Finland: Orienta-Konsultit.
- Engeström, Y. (1999). Activity theory and individual and social transformation. In Y. Engeström, R. Miettinen, & R. Punamäki (Eds.), *Perspectives on activity theory* (pp.39-52). New York: Cambridge University Press.
- Gao, X. (2008). You had to work hard 'cause you didn't know whether you were going to wear shoes or straw sandals! *Journal of Language, Identity, and Education*, 7, 169-187.
- Grossman, P. L., Smagorinsky, P., & Valencia, S. (1999). Appropriating conceptual and pedagogical tools for teaching English: A conceptual framework for studying professional development. *American Journal of Education*, 108(1), 1-29.
- Haneda, M. (2007). Modes of engagement in foreign language writing: An activity theoretical perspective. *The Canadian Modern Language Review*, 64(2), 297-328.
- Jonassen, D. H., & Rohrer-Murph, L. (1999). Activity theory as a framework for designing constructive learning environments. *Educational Technology Research and Development*, 47(1), 61-79.
- Jurdak, M. E. (2006). Contrasting perspectives and performance of high school students on problem solving in real world, situated, and school contexts. *Educational Studies in Mathematics*, 63(3), 283-301.
- Lantolf, J. P. (2000). Second language learning as a mediated process. *Language Teach*, 33, 79-96.
- Lantolf, J. P., & Genung, P. (2002). "I'd rather switch than fight": An activity-theoretic study of power, success, and failure in a foreign language. In C. Kramasch (Ed.), *Language acquisition and language socialization. Ecological perspectives* (pp. 175-196). London: Continuum.
- Lantolf, J. P., & Thorn, S. L. (2005). *Sociocultural theory and the genesis of second language development*. Oxford, UK: Oxford University Press.
- Leont'ev, A. N. (1978). *Activity, consciousness and personality*. Englewood Cliffs: Prentice Hall.

- Leont'ev, A. N. (1981). The problem of activity in psychology. In J. V. Wertsch (Ed.), *The concept of activity in Soviet Psychology* (pp. 37-71). Armonk, NY: Sharpe.
- Nelson, C. P., & Kim, M. K. (2001). Contradictions, appropriation, and transformation: An activity theory approach to L2 writing and classroom practices. *Texas Papers in Foreign Language Education*, 6(1), 37-62.
- Norton, S. J., McRobbie, C., & Ginns, I. S. (2007). Problem solving in a middle school robotics design classroom. *Research in Science Education*, 37(3), 261-277.
- Shen, L. B., Tseng, C. Y., Kuo, S. W., Su, Y. J., & Chen, M. Y. (2005). A preliminary study of college students' out-of-class English learning activities. *Chia-Nan Annual Bulletin*, 31, 464-475.
- Storch, M.(2004). Using activity theory to explain differences in patterns of dyadic interactions in an ESL class. *The Canadian Modern Language Review*, 60(4), 457-480.
- Vygotsky, L. S. (1987) *Thinking and speech*. New York: Plenum.

Structured Writing Activities for EFL Junior High School Students in Taiwan

Fang-chi Chang (張芳琪)

National Chiayi University

engivy@mail.ncyu.edu.tw

This paper reports the writing practice, a component of a 4-year EFL teaching project conducted in a junior high school. One of the four English class periods per week was scheduled for the structured writing activities. Three drafts were produced by each individual in one series of structured writing activities sequenced in three weeks in a row. Two series of the activities yielded 1068 drafts produced by 178 students. Most students were found progressing significantly from listing words in Draft 1 to listing sentences in Draft 2 to producing a piece of writing of paragraph nature in Draft 3. Moreover, the questionnaire responses from the students and the teachers show their positive attitude towards these structured writing activities.

INTRODUCTION

The author was contacted for the assistance with implementing a 4-year English teaching project in March 2008 by Pei-Yuan Junior High School in southern Taiwan. The project, designed by the team of four English teachers at Pei-Yuan and granted by the Ministry of Education (MOE), started in February 2007. The focuses of the project for the four years are vocabulary-reading instruction, vocabulary-writing instruction, speaking instruction and listening instruction respectively. Certainly the four language skills should be and are actually integrated into classroom instruction. This project is to conduct more various and structured activities as extension activities in class and on extracurricular themes as well for practicing one particular language skill.

The team administered the year 1 project without any assistance from the outside. To tackle problems encountered while implementing the project, the team contacted the author in March 2008 for assistance with the year 2 project. During the two semester periods from March to June and from September to November in 2008, the author met with the team once per month to assist with the in-class structured writing activities. The other activities, such as competitions of picture-book-creating, sentence-making, learning-portfolio-managing, etc., were designed and administered by the team for students to join at their will. This paper reports the structured writing activities and the findings.

One thing to be noted now and further elaborated in the method section is that the paper actually reports only the findings obtained from the limited data collected within six weeks from September 2008 to November 2008, though the year 2 project had been carried out from February 2008 to December 2008 for eight months. Besides having limited data, there were defects in implementing the writing project, but the findings suggest implications. Presented as a pilot study, this report is expected to contribute to pedagogical applications for classroom writing instruction and inspiration for future studies.

LITERATURE REVIEW

These structured writing activities in the current writing project were based on scholars' notions regarding the language development of young children in their native language. These notions should provide great implications for teachers to assist with the language development of SL/FL learners as well. Serving as the base for these structured writing activities, these notions are reviewed in the following.

A series of propositions stated by McNaughton (1995) explains how a growing child is socialized into literacy. McNaughton's first three propositions are about the opportunities, resources and activities which families provide to socialize children into the practices of literacy. The fourth proposition concerns the tutorial system, such as modeling, structuring activities, instructing strategies and commenting, provided as a scaffold for the writing development.

A similar notion on scaffolding was presented by Cazden (1996). Cazden interpreted Vygotsky's ZPD notion (1978) on children's language learning as "the difference between what a person can do alone and what s/he can do with assistance" (p.168). Assistance can be *demonstration, leading questions, and introduction to the initial elements of a task's solution* (Vygotsky, cited by Moll, 1990). As a result of the assistance, a child will be able to do by himself/herself tomorrow what he/she can do today with assistance. McNaughton's and Cazden's notions, suggesting modeling, structuring, strategy instructing, commenting, etc., provide pedagogical implications to assist with SL/FL development as well.

Besides, the Life Experience Approach (Tompkins, 2008), tallying with Vygotsky's ZPD, suggests that teachers offer appropriate assistance for the beginning writers to approach a further level. Five steps in an LEA activity are suggested by Tompkins (2008): a) provide an experience, b) talk about the experience, c) record the dictation, d) read the text, and e) examine the text. These steps are considered as well for the structured writing activities in the current project.

There is no best approach to teaching writing except for being flexible and supportive, according to Hyland (2002). The structured writing activities conducted in these junior high school classrooms were tried out based on the above-reviewed scholars' notions, with the author and the team's being flexible and supportive. The year 2 project is elaborated in the following section.

THE PROJECT

Setting

Located in Chiayi County, Public Pei-Yuan Junior High School (grade 7 to grade 9) is of 270 or so students, averagely 30 students per class. Four class periods are scheduled for the English subject per week, 45 minutes per period. The materials used are common language learning textbooks adopted by almost all the junior high school teachers in Taiwan, which are developed by the local scholars and teachers and approved by the Ministry of Education. Mandarin is the medium for the instruction in class.

Participants

Participating in the year 2 project were 88 junior 1 students and 90 junior 2 students respectively from three classes (grade 7 and grade 8 students, aged 12-14), the four English teachers and the author. Before attending Pei-Yuan, these students, having either Mandarin, Taiwanese or other non-English languages as their mother tongue, had received formal classroom English instruction for two years at their grade 5 and grade 6 years. Some of them might engage in extracurricular English activities at cram schools or at home with tutors. However, the team did not have the information for further explaining the students' English learning background.

Activities

The above reviewed scholars' notions contributed thoughts to the design of the structured writing activities for the year 2 project. This section starts with a general introduction to the two-semester project, followed by the elaboration on the activities for each class period.

The year 2 project was carried out for two semesters with a two-month summer vacation in between. Though the students were required to attend the one-month summer session for extra coursework, the whole project was not implemented during the summer session. Moreover, though the year 2 project was carried out for two semester periods, the administration of the structured writing activities in Semester 1 failed but fortunately contributed to activity modification for Semester 2. This situation is explained below.

The author met the team in late March 2008 for the first time. Through discussion, we designed two 4-period cycles of vocabulary-reading-writing activities for Semester 1 (one class period per week, from early April to mid June), to be explained further later. However, due to some factors, the second cycle was not completed. Therefore, collected were only the students' written products from the first cycle. Besides, according to the team, though we discussed together the steps for conducting the classroom activities, further details were not well confirmed among the teachers. Consequently, the activities were not conducted in a uniform way. The team thus abandoned analyzing the collected written products.

The experience obtained in Semester 1, fortunately, contributed to designing and implementing the semester 2 activities. The 4-period cycle of activities in Semester 1 and the 3-period cycle of activities in Semester 2 are introduced in the following. Though no data collected from Semester 1 was analyzed, the semester 1 activities are to be introduced too, in the hope of contributing thoughts.

Semester 1: March 2008 ~ June 2008

Period 1 Vocabulary Teaching → Text Reading The vocabulary words from the textbook unit were taught on the blackboard with the application of mind-map technique (which was introduced and practiced in the year 1 project); the text in the unit was then read with the teacher's guidance, for example, meaning and grammar explanation, etc.

Period 2 Brainstorming → Writing The students were given the opportunity to create their mind-map to record their brainstormed words, either which were taught in Period 1 or other words they would like to write down. Under the mind map the students then composed (e.g., sentences or paragraphs) to express, using their brainstormed words.

Period 3 Shared-Reading Some written products were selected and scanned to be displayed for shared-reading in class for input stimuli of different language levels and for further instruction given at appropriate moments.

Period 4 Rewriting The activities done in Period 2 were repeated.

Semester 2: September 2008 ~ November 2008

Only two 3-period cycles were completed in Semester 2, from late September to mid November. The team spent the first three weeks of Semester 2 designing the activities and preparing the materials. The data collected had to be analyzed for the report to be submitted to the MOE in early December. The schedule thus allowed only two 3-period cycles on two chosen themes *The Beautiful Ocean* and *Vincent Siew*, which were the extracurricular educational themes promoted by the school authority for the semester. The first theme is on environmental protection and the second, a local famous figure. These two themes were for both grade 1 and grade 2 students with materials of different levels being prepared for each grade. Elaborated in the following are the activities conducted on the first theme. These activities were repeated on the second theme.

Period 1 Vocabulary Instruction → Text 1 Reading → Drafting 1 Based on the Life Experience Approach, the teacher led the class to talk about their experience on *The*

Beautiful Beach. The talk was done in code switching and mixing of Mandarin and English. Meanwhile, the teacher and the class wrote interactively about their experience on the theme, *the Beautiful Beach*. The teacher created a mind map (Appendix 1) on the blackboard to record the brainstormed ideas and meanwhile introduce planned new vocabulary words related to the theme; the students were involved in contributing ideas, giving words and spelling out the words. Ending brainstorming, the teacher taught the new vocabulary words recorded in the mind map.

A short text (Text 1) on the theme (Appendix 2) written by the team of English teachers was distributed and read with teacher instruction. Completing the instruction on Text 1, the teacher collected Text 1 (to avoid copying) and distributed a blank sheet for the students to write whatever they would like to express on this theme. They were encouraged to write anything, recalling words, making sentences, or writing paragraphs, which was termed as Draft 1. They were also told that their writing wouldn't be graded and no penalty would be given on unsatisfactory performances.

Period 2 Text 1 Reviewing→Sentence Instructing→Sample Writing→Drafting 2

Text 1 was distributed to the class and reviewed. A Vocabulary-Sentence Sheet, prepared by the team, presenting vocabulary words, conjunctions and sentence patterns appearing in Text 1 was handed out for instruction. Then, a piece of Sample Writing (Appendix 3) written by the team on the theme was distributed for reading and instruction. After the three sheets (i.e., Text 1, the Vocabulary-Sentence Sheet, and the Sample Writing) were collected (to avoid copying), a blank sheet was given to the students for them to write on the theme for the second time, termed as Draft 2.

Period 3 Shared-Reading→Drafting 3

Choosing from drafts 1 and 2 some good products and a few sentences of mistakes (e.g., grammatical mistakes, semantic mistakes), the team typed them up to make the material for shared-reading. The material was distributed for reading and instruction, and Draft 3 was produced after the shared-reading material was collected.

Data Collection and Analysis

Six drafts by each individual were collected with three drafts on either theme. The students' written work was examined by the team without the author's participation. Later the team explained that considering many factors, for example, having only limited data, not being allowed much time for the report to the MOE, the results of a quantitative format being favored for evaluations of this kind, etc., the team decided to modify the method for data analysis originally planned for the semester 1 data.

The team sorted the written products into four categories:

- a) a blank sheet: only the writer's name was given, with/without drawing
- b) word writing (off context): the writer listed words, with/without drawing
- c) sentence writing: the writer listed simple sentences containing no conjunctions (e.g., and, but) and/or transition words (e.g., however) used at the intra-clause (e.g., I like the beach but it is dirty.) and/or intra-sentence level (e.g., I like the beach. However, I don't go there often.), and demonstrating no coherence, neither cohesion
- d) paragraph writing: the writer wrote sentences containing conjunctions and/or transition words used at the intra-clause and/or intra-sentence level and expressing coherence and cohesion

Though admitting that the judgment on "sentence writing" and "paragraph writing" was subjective, the team decided to adopt the criterion at the stage. Fortunately, according to the team, the sorting task was not difficult because it was easy to tell "sentence writing" apart from "paragraph writing" on the students' products. The sorting was made by each teacher on his/her own students' work. The sheets of each category were calculated for the quantitative analysis. No further analysis was done on accuracy or other aspects.

The other data source came from the questionnaire responses made by the students and the teachers on the year 2 project in early December. Responding to the team's request for suggestions on presenting more statistics to demonstrate how the objectives specified in the proposal to the MOE were fulfilled, the author suggested the administration of questionnaire and revised the questionnaire developed by the team. The questionnaires asked the students and the English teachers about their perception of all the English-learning activities conducted in the year 2 project. Further explanations to their responses were invited and encouraged but not required. The questionnaire responses provided data for quantitative and qualitative analysis.

FINDINGS AND DISCUSSION

Two cycles of 3-period structured writing activities were completed in two months. Six drafts were collected from each writer, with three drafts on either theme. The 1068 drafts were sorted into the four above-mentioned categories; two questionnaires were done and the responses were analyzed. The findings indicate that the structured writing activities lifted writing quality and were valued by the participants.

Lift Writing Quality

Table 1 and Table 2 present the quantitative data on junior 1 students' and junior 2 students' writing respectively. As shown in Table 1, on Draft 1 of Theme 1 produced after vocabulary instruction and Text 1 reading, 22 out of 88 junior 1 students put down only their names, 61 students recalled to list some words, four students listed sentences and four students produced paragraphs. In Period 2, a week later, after reviewing Text 1, studying the Vocabulary-Sentence Sheet and the Sample Writing, the students produced Draft 2. Only one student handed in a blank sheet, 17 students listed words, 64 students listed sentences and five students produced paragraphs. Obviously, more drafts were sorted into a higher-level category. On Draft 3, even more students (23 students) produced paragraphs.

On theme 2, similar changes were found. Moreover, more students produced paragraphs on Theme 2 than on Theme 1, even though Theme 2 was more difficult than Theme 1, to be explained later. The findings on the junior 2 students' performances were similar to those on the junior 1 students' performances.

Since the students move upwards in terms of levels, from leaving the sheet blank to listing words to listing sentences to producing paragraphs, this result indicates that the structured writing activities lift writing quality. Three drafts by two students of a high and a low proficiency level respectively on Theme 1 are presented in Appendix 4 as examples for writing progress

The students seemed to become more comfortable and competent (or simply more experienced) in making sentences and writing paragraphs on Theme 2 than on Theme 1. However, among the junior 1 students, on Theme 2, more students remained in the blank-sheet group. On Theme 1, 22 students handed in a blank sheet on Draft 1 and one of them remained in the blank-sheet group on Draft 2 and Draft 3. On Theme 2-Draft 1, though fewer students turned in a blank Draft 1 but five and three on Draft 2 and Draft 3 respectively remained in the blank-sheet group. A possible explanation is given below.

Table 1. *Statistics of writing categories on junior 1 (total 88 students)*

Theme 1 <i>The Beautiful Beach</i>					
	Blank	Word	Sentence	Paragraph	absent student
Draft 1	22 (22/88=25%)	61 (69.3%)	4 (4.5%)	4 (4.5%)	0
Draft 2	1 (1/87=1.1%)	17 (19.5%)	64 (73.6%)	5 (5.7%)	1
Draft 3	1 (1/82=1.2%)	10 (12.2%)	48 (58.5%)	23 (28%)	6
Theme 2 <i>Vincent Siew</i>					
Draft 1	12 (12/88=13.6%)	44 (50%)	26 (30%)	6 (6.8%)	0
Draft 2	5 (5/84=6%)	17 (20.2%)	48 (57.1%)	14 (16.7%)	4
Draft 3	3 (3/88=3.4%)	10 (11.4%)	45 (51.1%)	30 (34.1%)	0

Table 2 *Statistics of writing categories on junior 2 (total 90 students)*

Theme 1 <i>The Beautiful Beach</i>					
	Blank	Word	Sentence	Paragraph	Absent student
Draft 1	37 (37/87=42.5%)	28 (32.2%)	18 (20.7%)	4 (4.6%)	3
Draft 2	8 (8/82=9.8%)	7 (8.5%)	44 (53.7%)	23 (28%)	8
Draft 3	8 (8/81=9.9%)	3 (3.7%)	38 (46.9%)	32 (39.5%)	9
Theme 2 <i>Vincent Siew</i>					
Draft 1	8 (8/82=9.8%)	40 (48.8%)	25 (30.5%)	9 (11%)	8
Draft 2	9 (9/83=10.8%)	18 (21.7%)	24 (28.9%)	32 (38.6%)	7
Draft 3	2 (2/77=2.6%)	4 (5.2%)	30 (39%)	37 (48.1%)	13

On Theme 1, *The Beautiful Beach*, the vocabulary words taught were: diary, went, beach, dirty, something, make, litter, keep, clean, eat, seafood, hope, forever and on Theme 2, *Vincent Siew*: legend, vice-president, senior high school, university, graduate school, treat, encounter, trouble, example, efforts. The words for Theme 2 are more advanced. This might explain why more students (though only 2 more) remained in the blank-sheet group. The students also expressed their preference for Theme 1 through the questionnaire.

According to the team, the self-developed materials on these two themes were collected before the drafting task (to avoid copying from the materials unto the drafts). Between the three drafting periods in three weeks, the students had no access to the materials. So the progress of writing performance seemed to be the effect of the treatment (i.e., the conduction of the structured writing activities) rather than the effect of memorization.

Positive Attitudes towards Structured Writing Activities

Eighty junior 1 students (90.9%) and 81 junior 2 students (90.0%) expressed their positive attitude towards those writing activities when asked whether they liked the set of structured writing activities. The major reason given is, "they benefit writing ability." The majority of the students thought these activities facilitated

- vocabulary learning (through the mind-map technique, 94.2% of junior 1 students and 85.6% of junior 2 students),
- sentence pattern learning (through the Vocabulary-Sentence Sheet instructing, 87.5% of junior 1 students and 85.6% junior 2 students)
- writing performance (through shared-reading, 84.1% of junior 1 students and 87.7%

of junior 2 students).

Some of those holding a negative attitude explained, “Not interested in English; do not understand the content; do not like writing; too hard.”

The students were asked which activity (or activities) among the conducted writing activities they liked the most. The mind-map vocabulary instruction received the most votes (38 junior 1 students and 24 junior 2 students). The shared-reading, following the mind-map technique, was favored by 35 junior 1 students. The text-reading was preferred by 19 junior 2 students as the second, followed by the shared-reading preferred by 18 junior 2 students. They favored the mind-map technique because “the map holds the words in link, which benefits the memory of the words,” according to the students.

More than 44 percent of junior 1 students and 20 percent of junior 2 students expect the activities to continue; 46.6 percent of junior 1 students and 66.7 percent of junior 2 students responded with “no comments” towards the question on continuing these writing activities or not in future. It then means that these activities are welcomed by and acceptable to the majority of the students. However, 9.1 percent of junior 1 students and 13.3 percent of junior 2 students did not want the activities to continue. Some of them explained, “not interested in English.”

All four teachers in the team agreed that these planned writing activities promoted writing performance significantly. One teacher noted that the students at least improved in arranging the words of different lexical parts of speech appropriately while making sentences. What another teacher wrote in Mandarin is translated: These activities provide opportunities for the intermediate-level students and advanced-level students to explore writing and the low-level students to experience writing. According to the team, even low-leveled students showed their enthusiasm for these activities.

CONCLUSION

As explained earlier, the writing project reported was not a well designed and conducted research, neither were the limited data analyzed thoroughly to yield significant findings. However, the meager findings seem to provide one more piece of evident to verify Krashen’s Input Hypothesis (comprehensible input, 1985) and Swain’s Output Hypothesis (2000). In these writing periods, these students were provided with comprehensible input stimuli and they advanced a step to produce output which is more comprehensible.

Leki (2005) viewed writing as being better than any other kind of language instruction because it allows students of different proficiency levels to push FL output individually at their pace without holding other students back, which is often seen in conversation classes. In addition, some studies have shown that more complex language forms turn up first in a learner’s developing language system in text rather than in spoken interactions (Bardovi-Harlig, 2000; Weissberg, 2000), so writing seems to well reflect the learners’ best performance.

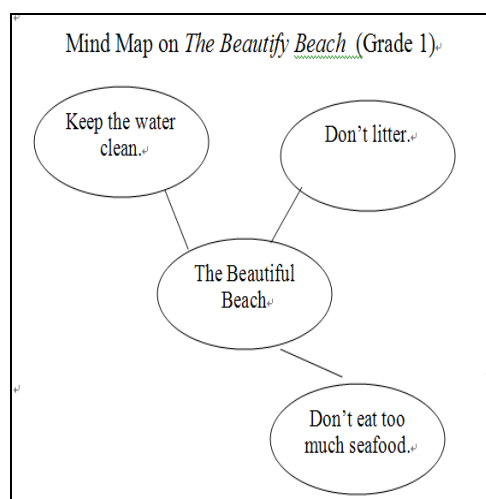
The readers might argue that the lifted writing quality is simply the result of teaching-learning effect or even the fresh-memory effect, since the students generated output right after receiving input. This teaching-learning effect should be a common and certain goal and a result expected in class. Fortunately, the current project provided the opportunities for this expected result to be detected. Significantly, this project shows that the structured writing activities managed to provide comprehensible input at appropriate moments for comprehensible output to be pushed out. As mentioned, reported as a pilot study, this reported project intends to inspire thoughts on incorporating writing instruction into the junior high curriculum and structuring the activities for writing instruction.

REFERENCES

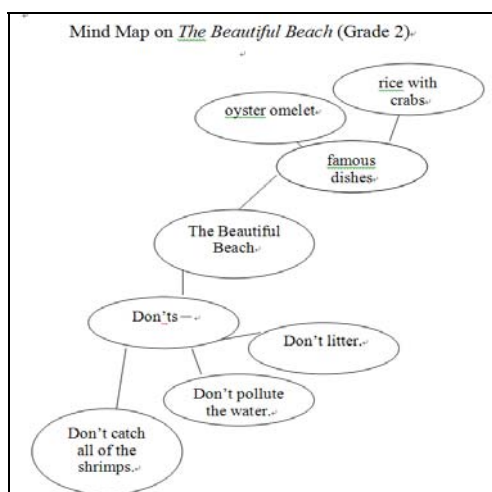
- Bardovi-Harlig, K. (2000). *Tense and aspect in second language acquisition: Form, meaning and use*. Malden, MA: Blackwell Publishers.
- Cazden, C.B. (1996). Selective traditions: readings of Vygotsky in writing pedagogy. In D. Hicks (Ed.), *Discourse, learning, and schooling* (pp.165-185). Cambridge: Cambridge University Press.
- Hyland, K. (2002). *Teaching and researching writing*. England: Pearson Education.
- Krashen, S. (1985). *The input hypothesis: Issues and implications*. New York: Longman.
- Leki, I. (2005). The challenges of teaching EFL writing. *Selected Papers from the Fourteenth International Symposium and Book Fair on English Teaching*. Taipei: Crane.
- McNaughton, S. (1995). *Patterns of emergent literacy: Processes of development and transition*. Auckland: Oxford University Press.
- Moll, L. (1990). *Vogotsky and education: Instructional implications and applications of sociohistorical psychology*. Cambridge: Cambridge University Press.
- Swain, M. (2000). The output hypothesis and beyond: mediating acquisition through collaborative dialogue. In J. Lantolf (Ed.), *Sociocultural theory and second language learning*. Oxford: Oxford University Press.
- Tompkins, G. E. (2008). *Teaching writing: Balancing process and product* (5th ed.). Columbus, Ohio: Pearson.
- Vygotsky, L. (1978). *Mind in society: Development of higher psychological process*. Cambridge: Harvard University Press.
- Weissberg, B. (2000). Developmental relationships in the acquisition of English syntax: Writing vs speech. *Learning and Instruction*, 10, 37-53.

Appendix 1 The Mind Map on The Beautiful Beach

a. For Grade 1



b. For Grade 2



Appendix 2 Text 1 on The Beautiful Beach

a. For Grade 1

Dear Diary:

Today I went to the beach. I am happy, but the beach is dirty. We can do something to make the beach beautiful.

1. Don't litter.
2. Keep the water clean.

3. Don't eat too much seafood.
I hope the beach can be clean forever.
(The underlined parts were the linguistic items to be taught.)

b. For Grade 2

Dear Diary:

Today I went to the beach. I ate two kinds of famous dishes at a restaurant there. For example, one was oyster omelet. The other was rice with crabs.

I am happy, but the beach is dirty. We can do something to make the beach beautiful. First, don't litter. Second, don't let the dirty waste pollute the water. Third, don't catch all of the fishes, crabs, shrimps and oysters.

If we can follow the tips above, we can enjoy a beautiful beach forever.

Appendix 3 Sample Writing on *The Beautiful Beach*

a. For Grade 1

I went to Bu-Dai port with my family today. The air is fresh, but the beach is dirty. I am sad. We can do something to save the beach. First, don't litter. Second, keep the water clean. Third, don't eat too much seafood. Finally, we can enjoy the clean beach forever.

b. For Grade 2

I went to Kenting yesterday. I went there with my family. The scenery there was beautiful, but the beach was dirty. I was very sad. If everyone could follow the tips, the beach would be beautiful again. First, don't let the dirty waste pollute the water. Second, don't litter. Third, don't catch all of the fishes, crabs, shrimps and oysters.

I also ate two kinds of seafood. One was rice with crabs. The other was oyster omelet. They were both delicious.

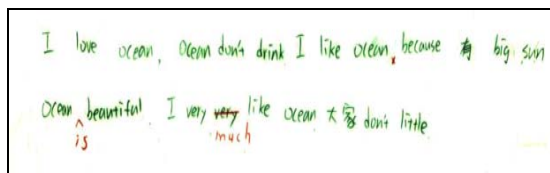
If we can keep the beach clean, we can enjoy the scenery and the food forever.

Appendix 4 Three Drafts by One Student

a. A high-leveled student

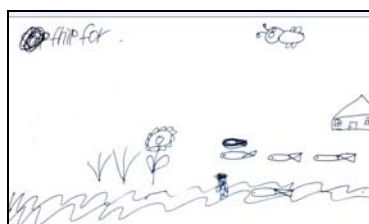


Draft 1

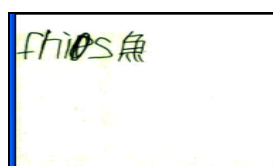


Draft 2

b. A low-level student



Draft 1



Draft 2

Sunday we went to the beach. I was happy because this is beautiful, but
people always fishes. His look Fisher don't take Fisher.
I love beach because very beautiful. I like beach because air and water very blue.
People love beach swim, play ball,
People
to

Draft 3

Today I go to I am Penny
the beach happy

Draft 3

遠距中英口筆譯教學個案分析

Hsiu-chen Chang (張秀珍)

National Taipei College of Business

folhcc@hotmail.com

科技的進步使得當今教育在二十一世紀呈現前所未有的普及與多樣化，從幼年到老年，從富商到貧民，超越年齡，打破階級，人類學習的夢想得以無限的伸展。遠距教學是全球教學的趨勢，傳統教學似乎受到不少威脅。然而虛擬教室真的可以取代傳統教室嗎？兩者之間的優缺點為何？學生選擇遠距教學的目的何在？遠距教學幫學生實現何種夢想？本篇論文希望藉由兩件筆者親身參與的中英口筆譯教學個案，亦即在中正大學教授的非同步網路翻譯教學，和筆者在中華電視錄製的英語口譯教學課程以及在空中學院實際教授的過程分析來解答以上的問題。

導論

Hope is the thing with feathers—
That perches in the soul—
And sings the tune without the words—
And never stops—at all—

And sweetest—in the Gale—is heard—
And sore must be the storm—
That could abash the little Bird—
That kept so many warm—

I've heard it in the chilliest land—
And on the strangest Sea—
Yet, never, in Extremity,
It asked a crumb—of Me.

(Emily Dickinson #254)

科技的進步使得當今教育在二十一世紀呈現前所未有的普及與多樣化，從幼年到老年，從富商到貧民，超越年齡，打破階級，人類學習的夢想得以無限的伸展。即使在最艱難的時刻，教育也可帶給人心希望與成長。遠距教學的議題過於龐大，全貌非一篇短文可以涵蓋，本篇論文重點在兩件筆者親身參與的遠距中英口筆譯教學個案分析。本文第一部分將對臺灣遠距教學做一簡單的歷史回顧；第二部份報告筆者在中正大學教授的通識課程網路翻譯教學的經驗，藉由學生的評鑑比較遠距與傳統筆譯教室的差異，並從教師觀點探討影響教學成果的因素；其次分析筆者在中華電視錄製的英語口譯教學課程以及在空中學院實際教授的過程；最後比較分析影響兩個案的要素：包括經費、教材編輯、班級大小、硬體設備、技術支援、學生的背景與學習動機、師生間互動、教師可能

遭遇之挑戰、學生對課程的建議與評鑑。

臺灣遠距教學歷史回顧

遠距教學在臺灣行之已經有四十餘年，茲簡述其歷史如下：

民國五十四年：二月教育部指定當時台灣省立台北商業職業學校試辦空中廣播課程，開授『商業概論』與『會計學』，由教育電台製播，進行教學。

民國五十五年：正式成立高級商業廣播實驗補習學校，

民國六十年：中華電視公司成立，空中教學步入電視教學階段。空中補校奉准改制為專科，設有國際貿易、企業管理、銀行保險、商業資訊及會計科；

民國八十六年：增設應用外語科英文組。除校本部外，另設有中壢、基隆、宜蘭、花蓮、台東等五個教學輔導處。

民國九十一年：六月奉准改制為空中進修學院，設應用商學系，招收二技學生。

空中大學為民國七十五年成立，與空中進修學院偏重商業不同，空中大學設有六個學系：人文學系、社會科學系、商學系、公共行政學系、生活科學系及管理與資訊學系，發給學士文憑。在基隆、台北市、台北縣、新竹、台中、嘉義、台南、高雄、宜蘭、花蓮、台東、澎湖、金門等地設立學習指導中心，學生可就近選擇學習指導中心參加面授教學、考試、課業及生活輔導或社團活動。

中正大學網路非同步翻譯教學

(一)緣起：國家教育政策發展：

一個新的教學方向常常需要政府政策的定位, 推動, 以及規範。二十世紀末, 教育部看到了網路發展與遠距教學結合是全球教學的趨勢, 民國八十八年將「專科以上學校開辦遠距教學作業要點」行文全台大專院校, 大致規範如下:

1. 網路遠距教學活動包括上課、討論與其它學習活動。不須同一時間、同一地點達成(即非同步)。網路教學系統須具備教學, 課程進度時程, 同學間, 與師生間的交流管道, 教學系統之使用說明, 與解惑多項功能。
2. 教學部份必須包含註冊、公告、教材、討論、作業、線上自我評量、對學生的評量、課程評鑑。討論區為避免混亂, 須以討論範圍區分。對學生的問題, 教師須確保於所定期限內給學生滿意答覆, 所有討論過程學生均應得以讀取, 作業的公告與討論均在網路進行, 學生作業可在網路繳交, 並能立即知悉是否成功送出。
3. 對學生之作業評量可包括作業評分、討論評分、個別小組評分、考試評分、學習態度評分、同學互評(同學彼此出題互考)、網路網頁閱讀之成效、網路學習活動(如討論)、網路學習系統記錄之學生學習歷程皆應作為學生學習評量之依據。
4. 為了解學生基本資料以及個別學習之情況, 系統必須有註冊機制, 以方便教師進行教學管理。
5. 為確保教學品質網路教學須訂定時程教師須公告教學進度, 作業、討論、專題、考試之時程和進度。同時公佈這些學習活動評量中有關時限與分數之關係, 以使同學了解整個時程。
6. 網路教學因教師無法與學生當面溝通 教學系統因此必須提供機制達成下列功能: 教師須了解個別學生之學習, 以及小組互動; 教師要能了解學習因素與學習效果之間的關係, 並作統計與分析; 示警機制: 教師必須知道那些學生以及小組須付出關懷。

(二) 中正大學通識教育中心：網路遠距教學夢想的推手

中正大學位居嘉義縣民雄鄉，交通不甚便利，網路最大功能之一是跨越距離，就長期發展而言，中正當可受惠於網路遠距教學。中心中正大學通識教育中心在八十八年教育部頒布「專科以上學校開辦遠距教學作業要點」後，立即推動為期三年之網路非同步教學輔助課堂教學計劃，每年挑選一些通識課程予以數萬圓經費補助。筆者之翻譯課程有幸獲選，參與此一教學新里程，以下為教學效果分析。

(三) 網路遠距非同步翻譯教學

有夢相隨，希望最美。筆者雖不精於電腦網路，然總是希望能利用現代科技，幫助學生以最小的代價，最省力，最輕鬆地，達到最高學習成效的夢想。然而中正大學和我皆是初次嘗試，結果是事倍功半，不甚完美。每位教師的教學成果皆會受到學生、教材、評量以及經費資源的影響，茲將此課程教師面對的挑戰分述如下：

1. 班級人數及學生程度

翻譯課如同作文課，由於須改作業，耗時費力，傳統教學是以小班為佳。然而通識網路課程是以大班修課為原則，此次非同步網路翻譯課程虛擬教室的學員，共有四十九位學生，分別來自歷史、哲學、數學、心理、勞工、政治、資訊、電機、財金、企管和法律各系，計有十一系之多。其中四十人是大一新生，九人為大二學生。學生的英文程度以及對電腦網路的熟悉度皆會引影響教學。資訊、電機的學生較熟悉電腦與網路；一年級和二年級，不同學門學生由於沒有能力分班，中英文程度也參差不齊。

2. 課程設計與教材選定：

一切課程設計皆須以學生為主體，訂立合理之教學目標，繼而選定教材。由於這是通識課程，不宜過難；網路大班教學，應以公佈答案，解說範例為主。翻譯涉及語音、語法、語意與文化多層面，十分龐雜，我選定一本相當好的柯平先生的英漢與漢英翻譯書作為翻譯課程之主要教材為，其它教材決定再依學生不同需求補充。預定十六週將緒論，翻譯的語意學，翻譯的過程，錯譯分析等四章共十六節的內容討論完。此外，開學之初，本人曾將國內外翻譯相關網站提供給學生，請學生針對所需，提出意見。學生意見五花八門，有要求增加英文字彙、新聞翻譯、跨文化溝通、以及口譯…等需求。以下列出同學討論中英文各一篇供參考。

學生一、

1. 絕大多數大學生字彙不足，應該是首要的問題吧，這或許也是中學的填鴨教育所帶來的缺失吧，中學時候，大部分的老師只要求成績上的表現，而忽略英文不該只是門科目，而是個語言，造成學生們上大學後很少接觸英文，所以希望老師可以在上課討論時，可以針對重點討論，提供一些字彙；其次就是關於不同文化背景，所造成在語言及行為上差異，這也是我為什麼會選這門課重要原因，相信同修這堂課心裡系同學們一定也是對這很感興趣，因為我們都知道，不同成長環境，可以造成不同人格發展及行為模式，這也就是為什麼不同文化背景人群相遇時，會產生溝通上困難；而翻譯，是跨文化溝通重點，如何能翻譯恰到好處(如音譯和意譯運用)，就要看翻譯者對於兩個文化了解程度，所以希望老師能在討論時候，能針對議題作不同文化背景比較；最後就是有關於口譯問題，其實，我也是對口譯心存畏懼，要在沒有草稿，不准翻字典情況下，把原文翻譯好說出來，並且還要了解文化差異以及發言者可能隱含寓意，這對我而言實在是有點困難，個人覺的這是個綜合性的問題，不曉老師覺得可以從什麼方向著手解決???

學生二、

Though I long to become an outstanding interpreter, I'm afraid that my English is not

good enough. So I think I'd like to start with some easy translation. It's easier than interpretation, isn't it? As to the content of our translation lesson, I'd like to try some things that are close to our daily life first. Then I don't need to learn the background knowledge that is new to me. For example, I don't expect to start with British foreign policy because I have no idea about that. It may take lots of specific vocabulary. I may need lots

of time doing preparation work. It's necessary when I'm trying to do this kind of translation, of course. But not now, not for a beginner. I hope that you could get the point of my statement.

由於此班網路翻譯教學學生人數眾多，這學期專重筆譯教學，口譯並未實施。此外因為學生各有不同的知識背景及興趣所好，所以本課程設計主要是提供學生筆譯有關語言文化部份的一般常識。其次以分組討論的方式，讓學生依個別不同的興趣，選擇重點做英文字彙，跨文化溝通、漢語拼音、音譯與意譯等不同議題深度研究來滿足學生不同的需求。從事網路教學，教師須付出極大心力與時間。就像是學生有一個二十四小時的家教，老師須迅速有效的為每一個學生解決問題。雖然有軟體可以幫忙提供一些服務，絕大的工作仍須老師親自來做。

3. 教學成效與建議

(A). 本文之教學成效研究範圍資料由十六週內學生上網登錄，使用，瀏覽的次數、回家功課、考試範圍、分組報告，學生討論與對課程的建議與評估的各部份所組成。教學網站上提供的功能包括：登錄次數排行，使用次數排行，學生瀏覽教材次數排行，教材被瀏覽次數統計，及學生使用情形。從這些排名，教師可以追蹤到學生的學習情形。瞭解哪些學生積極上網閱讀教材，參與討論；哪些學生進度落後，甚至完全沒有來上網上課。問題是這些表格只列出一到十五的排名，然而事實上班上最須注意的，應該是排名墊底，嚴重落後，甚至完全沒來上課的學生。此外，如何讓網頁活潑有趣吸引學生，增加聲光圖色，這些都有待學校對網路教學老師以及助理的技術指導與支援才能改進。

(B) 作業

網路翻譯教學的作業原先一半來自採用的教科書，另一半則為本人所出的課本以外翻譯問題。課本教材內的翻譯作業，學生回答的情形良好。課外的部份，一半以上的學生都沒有作答，理由是問題太難，通識課程不應太專業等等。

由於網路上學生的翻譯，成果保留完整，這對喜歡作翻譯分析研究的教師十分有幫助。學生普遍的反應，英譯中的回譯比中譯英翻譯更困難。這現象可能顯示，傳統文化在年輕的一代，似乎有逐漸流失的現象。所以會將一般日常成語 "人外有人，天外有天" 硬譯為 "每一偉人背後，都有另一偉人。" 此外 "Govern a big country as you would cook a small fish -- very gently" 應是來自老子道德經第六十章 "治大國，若烹小鮮。" 誤譯的原因除了對 "govern" 這字的意義不清楚外，很重要的原因是學生之中，無人知道老子的這句話。這也顯示出，不只是英文，學生的中文，也急需加強。

教師在出作業時常遇到 "You have no permission" 的回應。此外有時收到假訊號說 "作業製作完成"，事實上確是失敗的。而有時學生只能看到問題，卻無法作答，或是繳交作業。學生因此隨便跑到任何一個可以作答的作業區胡亂繳交，發生將 A 作業交到 C 作業區，C 作業交到 F 作業區的混亂情形。關於繳交期限，有時學生遲交一至三個月，隨時要求教授回頭改作業給成績，建議系統應該制止這些混亂的行為。此外，評分的系統九至十一月都無法運作，運作之後也經常有狀況產生，這些都應改進。

(C) 考試：

學期開始時，翻譯的測驗問題一半來自教材課本，另一半來自課外，英文程度好的學生可以得到滿分，然而也有不少的學生得到零分，因而產生極大的挫折感。解決之道是百

分之九十的問題都來自教科書，以提昇學生的自信心。中翻英及英翻中應該並重，讓學生瞭解翻譯課程牽涉中文及英文的兩種語言能力，至於難易和內容應視學生的興趣以及語言能力作適當調整。此外，作業十分容易請人代打，因此建議在網路教學，測驗一定要在教室內實施，確定學生的學習以及評分的可信度。

(D)分組報告

網路教學的弱點之一是缺少互動，分組活動施行，可以有效增加學生間的互動。此外還可以讓學生了解翻譯可以是團隊活動，學習有效的分工合作，收集資料，解決問題。另一好處是同組之間學生可以彼此激勵，督促對方的學習。

各小組資料蒐集十分豐富，這些工作是個人很難獨力完成的，因此應當增加分組活動。然而相當多數的同學將網路上的資料整篇下載，大量抄襲，不分析消化整理，可以看出學生嚴重缺乏尊重著作權的觀念以及寫作研究的原創性，這些是教師應該予以重視。

(E)學生討論

網路翻譯教學網站設定了多項討論區：包括每週作業討論區，分組報告討論區，以及問答區。然而，討論情形並不十分踴躍，學生提供的理由是中文打字太麻煩，不方便上網，這些問題亟待解決。此外學生也有在討論區發生口角爭執，網路發言也應訂定禮儀規範，以避免與課程無關的爭執。

(F)學生的評鑑：比較遠距與傳統筆譯教室的差異

曾經在課堂中，針對網路教學之優缺點加以討論。對課程評估 5%很滿意，70%對課程滿意，20%對課程尚可接受，5%不適應。以下節錄學生反應的重點：

優點

- 1 網路課程不受時間限制，如夜貓子即可免除日夜顛倒的問題(不過這只適用於有自制能力的人)，大家可以自由選擇最適合自己的時間上網討論。
- 2 不受空間限制，網路教學不須有形的教室，更可以接近處處是教室的理想，即使在家中也可以學習。
- 3 對於個性比較害羞的同學，以網路教學可使大家更勇於發表意見。
- 4 舒服，可以一邊上課，一邊吃早餐，不會有不尊重老師的問題發生。在現今科技這麼發達的時代，我們可以充份享受科技帶給我們的進步與便利。
5. 電腦網路可像 BBS 一樣將談話、研究、討論的內容記錄下來，方便我們日後查閱。

缺點

- 1 容易受物質方面影響(如沒有電腦或網路的同學就較難使用)。
- 2 不諳電腦的同學很吃虧。
- 3 會缺乏人與人之間的接觸及互動，老師也不能觀察到學生不積極於課堂上的參與，學生有沒有自信。
4. 如果以網路教學，學生整天面對的是冷冰冰的電腦，我們和電腦都是以相同的模式相處，就算是上不同的課也不會有所不同。但如果是面對面的教學就可以接觸到不同的老師及同學。大學是一個小型的社會，多和各種人接觸，可以提早適應社會，因為當以後我們踏出社會，我們去找工作面試時，面試的人不是電腦，是活生生的人。
5. 網路教學，我認為對自動自發的學生是可行的；而一些較懶散的同學，則無太大用處。
6. 網路教學勢必是一個未來的趨勢。但其中的困難還有很多，包括個人對網路的基本認識、基本的使用方式、規範…等等。所以現在只是剛起步的階段，必須要好好發展才是。而現在翻譯在網路教學當然可行！不僅可以當成一般人學習的教材，更有助於我國對外語程度的提升。
7. 我覺得這本書寫得不錯，蠻專門、深入的。但這樣的一本書應該是基本教材，等到真

正進行翻譯時，還是要靠老師從旁引導修正才對。網路上課不算辛苦，若非真有心想唸，同學們抄抄答案，交出作業根本也學不到什麼，考試也是一樣。網路教學效果不大，總覺得不是教材或老師不好，只是學生素質不齊，無法配合。

8. 網路教學雖然比較方便，可是很難使得師生間產生良好互動，因為每個人上網時間不同，而且如果不是學生間熟悉度較高的話，可能也很難產生討論課業的效果。

9. 網路教學應建立在一套完善的設備及規定。我們學校網路出錯率極高，所以網路教學極不合適。

10. 網路使上課輕鬆，但缺乏約束力，或許加上嚴格的規範與條例，才比較容易實行。這門翻譯課學到一些概念，但沒有重覆運用，相當容易忘記。

11. 這門課學到一些中譯英和英譯中的問題，網路教學不錯啦！很方便！可

是如果要上翻譯，我會比較喜歡到教室上課。

12. 網路教學是可行的，只不過大家現在比較沒有上網上課的習慣，可能還需一段時間培養。而翻譯這一方面，其實我比較喜歡老師之前無範圍的出題，雖然可能考得不好，但會去思考英文和中文翻譯的關係。至於如何去改進，我也不知道，可能需要時間吧，至少網路系統這方面，需要好好去改進的，要不然常常當，也不是辦法。

13. 教材對翻譯作了詳細的介紹，例子很多，幫助我們理解蠻有用的。網路教學當然可行，可打破時間和空間的限制，但前提是必須每個人都有強烈的學習動機，在網路上熱烈的討論，才能引起共鳴。

從教師觀點探討影響教學成果的因素：

經過一學期網路翻譯教學，分析發現，多數學生對翻譯教材大致可以接受，若能加強傳統教室學習與網路教學搭配，學生對教材的理解與接受程度將會增加。其他問題，例如技術支援，電腦與網路的知識、上網的方便性、學習的動機、班級學生人數的多寡、能力分班等才是可能影響翻譯教學成果的因素。茲討論如下：

(1) 技術支援：學校的教學網頁只包括：課程大綱、教材、討論區、公佈欄、作業區、成績公佈。缺乏教學系統的使用說明與解惑功能，為求畫面齊整以及網站的便於管理，網頁之設計早已設限，教師僅能就字型大小、畫面顏色作更改，本當有的html程式語言功能卻付之闕如。造成學生在上網時，對網頁設計之呆板、功能簡陋，頗有微詞。此外，網站所有的功能經常出現使用困難的問題。甚至學生關切的成績公佈重要功能一直到期末都無法應用。學生一半的問題都是與網頁的功能有關，造成師生共同的挫折與困擾。所以，沒有很強的技術支援，網路翻譯教學絕對無法成功。

(2) 電腦與網路的知識、

學生上網討論情形非常不踴躍。經過與學生討論，追究原因，學生皆回答：中文輸入太難或太慢。英文又不夠流暢來表達自我，所以裹足不前、不常上網討論，甚至不交報告。需要分組討論時，常以電話，或約出來面談的方式解決。此外，學生有問題，也經常到老師研究室面對面的討論，堅持這才是迅速、有效的溝通方式。學校當局應有機制由電算中心設立專職解決同學的疑難雜症，給予必要的技術支援，以解決教師在網路教學的額外負擔。

(3) 上網的方便性

遠距教學原是為了上班族解決上班、上課的時間衝突；以及空間上，旅途費時奔波之苦。只需一台接上網路的電腦，家庭即是教室，經濟又舒適。然而經濟拮据的學生常無法承受額外支出，造成上網討論、交作業與報告的困難。上網的方便性因此也是決定學習成果的要件之一。這些問題學校當局應認真考慮，積極解決。

(4) 學習的動機：網路教學由於沒有師長課堂的督促、壓力，同學的鼓勵和競爭，缺乏

自制力的同學十分容易怠惰、進度落後。本課程平均只有百分之五十的學生按時繳交作業，以及參予課堂考試。因此網路教學，學生學習的動機強烈與否，與學習的成果有重要的關聯。

(5) 班級學生人數的多寡：翻譯教學無論口、筆譯，教師對學生皆需付出極大的心力與時間。傳統上是以小班教學為宜；然而學校行政人員堅持大班教學。以此次網路教學的班級為例，四十九名學生，教師除了每週更正學生的翻譯練習，每個月有一至二次的課堂測驗評分，每天須回答學生各種問題，還要追蹤學生的學習狀況。當學生發生學習嚴重落後時，還要以電話或電子郵件與學生系辦行政或導師聯絡，這些對教師而言是極沉重的負擔。所以，有意提供網路翻譯教學課程的教師要慎重的考慮班級大小，學生多寡。

(6) 能力分班：由於學生來自不同系別，中英文能力有極大的差異。對教學活動、教材理解、作業與測驗的接受能力與學習成果，有兩極化的現象產生。亦即對某些學生極為容易的作業與測驗，某些同學得到滿分，有些卻得零分。讓老師在教學活動課程、教材、作業與測驗的準備，倍感困難。所以能力分班也是影響翻譯教學的重要因素之一。

空中學院應用外語科英語口譯：

(一) 緣起：

九十三年空中學院應用外語科決定新增開英語口譯課程，九十三年底筆者授命籌劃課程，除了編纂教材，錄製電視與電台講授節目，還須編寫函授補充教材，習題與解答。九十四年九月開課至今，筆者並擔任台北授課教師，教導兩班學生。以下詳述中華電視錄製的英語口譯教學課程以及在空中學院實際教授的過程，並分析從學生及教師觀點探討分析影響課程要素，包括經費、教材編輯、班級大小、硬體設備、技術支援、學生的背景與學習動機、師生間互動、教師可能遭遇之挑戰、學生對課程的建議與評鑑。

(二) 學生的背景與學習動機

(1) 學生的背景空中教育為學習回流之教育，八十七年度開始，採登記入學方式，使得就學管道更加暢通，嘉惠更多社會在職人士高階進修機會，以提昇國民文化水準。學生的年齡從十八到六十八老少皆有；口譯班學生的職業涵蓋空中小姐、職業軍人、船務公司老闆、書店店長、導遊領隊，進出口公司業務員、幼稚園教師、水電工、小販、家庭主婦、退休人士、各行各業，應有盡有。每位學生的英語程度差異比起中正的學生更大，從在英語系國家多年居住，到幾乎無法開口說英語，差距不可謂之不大。97年調查選擇空院的原因如下：

- a、時間有彈性，只需假日上課
- b、可兼顧家庭及工作
- c. 免試入學
- d. 學費低廉
- e. 提升學歷
- f、朋友介紹
- g. 自我成長
- h. 給孩子以身作則

(2) 學習口譯動機：

學習動機強烈與否與學習成果有密切關係。每位學員來到空院，選擇應外系，選擇這門課背後皆有一個故事，有些人只是陪姐姐妹妹來，也有人是有強烈之事業人生規劃，茲將動機分類並選擇紀錄一些學生的問卷答案如下：

a. 事業規劃：就長期目標來看，我希望自己未來8-10年內能拿到翻譯所碩士學位或通過外交領事考試。專科學歷是我報考這些考試的基本條件，選讀本課可以讓我學以

致用。短期而言，會去從事水電工作，外國的技術手冊和法規閱讀也需加強英文。(我深深祝福這位同學，希望本課程能幫他實現夢想！)

b. 工作需求：做出口貿易，工作需要。且提升能力可增加工作機會。

c. 生活需要：小孩念美國學校，去學校需要和老師溝通，而且日後陪小孩出國唸書也需要用到。

d. 興趣：就是喜歡英文，學這門課可增加國際觀。

f. 出國旅遊會用到，退休後學習出國可用到。

(三) 行政管理，技術支援，以及經費：

空中學院由於有數十年的遠距教學經驗，行政有課務，總務，註冊組分工；技術上有華視支援，經費除了有教材編制費，電視以及電台播出時皆可邀請專家學者或製作短劇，讓教學節目豐富有趣。

(四) 課程設計與教材編寫：

a. 五環式教學：空院課程採五環式教學，即電視、廣播、網路、函授及面授等方式進行。媒體教學的電視、廣播教學，由中華電視公司教學處及國立教育廣播電台、漢聲廣播電台負責專業製播。此外網路教學係因應近年來網際網路的蓬勃發展，逐漸成為遠距教學重要之一環，中華電視公司有鑑於此，特建立『網路學習中心』之網路教學環境，在 e-learning 的學習環境中，學習者透過網路即可閱讀教材參與學習，使遠距教學之學習空間更寬廣，可謂處處是教室，隨時可讀書。函授教學為基本方式，教材由中華電視公司教學事業處負責編輯並出版雙週刊，學生對於疑難問題可藉由通訊方式函詢。每月面授教學亦為重點方式，可增進師生互動，提昇學生學習效果。

b. 教學設計：為了滿足不同程度學生的需求，課程設計主要是由淺入深，內容豐富但不艱澀。主要從語言能力提升，技巧訓練，知識增加，實務演練，職場和證照考試和升學管道介紹多方面考量編寫教材。製作每週電視教學也常邀請學者及業界工作者。以下為我上下學期各十八周的教學課程設計：

英語口譯上：注重語言能力提升，技巧訓練，及實務演練。

第一週：Introduction

第二週：Pronunciation problems for Chinese students, part I

第三週：Pronunciation problems for Chinese students, part II

第四週：Public speaking skills

第五週：Shadowing

第六週：Reading aloud

第七週：Speech analysis, part I

第八週：Speech analysis, part II

第九週：Gist extraction

第十週：Paraphrasing

第十一週：Note-taking

第十二週：English numerical system

第十三週：Formulaic expression

第十四週：Idioms and proverbs

第十五週：Sight translation

第十六週：Short and long consecutive interpretation, Part I

第十七週：Short and long consecutive interpretation, Part II

第十八週：Conclusion

英語口譯下：理論介紹，知識增加，實務演練，職場和證照考試介紹。

- 第一週 :Introduction
- 第二週 :How to build up our English vocabulary
- 第三週 :To be faithful or not to be faithful
- 第四週 :to be elegant or not to be elegant: language register
- 第五週 : Tour interpreting, part I
- 第六週 : Tour interpreting, part II
- 第七週 : Tour interpreting, part III
- 第八週 : Escort interpreting
- 第九週 : Telephone interpreting
- 第十週: Whisper interpreting
- 第十一週: Liaison interpreting, part I
- 第十二週: Liaison interpreting, part II
- 第十三週: Conference interpreting, part I
- 第十四週: Conference interpreting, part II
- 第十五週: Business interpretation and negotiation
- 第十六週: Graduate institutes for translation and interpretation
- 第十七週: Accreditation
- 第十八週: Conclusion

測驗評量：上學期期中和期末考，通常會考短逐步以及中英數字轉換；下學期期中和期末考，考導覽口譯（tour interpreting）和視譯（Sight interpreting）。考前會給同學充分練習及準備。成績表現 95 優到 60 分甚之不及格皆有，然同學大致滿意接受此種考試。

期末課程評鑑：97 年調查，百分之九十五的同學認為本課程對於其工作或語言能力成長有助益，百分之五的同學認為本身程度太差，不適應。95%選修本課程的學生強力推薦此課程，以下為學生評語摘要：

1. It really helps me a lot! There are quite a few training activities to help students .
2. It's very professional!
3. It's unique! 挑戰潛能。
4. 內容很實用。
5. 教學生動，內容充實。
6. 可以讓自己的英文進步神速。
7. 上課活潑，有趣，課程實用。
8. 實戰演練，刺激緊張的一門課。
9. 讓自己從不敢開口說英文到輕鬆把英文說出來。
10. 我會推薦這門課，因為有趣。很棒！
11. 藉這門課幫助，使大家更有信心說英文。
12. 教法很互動，不會打瞌睡，而且增加了我口語能力，打開我對英文口譯的視野。
13. 非常實用，對生活上，工作上皆有幫助，也能獲得更多資訊，了解不同的文化，更有國際觀。
14. 沒學口譯，連自己發音對不對都不知道。要說的流利就一定要學口譯。
15. 充實豐富，這門課是語言學習路上的激勵大師。

16. 學英文一定要敢講, 一定要開口說, 不接觸口譯就不敢講。
17. 我會推薦大家選修“英語口譯”, 因為這門課不需死背, 而是靈活運用, 老師上課時, 會教我們很多口譯的技巧, 生動有趣的上課方式, 讓大家覺得“物超所值”。

結論

The secret garden of open and distance learning has become public, and many institutions in the UK are moving from single conventional-mode activity to dual-mode activity, that is to say offering a range of modes of study from the full/part-time and conventional/distance spectrum.

(Alan Tait)

從以上兩個案來看, 遠距教學有其方便實用性, 然 網路非同步翻譯教學師生滿意度皆不及空院口譯課。網路翻譯中, 主要可分為網路與翻譯兩部份, 然而從報告中可以看出老師、教材、學生都受限於網路, 技術的支援與突破重要性超過傳統的翻譯教材。此外, 網路翻譯教學缺乏師生面對面的互動, 如何使翻譯課程設計得活潑有趣(例如: 學生建議增加聲光圖片、遊戲軟體)是高難度的挑戰。教師除了翻譯本身的專業知識還須學習電腦、網路科技。非同步網路教學, 活動受到不少的限制, 應加強聲光多媒體, 讓活動多元化、趣味化。

空院口譯課教師可能遭遇之挑戰是大班教學, 然教師不須花費精神在行政技術上, 經教學設計後, 五環式專業教學, 即電視、廣播、網路、函授及面授等方式進行, 師生滿意度較單純網路教學高。其它成功原因包括經費較充裕、教材編輯豐富、硬體設備良好與行政技術支援極佳、師生每月見面、課堂間互動良好。

遠距教學無法取代傳統教學, 然兩者不須對立, 反之應相輔相成, 如此當可帶給學生無限夢想與希望, 更多的樂趣與成長。

參考文獻

張達聰(1989) *翻譯之原理與技巧*, 台北:國家出版社。

張培基等(1993) *英漢翻譯教程*, 台北:書林出版社。

柯平(1994) *英漢與漢英翻譯*, 台北:書林出版社

金莉華(1995) *翻譯學*, 台北:三民書局。

教育部(1999)「專科以上學校開辦遠距教學作業要點」。台(八八)電字地八八〇四九二三八號。台北:教育部。

<http://www.future-china.org.tw/fcn-tw/199907/1999072801.htm>

網路翻譯家網站 <http://www.hello.com.tw/~hengtong>

Johnson, Thomas (1955) ed. *The Poems of Emily Dickinson*. Cambridge UP.

Threlkeld, R., and Brzoska, K. (1994) "Research in Distance Education." *Distance Education Strategies and Tools*, ed. Barry Wills. New Jersey: Educational Technology Publications.

Tait, Alan. (1999). "The convergence of distance and conventional educational education: Some implications for policy." *The convergence of Distance and Conventional Education: Patterns of flexibility for the individual learner*. Ed. Alan Tait and Roger Mills. London and New York: Routledge Studies in Distance Education.

Early English Writing Activities in EFL Grade-one Classes in Taiwan

Shu-i Chang^a (張淑儀), Fang-chi Chang^a (張芳琪), Tien-cheng Shen^b (沈添鉦),
Hsiu-fen Hsu^c (許秀芬)

^a National Chiayi University

^b Diwan Universtiy,

^c Siang-He Elementary School

sic@mail.ncyu.edu.tw

engivy@mail.ncyu.edu.tw

victorshen1995@gmail.com

ceritafen@yahoo.com.tw

English writing has been a skill that is usually introduced after other language skills. In this study, however, we incorporated writing activities fairly early on in three beginning EFL classes for students to explore writing, not in a conventional way, by using limited English along with other symbols, such as drawing. In this report, we will present how early English writing activities were incorporated into the English curriculum as well as the findings from September 2006 to January 2007.

One general research question guided this study, that is, what can we learn about practicing writing activities in beginning EFL classes? There were 106 grade-one students included in the study from a public elementary school in southern Taiwan. Three English classes were scheduled each week as a regular English curriculum in this school. The incorporation of writing activities merely took 10 minutes or so each week, a considerably small portion of the English class. Throughout the semester, the writing activities were carried out in a two-week writing-sharing-instruction cycle.

In this study, we found encouraging results as follows. First, the differences were observed in the students' works over the time that showed their progress. Second, the change of the participating teacher's and students' attitudes toward writing was detected. Then, these young EFL students shared some similar developmental characteristics in writing with native-English-speaking emergent writers reported in Clay's study (1975). In addition, the students were found using code mixing in their writing. Learning to write is a long journey. We provided opportunities for the beginning EFL students to explore writing in their own ways while we were learning with them what they could do and how writing could be better incorporated into their English curriculum.

INTRODUCTION

Writing is considered more difficult than other language skills (Scott & Ytreberg, 1990). Therefore, for foreign language learners with limited English ability, it seems pretty reasonable not to introduce writing until they arrive at a certain level of English proficiency. In Taiwan where English is a foreign language, the majority of elementary English teachers generally place their focus on the teaching of spoken language. Not many teachers would offer instruction to and expect elementary students, particularly beginning EFL students, to be able to write.

Generally speaking, people seem to define “writing” as being able to put organized and coherent ideas together with accurate spelling and grammatical sentences. As a matter of fact, “writing” to many young children can take place in many different forms, including scribbling, drawing, writing letters or words, which should all be encouraged before they are taught to write conventionally (Schickedanz, 1986). Young language learners actually experience many developmental stages with different forms of writing before they become conventional writers, which may also be applied to EFL learners who start learning English at a fairly young age.

As pointed out by Paul (2003), “writing” should not be taken away from foreign language lessons simply because it is a relatively difficult skill compared to other language skills. For young EFL learners, teachers may start out with a helpful attitude toward children’s writing behavior. Schickedanz (1986), for instance, suggested teachers to look at writing the same way they look at children drawing or building blocks which provides them opportunities to explore and experiment.

Given opportunities, EFL learners with limited English ability could also write. Wright (2002, p. 5) used an example in his book to demonstrate how such learners can write:

big ... dog ...small...cat...run...jump...tree

Not using complete sentences, a child still could use single words to depict a story. Wright even encouraged EFL children to use drawing during writing in order to supplement ideas and make writing more fun. On the same line of argument, DeFord (1980) pointed out that language learners should be allowed to explore writing to have their concepts and strategies become sophisticated and refined to reflect conventions of the written language.

Presently, it is very rare that researchers or teachers in Taiwan would carry out studies regarding writing practice in elementary level because of various reasons. No studies so far have clearly illustrated what would happen if writing activities are introduced to beginning EFL learners, a practice that would be supported by scholars like Linse (2005), Paul (2003), Scott & Ytreberg (1990), and so on. Therefore, in September 2006, we conducted a study in which we designed and implemented writing activities in three first-grade classes to find out what beginning young EFL students could do with writing. In this report, we will focus on reporting the students’ writing performance and the participating teacher and students’ reactions to the writing activities between September 2006 and February 2007.

THE CONTEXT OF THE STUDY

The study took place at Siang-he Elementary School located in Chiayi County. Siang-he is a fairly young public school founded in 2003. One of the unique characteristics about Siang-he is its English education that starts in grade one. Three English classes are scheduled each week for all grades at this school. In the participating classes, two periods of English class were co-taught by Cerita, a native speaker of Mandarin and Myra, a native English speaker without knowledge of Chinese, while one period was taught by Cerita alone. Cerita was solely responsible for carrying out the practice of writing activities for 10 to 15 minutes in one of the periods, while Myra helped with videotaping and some classroom management.

The participants in this study included four college teachers as researchers³ and one qualified elementary English teacher, Cerita, along with 106 first graders in three different classes. The materials used were a song book, a self-developed textbook (mainly for alphabet learning) and a self-developed workbook. Of the scheduled three periods for English, only ten minutes or so in one period was used for the planned writing activities.

³ One of the teachers is not engaged in the composition of this report; however, her involvement in the design of the writing practice is highly appreciated.

The Incorporation of Writing Activities

As mentioned earlier, there were no existing empirical studies with suggestions as to how English writing activities can be designed and implemented for early EFL learners. We had to rely on our professional knowledge to design what we thought would be appropriate for beginning EFL students. During the process we had to be flexible about the design since information from fieldnotes, videotaped classes, and discussions with Cerita would inform us whether our design was suitable to the students. Throughout the semester, nine writing activities were carried out, which means, generally each student from three classes engaged in writing for nine times.

The pattern for carrying out the writing activities was that the students would be engaged in writing the first week followed by written work sharing and scaffolding instruction from Cerita the next week, which formed a two-week writing-sharing-instruction cycle for nine times within a semester. These nine cycles can be divided into two periods because of different natures in writing topics, with cycles one to six in the first period and cycles seven to nine in the second period.

In the first period, the students were told to write or draw freely about something they learned in English class. In other words, no specific topics were given to the students to work on. They could recall what they just learned before the writing activity or what they learned in previous classes. In the second period, the students were given specific topics selected from the song book. For example, a topic of “Four Seasons” was given to the students to write about what they knew or how they felt about four seasons. In this period, though specific topics were given, the students were still allowed to write and/or draw something not related to the given topic if they could not produce anything about the topic.

Right from the beginning, the students were told not to copy anything onto their papers; however, this rule was not strictly practiced at first and the students would sometimes copy words from neighboring students or the environment. Gradually, the students realized that there were no standard ways to write and they were given freedom to produce their work in various forms, so it was not necessary for them to copy from others or the environment at all.

The students’ works were examined in order to select some samples to be shared with proper scaffolding instruction in the week following the writing activity. A variety of samples were selected to represent different language abilities, for example, samples with pictures only, with letters, with words, and so on. Scanned works were displayed to elicit dialogues between Cerita and the students regarding the contents of the works. For each work, Cerita would provide scaffolding instruction to encourage the students who produced similar work to try to move a step further in their next writing. For example, for those who drew pictures (an apple, for instance) only, Cerita would encourage them to produce letters (“a” for an apple, for instance) to go with pictures next time; for those who wrote letters (“a,” for “apple,” for instance), Cerita would encourage them to write other letters in the word (“ap” for “apple,” for instance). Such scaffolding instruction provided the students with opportunities to improve their next writing at least a step further than what they could currently do.

Data Collection and Analysis

Between September 2006 and February 2007, nine writing activities were carried out in three first-grade classes with about 106 students at Siang-he Elementary School. There were approximately 950 copies of written work collected and examined. This helped with our understanding about what the students could do and their progress in writing.

Information from fieldnotes, recorded video clips, and informal chats and discussions with the English teacher were also collected and analyzed. This helped us learn more about what happened in the classrooms and how the participating teacher and students reacted to the incorporation of writing activities.

FINDINGS AND DISCUSSION

A few interesting and encouraging phenomena were found in this study. First of all, it is the differences observed on the students' production that showed their progress. Second, the change of the participating teacher's and students' attitudes toward writing was detected. Then, these young EFL students shared some similar developmental characteristics in writing with native-English-speaking emergent writers reported in Clay's study (1975). In addition, the students were found using code mixing in their writing. "The change of attitudes" will be reported first below since without the attitude change, the study would not have continued for an entire semester and further gone into its third year now.

Changing attitudes

Both the participating teacher and students did not have much experience in English writing; therefore, they did not hold positive attitudes towards the incorporation of the writing activities at the beginning. When the researchers first suggested to Cerita to integrate writing into the curriculum, she hesitated. She agreed to implement the plan in only one single class of about 35 first graders rather than all her three classes of about 106 students. However, to keep the same weekly schedule, she had no choice but hesitantly accepted this proposal.

The students also had various worries when they were asked to write in English. In Mandarin, they expressed their concerns and asked questions:

"I only learn a little. How can I write?"

"We just start learning. Why are we having a test?"

"Is this a test?"

"May I copy from the book?"

"I don't know what to draw."

"If I don't do well, what will be the punishment?"

"May I ask my neighbors [questions]?"

"If I don't do well, how many times should I copy the words?"

Most of the students then either looked around to copy from the environment (such as words on the wall, pencil, eraser, pencil case, etc.) or neighbors, or peeked into the textbook to copy something out of it. Some drew or wrote a little; some ignored this task, switching attention to something else that was more fun to them, such as playing with their pens, hands, toys, etc.

Six weeks later, after completing Writing Three, Cerita expressed her and the students' positive attitude towards the writing activity. According to her, the students showed their excitement when it was time to write. They liked to know whose sheets would be shown on the TV screen. They also were eager to have their works shown and shared in class. Cerita said when she announced "writing time," quite some learners shouted "Hurray." We also observed these scenes on the recorded videos or in person. The attitude change was obvious, which encouraged all of us engaged in the study.

In this study, it was too early to conclude the contribution of the writing component to students' learning achievements. However, adding the writing component seemed to have some positive effects on learning. It was found that the students became more productive and their written works became more meaningful over the course of the study. There were several observable stages of writing development (from drawing to expressing meanings to initial letter writing to phonics spelling to word writing) that will be illustrated below.

Becoming more productive and producing more interpretable messages

At the beginning of the writing activities, some students either scribbled their names in English and left their paper blank or had a few drawings. Gradually, drawings, letters, words and/or sentences appeared. They became more productive in the following writing tasks. Figure 1 presents Ben's productions as an example. Ben Started out with two pictures (Writing 1) to one letter of its capital and lower case (Writing 2), to a list of letters (Writing 3) and to sentences (Writings 8 and 9), which revealed Ben's gradual progress. During the

writing process, Ben seemed to become more productive and the content contained more interpretable messages.

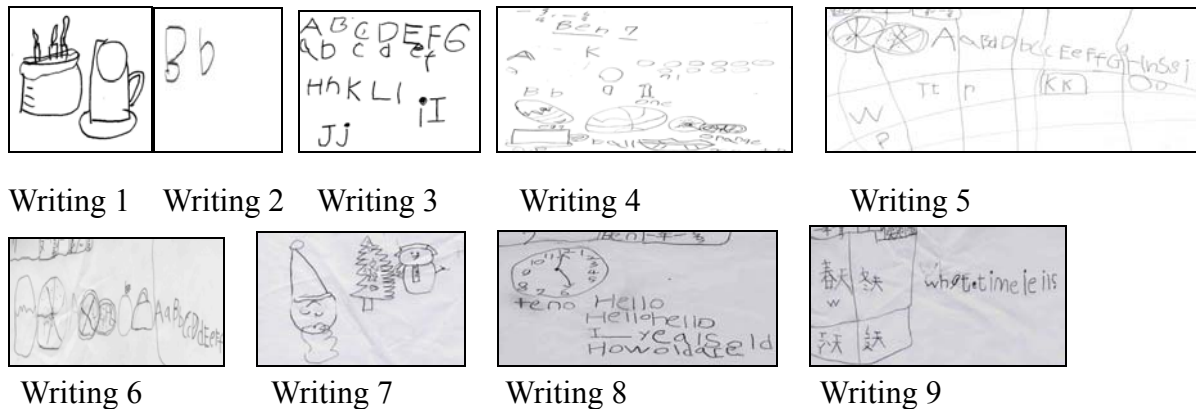


Figure 1 Ben's Productions

The observable stages of writing development

Spelling development was one obvious element that was observed among students to progress from the preliminary stage to more advanced stage. Hill (1999) proposed several stages of spelling development that are not necessarily manifest sequentially among children, including *prephonic spelling* (drawings/symbols), *semiphonetic spelling* (one or two letters representing a word), *phonetic spelling* (inventing spellings phonetically based on the speech sounds), *transitional spelling* (using more common letter patterns in words), and *independent spelling*.

An abundance of data was found in the current study to represent the first three stages, i.e. prephonic spelling, semiphonetic spelling, and phonetic spelling. For instance, Ben's productions presented in Figure 1 show some development from drawing (prephonic spelling—Stage 1) to one single letter writing (semiphonetic spelling—Stage 2) and then skipping other stages to sentence writing at the end of the semester. In Ben's production (Figure 1), we see the gap between his letter writing and sentence writing. The nearly conventional sentence production on Writings 8 and 9 might come from Ben's rote memory since those sentences were practiced frequently in class. Spelling is of course much more than just memorizing a predetermined set of words. However, remembering language items is rather important because it could help students "bridge the gap between language study and language activation" (Hamer, 2001, p. 165).

In Ben's case, we did not find a linear spelling development from stages one to three, which is normal in many children's spelling development. Meanwhile, in many other emergent writers' works in this study, it was found that their spellings exemplified the features of the spellings suggested by Hill in the first three stages of development, also not in the manner of linear development. Quite some students, for example, started out drawing (Stage 1) and scribbling letters (Stage 2) and words of invented spelling (Stage 3). It is also rather common to see students write initial letters of the pictures. When students wrote words, it was common to see that words were spelled with correct initial and final consonant letters and incorrect vowels in the middle (e.g. *foll* for *fall*, *het* for *hat*, *pancil* for *pencil*, etc.). In this particular setting, the students were given opportunities to write in their own way through which they explored spellings at their own pace; their spelling development was observable.

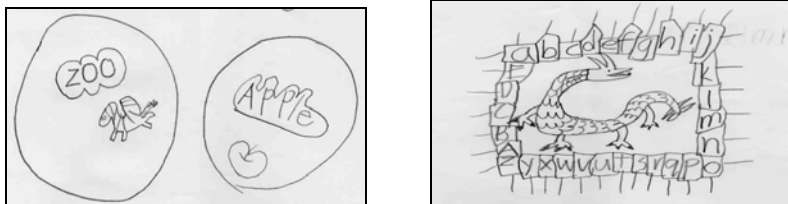
The Characteristics of the Writing by Emergent ENL and EFL Writers

Some of the characteristics reported by Clay (1975) on the writing of young native English speakers were observed on the writing of these young foreign language students in the current study. Emergent writers usually carry some concepts and apply some principles

when they start writing.

It is found that the principles of directional and inventory (Clay, 1975) were applied by the young EFL students in the study. The directional principle, writing from the top-left position, could be an influence by the exposure to both Mandarin and English print. Interestingly, it is found that some writing was done from left to right and supplemented by the downward movement before exceeding the right margin of the paper. The same phenomenon that occurred to native-English-speaking children was also reported by Clay. Some students in the study were found to write English at a vertical direction, which could be a strong influence by Chinese writing, a common practice for the students to practice writing Chinese characters vertically. However, students who placed their drawing or writing at the center of the page seemed to hold different principles for writing and drawing, which is beyond our current understanding.

Furthermore, young writers turn letters around, decorate them and evolve new signs as they explore the limits within which a sign can vary (Clay, 1975), which also occurred in the EFL students' works. Word decoration, for example, is seen in Figure 2-a and letters can even be used as décor to create some image in Figure 2-b to tell a story. The student who created the work 2-b said in a later interview that he made a cage to block the dragon so the dragon would not come out.



a. Decorated words

b. Decorating letters

Figure 2 Decorated and decorating letters/words

Lists of letters, words and sentences appeared frequently in students' works, an inventory principle, according to Clay (1975), applied by the students. The inventory principle has its value (Clay, 1975). A young writer can list all the words he can produce without copying, though sometimes these inventories are actually copied lists. They can be proud of themselves, as if they were saying "These are all the words I can write (Clay, 1975, p.32)." The young writer may be able to systematize those items which he can recall and which have become part of his written repertoire.

According to Clay, one type of this structure is to group all the members of a set such as the alphabet (writing out all the letters in sequence) or the members of a limited set—family, friends, words beginning with a particular letter, words ending similarly and so on. Ben's work (Writing 3 in Figure 1) is an example for the letters in sequence. Such examples were also found in many other students' works.

Mixing Codes in Writing

Since the students were allowed to produce their works freely it is found that many of them would put various combinations of symbols on papers, including English letters/words, Mandarin phonemic symbols, Chinese characters, and drawings, which is considered a characteristic of writing for beginning writers. Figure 3 presents an example of the combination of drawings, initial letters, English words (invented spelling and conventional spelling), Chinese characters, and Mandarin phonemic symbols.

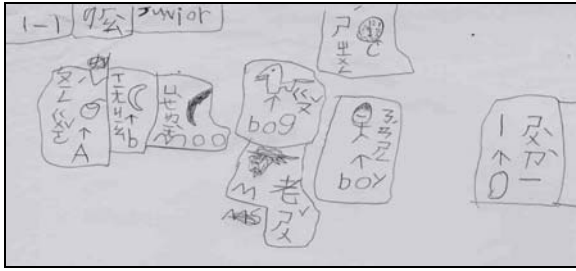


Figure 3 Code mixing

Mixing codes from different languages (e.g. English and Mandarin) and different writing systems (e.g. drawing, word/character, language-specific phonemic symbol) is a characteristic observed in the written data collected in this study. In a context where writing is a process for students to explore what they could do and what they wanted to do, we found that code mixing was employed by the majority of the students. This phenomenon did not alarm us and we allowed the students to continue producing their works using any symbols they wanted. Learning to write is a long process; it was not necessary to limit students to writing conventionally when they first encountered a foreign language.

CONCLUSION

Writing, like other language skills, could be a part of English curriculum, even for beginning learners. When teachers expect young learners to be able to write, they provide opportunities and assistance to encourage and foster writing behavior through which teachers observe how learners behave, which in turn informs further instruction. On the contrary, when teachers do not expect young learners to be writers, they probably would place focus on other language skills instead of writing and therefore writing would not be likely to automatically take place among learners. Naturally, teachers' experiences with young learners shape their teaching practices, which in certain situations may prevent them from taking new challenges, leaving learners' potentials uncultivated.

In the current project, we challenged the habitual practice of leaving writing out at the beginning stage for young foreign language learners and persuaded one bold elementary school teacher to also step out of her comfort zone to integrate writing into her classes. We cannot pinpoint what exact contributions to the whole language learning were made by the addition of the writing component. However, what had been observed revealed positive outcome that encouraged the continuation of practice with these young EFL students to further explore the world of writing.

Given opportunities with expectations and assistance to write, these young EFL students drew, scribbled, copied, invented and wrote letters, words, and/or sentences. Their attitudes toward writing turned positive; their writing development was observed and was probably even accelerated.

According to Clay (1993), while children explore writing they will

- a) attend closely to the features of letters and to learning letters;
- b) construct 'their own words', letter by letter;
- c) direct attention to special features like serial order and spaced between words;
- d) work within the order and sequence rules of print, revealing these to themselves while constructing messages;
- e) break down the task to its smallest segments while at the same time synthesizing them into words and sentences;
- f) engage in their own form of segmenting sounds in words in order to write them.

The result from the current project adds support to Clay's findings on her observation of

native-English-speaking emergent writers' development of writing. In addition, the findings of this study seem to also reveal the similarity between the writing development of native-English-speaking emergent writers and that of English-as-a-foreign-language emergent writers.

As mentioned earlier, learning to write is a long journey. We provided opportunities for the beginning EFL students to explore writing in their own ways while we were learning with them what they could do and how writing could be better incorporated into their English curriculum. This journey just began and the encouraging findings from this study motivated all of us to continue our further expedition in the world of writing.

REFERENCES

- Brewster, J., Ellis, G., & Girard, B. (2002). *The primary English teacher's guide*. Essex, England: Pearson Education Limited.
- Clay, M. M. (1993). *Reading recovery: A guidebook for teachers in training*. Auckland: Heinemann.
- Clay, M. M. (1975). *What did I write? Beginning writing behavior*. Portsmouth, NH: Heinemann.
- Clay, M. Marie. (2001). *Change over time in children's literacy development*. Portsmouth, NH: Heinemann.
- DeFord, D. E. (1980). *Young children and their writing*. *Theory into Practice*, 13(3), pp.157-162.
- Harmer, J. (2001). (3rd ed.). *The practice of English language teaching*. NY: Longman.
- Hill, S. (1999). *Phonics*. Armadale Vic, Australia: Eleanor Curtain Publishing.
- Jackson, A., Ralphs, W., Clark, S. (2002). *Music world for children*. Taipei: Ur Best.
- Krashen, S. (1985). *The input hypothesis: Issues and implications*. New York: Longman.
- Paul, D. (2003). *Teaching English to children in Asia*. Quarry, Hong Kong: Longman Asia ELT.
- Linse, C. T. (2005) *Practical English Language teaching young learners*. NY: McGraw Hill.
- Schickedanz, J. A. (1986). *More than the ABCs: The early stages of reading and writing*. Washington DC: NAEYC.
- Scott, W. A. & Ytreberg, L. H. (1990). *Teaching English to children*. NY: Longman.
- Wright, A. (2002). (4th impression). *Creating stories with children*. Oxford: Oxford University Press.

Professors' Use of Questions in English Lectures: An Interdisciplinary Comparison

Yu-ying Chang (張玉櫻), Wei-ling Liu (劉緯玲)

Yuan Ze University

yymeichu@yahoo.com.tw

Although the use of questions to facilitate interaction in academic lectures at tertiary education has been investigated in several previous studies (e.g., Morell, 2004; Crawford Camiciottoli, 2008), the issue how disciplinary cultures influence the use of questions in these lectures has received little attention. Thompson's study (1998) was the only interdisciplinary comparison of questioning in academic talks that we found during our library search. However, the data used in her study lumped together two different genres (i.e. lectures and research presentations), thus could not really provide us with a clear picture regarding the disciplinary differences in the use of questions in English lectures.

Therefore, this study aims to focus on the investigation of interdisciplinary differences in professors' use of questions in terms of question form and function. The corpus used in this study includes 15 small English lectures (i.e. classes with less than 40 students) from three academic divisions—Humanities & Arts (HA), Social Sciences & Education (SS) and Physical Sciences & Engineering (PS). The results of this study reveal that the two most common question patterns used in the three divisions are generally the same: (1) questions in the form of *declarative/Imperative + word tag* used to check students' lecture comprehension and (2) *Wh-questions* used to elicit students' responses. However, interdisciplinary differences in question use were also found. The most obvious distinction is that compared to their colleagues in hard fields, professors from the soft fields in HA and SS seem to ask more questions which serve to build closer rapport with their students and to enhance the classroom management. On the other hand, professors in the hard fields in PS appear to use questions as attention focalizing device to highlight upcoming new information more often than their soft field counterparts.

INTRODUCTION

Due to the trend of internationalization in higher education and the increasingly widespread use of English as a medium for instruction at the tertiary level in non-native English-speaking countries, spoken academic English has recently gained far more importance in the research of English for Academic Purposes (EAP). Among the various genres of spoken academic English, lectures have been considered the most prominent format of instructional activity in higher education worldwide (Bamford, 2005; Fortanet, 2004a), hence attracting the greatest amount of EAP research attention.

Previous EAP and applied linguistics studies on lectures have distinguished various styles of lectures. For example, Dudley-Evens and Johns (1981) provided a distinction among *reading style*, *conversation style*, and *rhetoric style* (Dudley-Evens & Johns, 1981) while Goffman (1981) provided one among *memorization style*, *aloud reading style* and *fresh talk style*. Among the various styles, it has been observed that the adoption of a more interactive style (i.e. a more informal and conversational style) is the current trend (Bamford, 2005; Flowerdew, 1994; Fortanet, 2004a). The informal, conversational style of lecture has been found to allow for higher degree of interactions between the lecturer and the audience

(Crawford Camiciottoli, 2004, 2008; DeCarrico & Nattinger, 1988; Dudley–Evans, 1996; Morell, 2007). In addition, various linguistic features enhancing the degree of informality and interaction have been reported to have positive effects on non-native English-speaking students' lecture comprehension (Crawford Camiciottoli, 2004) and on their learning of subject content (Northcott, 2001).

Given these findings, an increasing number of applied linguistics studies focusing on the investigation of how various linguistic and textual features are used to enhance the degree of lecture interactivity has been appeared. The linguistic and textual features examined so far include personal pronouns (e.g., Fortanet, 2004b; Morell, 2001; Rounds, 1987a, 1987b), modality (Crawford Camiciottoli, 2003), interactive lexio-syntactic patterns (Crawford Camiciottoli, 2004), discourse markers (Morell, 2001) and questions (Bamford, 2005; Crawford Camiciottoli, 2008; Fortanet, 2004a; Liu and Chang, 2009;Thompson, 1998).

Among the various linguistic devices, questions seem to be among the most obvious ones to invoke human interaction in general because a question presupposes an answer (Goody, 1978); in order to form a complete proposition, both question and answer are necessary (Bamford, 2005). Additionally, questions have long been recognized as an important interactional device employed by teachers to activate and facilitate the teaching and learning processes (Crawford Camiciottoli, 2008; Gall, 1970; Launspach, 1998; Long & Sato, 1983). Aschner (1961) even called the teacher “a professional question maker” and indicated that asking questions is “one of the basic ways by which the teacher stimulates student thinking and learning” (originally cited in Gall, 1970).

Due to this prominent role of question in educational settings, studies on the use of questions in classrooms have been conducted by scholars in many different disciplines. However, it is rather surprising to see that, till present, the use of questions in English academic lectures at tertiary level has attracted the attention of only few EAP scholars (e.g. Bamford, 2005; Crawford Camiciottoli, 2008; Csomay, 2002; Fortanet, 2004a; Morell, 2004;; Liu and Chang 2009 ; Thompson, 1998;). Among the studies located during our library search, Csomay (2002), Fortanet (2004a), and Morell (2004) only included questions as one of the interactive features in their text analyses, rather than focusing specifically on the examination of the use of questions in lectures. Thompson (1998), although focused on the analysis of questions, used a mixed corpus consisting of both academic lectures and presentations. Because her focus was the use of questions in academic talks in general, she did not distinguish the differences in the use of questions in these two genres. We will discuss this study in more detail later,

On the other hand, Crawford Camiciottoli (2008) compared the use of questions in academic lectures and written instructional materials (including print and on-line textbooks as well as on-line learning resources) collected from business studies. Her main purpose was to observe how communicative modes influence the use of questions. The results of her study revealed that although the frequency of questions was strikingly similar in these two sets of data, there was marked variation in question form and function across them.

Bamford (2005) and Liu and Chang (2009) are the only two studies we managed to find that focus specially on the use of questions in English academic lectures. Bamford (2005) examined the functions of question/answer adjacency pairs in academic lectures in economics. She pointed out that question/answer sequences in academic lectures are characterized by one interactant (i.e. the lecturer) performing both the questioning and the answering roles; that is, although the lecturer already possesses the information the question posits, she/he still asks the question and provides her/his students with the answer. Based on this observation, her major effort was to illustrate how lecturers use their control of both question and answer as effective attention focalizing mechanism and how the echoing of the prosody of spotaneous conversation question/answer sequences can create an atmosphere of interactive sharing of

ideas and information with the audience.

Liu and Chang (2009) investigated gender influences on professors' use of questions in English lectures based on an analysis of 12 small lectures collected in the Michigan Corpus of Academic Spoken English (MICASE). Their results showed distinct differences between female and male professors in their question use. For example, it was found that female professors tended to ask far more audience-oriented questions than their male counterparts to involve the students' participation both literally and metaphorically. On the other hand, male professors asked remarkably more content-oriented questions in order to control the development of the new topic; when asking such questions they seemed not expect response from the audience because they often answered the questions themselves.

From the above review, we can observe that although many EAP studies have revealed the important influence of discipline on discourse choices (Hyland, 2000), it is rather unexpected to see that very few researchers have paid their attention to interdisciplinary variation in the use of question in academic lectures. In fact, Thompson's study (1998) was the only interdisciplinary comparison found during our library search. Thompson (1998) compared the use of questions in academic talks in general (i.e. lectures and research presentations) in applied linguistics and applied sciences. Her results show that compared with applied sciences, applied linguistics seems to allow more space for the audience to take an active part in academic talks. According to Thompson (1998, p.147), one probable explanation for this difference lies in the organization of the linguistic and science talks: The applied science lecturers and presenters put more emphasis on "telling the audience about scientific entities and process, experimental procedures and the results of experimental work." In contrast, the applied linguistics ones emphasize more on raising linguistic issues and problems by creating a dialogic effect to discuss these issues and problems.

Thompson's (1998) study is noteworthy for the fact that comparative studies on spoken academic discourse have received relative little attention in EAP research. However, she did not specify the size of the lectures and presentations in her study. As evident in previous research, the number of participants effects the interaction in lectures (Hansen & Jensen, 1994). In addition, the data used in her study covered two different academic spoken genres; in her analysis; she did not distinguish the differences in the use of questions in these two genres. Given this indiscrimination of data belonging to different genres plus the common awareness of possible genre differences in EAP discourse (Swales, 1990), we can suspect that the results of her study cannot really provide us with a clear and appropriate picture regarding how questions are used in lectures in different academic disciplines. Therefore, the purpose of this study is to focus on the investigation of interdisciplinary differences in the use of question forms and functions by professors from three academic divisions in their small English lectures (i.e. classes with no more than 40 students).

METHOD

The Corpus

Being aware of the potential influence of class size to lecture interaction, the data of this study consist of the transcripts of 15 small lectures from MICASE. Based on the definition given in MICASE, small lectures are those given in classes with no more than 40 students. These small lectures were collected from three academic divisions—Humanities & Arts (HA), Social Sciences & Education (SS), and Physical Sciences & Engineering (PS)—as defined in MICASE Manual. In addition to class size, when selecting lecture samples from MICASE, we also tried to balance the distribution of the different participant levels (cf. Hansen & Jensen, 1994) and gender of the lecturer. At the end, for each division, five small lectures were selected, three of which were given by male professors and two were given by female professors. These lecturers were all native speakers of American English and the student

audiences were also composed mostly of native speakers of American English.

Table 1

MICASE small lecture corpus

Academic division	File	Topic	Gender	Participant	Words	Length (min.)
Humanities & Arts (HA)	1	Graduate Online Search and Database	F	Junior Graduate Students	14,910	147
	2	Visual Source	F	Senior Undergraduates	8,299	69
	3	Beethoven Lecture	M	Mixed Graduates	6,747	75
	4	American Literature	M	Senior Undergraduates	13,514	99
	5	Historical Linguistics	M	Senior Undergraduates	12,117	69
Social Sciences & Education (SS)	6	Ethics Issues in Journalism	M	Senior Undergraduates	5,662	83
	7	Labor Economics	M	Junior Graduate Students	11,204	77
	8	Statistics in Social Sciences	F	Mixed Graduates	13,831	109
	9	Honors Intro Psychology	M	Junior Graduate Students	5,430	49
	10	Intro to Psychopathology	F	Senior Undergraduates	7,863	52
Physical Sciences & Engineering (PS)	11	Graduate Physics	M	Mixed Graduates	11,757	105
	12	Intro to Groundwater Hydrology	F	Junior Graduate Students	12,295	82
	13	Radiological Health Engineering	F	Senior Undergraduates	11,865	98
	14	Professional Mechanical Engineering	M	Junior Graduate Students	12,013	90
	15	Intro Programming	M	Senior Undergraduates	7,442	50

The Analysis

Questions can be categorized on a formal level according to their structural characteristics or linguistic forms. They can also be defined on a functional level, based on the speakers' intentions when uttering a question (Athanasidou, 1991). Since both aspects are of interest to this study, the transcripts were searched manually with the reference of their corresponding sound files to determine both the form and function of the questions asked by the lecturers.

Classifications of questions in previous studies

In Thompson's (1998), questions were classified as having two different orientations. The first type is audience-oriented questions, where the audience is at least symbolically expected to respond the lecturer's questions. The other type is content-oriented questions, in which no audience response occurs or seems to be expected. Audience-oriented questions were further subcategorized by Thompson into three functions: (1) to *check* if the audience can perceive and understand the speaker's utterances, (2) to *evoke audience response* and (3) to *seek audience agreement*. Content-oriented questions, on the other hand, were subcategorized into two functions: (1) to *raise issues* or (2) to *introduce information*. Later in Fortanet's (2004a) study on university law lectures, he proposed two question types—rhetorical questions, which expect no response; non-rhetorical questions, for which an answer is feasible.

Following Thompson (1998), Crawford Camiciottoli (2008) also categorized questions into audience-oriented and content-oriented. In Crawford Camiciottoli's (2008) comparative study on the use of questions by lecturers and material writers, three kinds of question functions were identified to be audience-oriented: *eliciting response*, *requesting confirmation/clarification*, *soliciting agreement*. On the other hand, the functional subcategories under content-oriented question are *focusing information* and *stimulating thought*. Finally, in Liu and Chang's (2009) study on how lecturer's gender influences the use of questions, Liu and Chang (2009) adapted Crawford Camiciottoli's (2008) taxonomy of question functions by adding two additional subcategories (*classroom*

management/engagement and checking comprehension) under audience-oriented questions.

The taxonomy of question forms in this study

The analysis of the question forms in this study was based on Liu and Chang (2009). Question forms were classified into six categories (based on the lexical and syntactic features of the questions): *Wh-question*, *Yes/no question*, *Tag question*, *Declarative/imperative+ word tag*, *Alternative question* and *Incomplete question* (please see Table 2 for examples). For the rationale of this classification, please see Liu and Chang (2009).

One thing should be noted here is that tag question in this study refers to the canonical type of tag question with reversed or constant polarity (e.g., *students are tough aren't they; so this is the letter he sent you is it?*), while *declarative/imperative+word tag* is formed by statement followed by a word such as *ok, alright, right*. Finally, an incomplete question refers to an incomplete utterance ends with a pause, which hints the audience to complete the utterance.

Table 2

The taxonomy of question forms

Questions Forms	Examples
1. Wh-question	what phases are you comparing? (<i>PS—Lecture 11</i>)
2. Yes/no question	you wanna stop it here? (<i>SS—Lecture 10</i>)
3. Tag question	it's very moving isn't it? (<i>HA—Lecture 5</i>)
4. Declarative/Imperative+word tag	well it depends on a lot of things right? (<i>SS—Lecture 7</i>) and notice also that it is a fountain. okay? (<i>HA—Lecture 2</i>)
5. Alternative question	is it cold or warm? (<i>HA—Lecture 1</i>)
6. Incomplete question	oh really you were? (<i>HA—Lecture 2</i>)

The taxonomy of question functions in this study

Similarly, in this study, we adopted Liu and Chang's (2009) taxonomy of question functions (see Figure 1 for illustration). The following is a brief description for each of the functions.

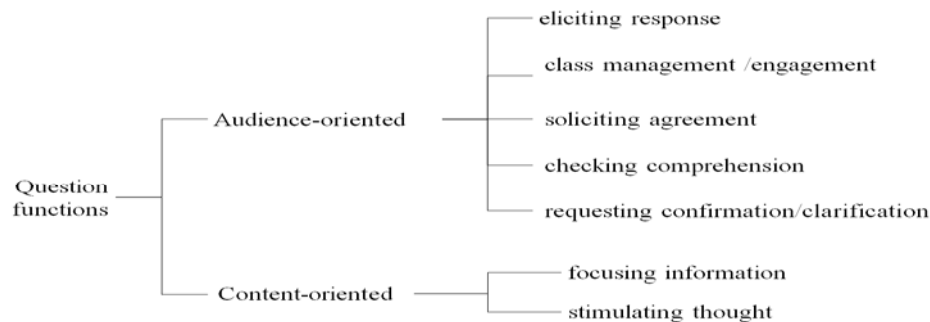


Figure 1 The taxonomy of question functions (adapted from Liu and Chang, 2009, p.7)

A. Audience-oriented questions

1. Eliciting response: Liu and Chang (2009) defined this type of questions as questions used to invite the audience to supply a piece of information related to the course content. The following are two examples:

what are some examples of the specifics, when dealing with photographs? (*SS—Lecture 6*)

what kind of a cemetery might it have been? (*HA—Lecture 2*)

2. Class management/ engagement: In contrast with questions to elicit response, when asking class management/engagement question, the lecturer does not expect the audience to answer the question based on course content; instead, this type of questions is used to

manage the class in order to enhance teaching flow (example 3) or as means of building rapport or increasing interaction with the student (example 4).

well one question at a time alright? (PS—Lecture 11)

do people wanna take a five minute bathroom break? (SS—Lecture 10)

3. Soliciting agreement: This refers to the type of questions used by lecturers when they attempt to appeal to the student to agree with their propositions (Thompson, 1998, p. 141). Example 5 shows that the lecturer assumes a certain statement to be true thus is inviting the students to confirm this.

that would be the other extreme right? (SS—Lecture 6)

4. Checking comprehension: This type of questions is typically formed by word tags (e.g., *okay?*, *alright?*) or by utterances like *Do you understand?*; they are questions used to ensure whether students are able to understand the lecturer's message.

everybody get that idea ? (PS—Lecture 12)

in other words it actually occurs on both sides of the linguistic border. alright? (HA—Lecture 6)

5. Requesting confirmation/clarification: This type of questions are used to check if the lecturer has correctly or incorrectly understood/heard the student's previous utterance (i.e. requesting confirmation; example 8) or used to request the student's reiteration of a previous utterance (i.e. requesting clarification; example 9).

did- did you say "distance"? (SS—Lecture 6)

student: uh... two-thirty-eight

lecturer: pardon me?

student: two-thirty-eight (PS—Lecture 13)

B. Content-oriented questions

Contrary to audience-oriented questions, in which the audience is at least symbolically given an opportunity to provide an actual verbal or non-verbal response, content-oriented questions are the ones that require no immediate response from the audience (Thompson, 1998, p. 143). The two major functions of content-oriented questions are *focusing information* and *stimulating thoughts*.

1. Focusing information: This is a type of questions that is introduced and immediately answered by the lecturer, for example:

what's gonna happen here? it's gonna start killing the signal. (PS—Lecture 11)

you have an incident photon it disappears it must take place, uh, in the electric field of a nucleus why? because the nucleus has to, conserve momentum and get a little bit of, recoil. (PS—Lecture 13)

2. Stimulating thought: This type of questions appears to be 'big' questions or issues to which there is no easy answer. The lecturer does not immediately provide an explicit answer to his or her own question, though a commentary on or evaluation of the question may be provided to encourage audience's reflection (example 12).

are there linguistic change that can be attributed to the, to language contact? This is a fairly controversial uh area within the area of Romance linguistics...(HA—Lecture 5)

RESULTS AND DISCUSSION

In this section, the results will be reported based on the overall frequency, forms and functions of the questions found in the data. The results are displayed as raw totals (N) and as frequency of occurrences per 1,000 words.

Interdisciplinary differences in question frequency

As shown in Table 3, significant variation can be found in the frequency of question occurrences. In terms of average, while HA professors asked questions most frequently (13.3

questions per 1000), those in SS had the lowest frequency of questions (8.6 questions per 1000); in between the two extremes was PS with a frequency of 9.9 questions per 1000 words. However, if we scrutinize the data shown in Table 3, we can also observe a great variation among individual lecturers in each of the divisions. Both the highest question frequency (26.8 questions per 1000 words in Journalism) and the lowest frequency (1.8 questions per 1,000 words in Psychopathology) were found in the SS division. This finding implies that the interpretation of the results pertaining to question frequency should be made carefully. Future studies with larger sample size are needed to clarify this issue.

Table 3
Frequency of questions in MICASE small lecture

Academic	Topic	Words	N	(per 1000)
Humanities & Arts (HA)	Graduate Online Search and Database	14,910	355	23.8
	Visual Source	8,299	138	16.6
	Beethoven Lecture	6,747	90	13.3
	American Literature	13,514	98	7.3
	Historical Linguistics	12,117	69	5.7
Average				13.3
Social Sciences & Education (SS)	Ethics Issues in Journalism	5,662	152	26.8
	Labor Economics	11,204	102	9.1
	Statistics in Social Sciences	13,831	44	3.2
	Honors Intro Psychology	5,430	12	2.2
	Intro to Psychopathology	7,863	14	1.8
Average				8.6
Physical Sciences & Engineering (PS)	Graduate Physics	11,757	222	18.9
	Intro to Groundwater Hydrology	12,295	202	16.4
	Radiological Health Engineering	11,865	86	7.2
	Professional Mechanical Engineering	12,013	53	4.4
	Intro Programming	7,442	18	2.4
Average				9.9

Interdisciplinary differences in the distribution of question forms

Table 4
Interdisciplinary difference in the distribution of question forms

	Humanities & Arts	Social Sciences & Education	Physical Sciences & Engineering
Yes/no question	310 41.3%	Declarative/imperative +word tag 122 37.7%	Wh-question 238 41.0%
Wh-question	235 31.3%	Wh-question 96 29.6%	Declarative/imperative +word tag 166 28.6%
Declarative/imperative +word tag	182 24.3%	Yes/no question 90 27.8%	Yes/no question 162 27.9%
Tag question	11 1.5%	Tag question 14 4.3%	Tag question 6 1.0%
Alternative question	7 0.9%	Alternative question 2 0.6%	Alternative question 5 0.9%
Incomplete question	5 0.7%	Incomplete question 0 0%	Incomplete question 4 0.7%
Total	750 100%	Total 324 100%	Total 581 100%

As shown in Table 4, *wh-question*, *yes/no question* and *declarative/imperative+ word tag* are the three most common question forms used in all the three divisions; they constitute more than 95% of all the question forms in this study. However, their frequency orders were

not identical in each of the divisions. In HA, the form used the most frequently was *yes/no question*, whereas in SS, it was *declarative/imperative+ word tag*, and in PS, *wh-question*. As for the other question forms are concerned, they seemed to be used rather sporadically in academic lectures in general.

Interdisciplinary differences in the distribution of question functions

The distribution of question functions across the three divisions is tabulated in descending order in Table 5. As can be seen in Table 5, more than 80% of the questions asked by the professors in all the three divisions were audience-oriented questions (HA: 90.0%; SS: 89.2%; PS: 80.6%). However, professors in PS asked relatively more content-oriented questions than those in soft fields (PS:19.4%; HA: 10.0%; SS: 10.8%).

Table 5

Interdisciplinary differences in the distribution of question functions

Humanities & Arts		Social Sciences & Education		Physical Sciences & Engineering	
Audience-oriented	675 90.0%	Audience-oriented	289 89.2%	Audience-oriented	468 80.6%
eliciting response	233 31.1%	checking comprehension	92 28.4%	eliciting response	199 34.3%
checking comprehension	184 24.5%	eliciting response	77 23.8%	checking comprehension	155 26.7%
class management/engagement	177 23.6%	class management/engagement	67 20.7%	class management/engagement	57 9.8%
soliciting agreement	43 5.7%	soliciting agreement	44 13.6%	soliciting agreement	42 7.2%
requesting confirmation/clarification	38 5.1%	requesting confirmation/clarification	9 2.8%	requesting confirmation/clarification	15 2.6%
Content-oriented	75 10.0%	Content-oriented	35 10.8%	Content-oriented	113 19.4%
stimulating thought	41 5.5%	focusing information	18 5.6%	focusing information	84 14.5%
focusing information	34 4.5%	stimulating thought	17 5.2%	stimulating thought	29 5.0%
Total	750 100%	Total	324 100%	Total	581 100%

In terms of audience-oriented questions, the major purposes of using this type of questions by professors from all the three different divisions were to elicit students' responses to content-related questions and to check students' comprehension of the lecture content. Given this result, we might speculate that *eliciting response* and *checking comprehension* seem to be the major functions of question used in academic lectures in general. However, it can be noted that in the soft fields in HA and SS, the percentages of questions devoted to class management/engagement were more than double of that in the hard field in PS. (HA: 23.6%; SS: 20.7%; PS: 9.8%). It seems that professors from the soft fields pay more emphasis on class engagement and rapport enhancement than those from hard fields.

Moreover, it can also be observed that SS seemed to solicit students' agreement (13.6%) more frequently than professors in HA (5.7%) and PS (7.2%). According to Thompson (1998), when using questions to seek agreement from the audience, the speaker is moving from the power associated with information giving to a more equal relationship, in which the speaker asks the audience to evaluate and confirm the speaker's judgment. However, she also points out that the act of *soliciting agreement* puts certain amount of pressure on the audience to agree with the speaker. It can thus be a potential face-threatening strategy if it is used to seek the audience' agreement with a controversial point. However, despite of its face-threatening nature, SS professors asked questions to seek agreement the most frequently across the three divisions. This might be related to the more complex and fluid nature of social science knowledge which often encompasses more various perspectives and possible interpretations than that in PS while at the same time allows a lower degree of subjectivity than the knowledge construction in HA.

If we consider content-oriented questions, it can be found that PS professors tend to use them more frequently than those in soft fields. While content-oriented questions made up

nearly 20% of all the questions in PS, they comprised only around 10% in HA and SS. Especially professors from PS are proportionately more likely than those in the soft fields to use content-oriented questions to focus information than to stimulate thoughts (see examples 11 and 12). This finding could be explained by the fact that the abstract nature of hard knowledge seems to impel PS professors to use these questions as an effective attention focalizing mechanism and as topic highlighter to introduce the new upcoming sub-topic in order to facilitate students' lecture comprehension.

However, Thompson (1998) argues that questions which function to highlight the new information are less dialogic than those used to stimulate thought because the speaker both asks and responds to the question himself or herself. This also makes this type of question more controlling, in that the speaker asks the question s/he knows that the audiences might want to ask and immediately answer it herself/himself. Similarly, Hodge and Kress (1993, p. 99) also indicated that questioning is related to issues of knowledge and power; they further argue that "the right to define the answer to a question comes from social power" (originally cited in Thompson, 1998, p. 144). In this view, the result that PS professors ask questions to focus information more often than those in the other two soft fields seems to confirm Thompson's (1998, p.144) argument that the speaker of academic talks in hard fields is projected as more omniscient being who answers the question s/he asks.

The mapping of question forms and functions

In the previous two sections, we have discussed the similarities and differences in the use of different question forms and functions among the three divisions. In order to further synthesize these results, based on the mapping of question forms and functions, the most frequent question patterns in each division were also found. As shown in Table 6, the two most frequently used question patterns across the three divisions are the same, although with different frequency orders in soft and hard fields: questions in the form of *declarative/Imperative + word tag* used for comprehension checking (example 13) and *Wh-questions* for response eliciting (example 14).

it's called the transition to the transition. okay? (*HA—Lecture 3*)

this X variable was, gender. okay? (*SS—Lecture 8*)

when i try to quantize the system, i'm able to quantize it... alright? (*PS—Lecture 11*)

Mr. Froid, what's a rescue opera? (*HA—Lecture 3*)

what effect does that have on the reader? (*SS—Lecture 6*)

what's the definition of enrichment? (*PS—Lecture 13*)

should I spell that for you? (*HA—Lecture 2*)

what is that carbon? uh that's_ remains a mystery. (*PS—Lecture 14*)

Table 6

Top five common question patterns across divisions

Humanities & Arts <i>N=750</i>		Social Sciences & Education <i>N=324</i>		Physical Sciences & Engineering <i>N=581</i>	
Declarative/Imperative+word tag for comprehension checking	19.3%	Declarative/Imperative+word tag for comprehension checking	26.2%	Wh-question for response eliciting	22.9%
Wh-question for response eliciting	17.9%	Wh-question for response eliciting	17.3%	Declarative/Imperative+word tag for comprehension checking	19.4%
Yes/no question for class management/engagement	16.7%	Yes/no question for class management/engagement	14.5%	Wh-question for focusing information	12.9%
Yes/no question for response eliciting	12.1%	Declarative/Imperative+word tag for agreement soliciting	9.3%	Yes/no question for response eliciting	10.5%
Wh-question for class management/engagement	4.1%	Yes/no question for response eliciting	5.9%	Yes/no question for comprehension checking	6.9%

In addition, echoing the previous discussion, as also shown in Table 6, while the third most common question pattern found in the soft fields in HA and SS is Yes/no questions for class management/engagement (example 15), it is Wh-question for information focusing (example 16) ranks the third in the hard fields in PS. This result further confirms the contrastive finding that while HA and SS professors' seem to invest greater efforts in asking questions to build closer rapport with their students and to enhance the classroom management, PS lecturers appear to more questions to highlight upcoming new information. Moreover, as we can also note in Table 6, while the form *declarative/imperative+ word tag* for agreement soliciting ranks the fourth in SS, none of the top five question patterns in HA and PS falls into this category. This again confirms the previous finding that SS professors seem to solicit agreement from their students more frequently than professors in HA and PS.

Finally, based on the data shown in Table 6, we can observe that the three divisions share certain similarities in the realization of the two most common question functions (i.e. eliciting response and checking comprehension): In all the three divisions, the most common question form used to elicit students' responses to content-related questions is *Wh-question* and the second common one is *yes/no question*. Similarly, in all the three divisions, the most common form used to check students' lecture comprehension is *declarative/imperative + word tag*. Due to the space limit, we will provide a more detailed discussion on the issue how different question functions are realized linguistically in these three divisions in another paper.

CONCLUSION

The results of this study show both similarities and differences between soft and hard fields in the use of questions in academic lectures. In all the three fields, over 80% of the questions used are audience-oriented. This result implies the professors' great efforts in creating an atmosphere of interactive sharing of ideas and information with their students. Since the major functions of academic lectures include (1) knowledge transmission, (2) guiding and encouraging students through interaction (Cammiciottoli, 2004, 2008; Morell, 2007), and (3) enculturating students into their disciplinary communities. It is thus not surprising to see that in the three divisions studied, questions are used the most frequently to elicit students' responses to content-related questions and to check students' understanding of lecture content.

In addition, in order to enculturate students into their disciplinary communities, lectures often reflect different disciplinary cultures and different natures of disciplinary knowledge. In hard fields, the process of knowledge production is accumulative in nature; more shared background knowledge and standard procedures of knowledge making can be established, hence creating a less interactive and more technical style of discourse. In addition, the knowledge of hard fields is often more abstract than that of soft fields. In order to facilitate students' learning process, PS professors thus tend to ask content-oriented questions (which is more controlling and less dialogic in nature) more often than their soft-field colleagues as attention focalizing device to highlight the upcoming information.

On the other hand, the research targets in soft fields are very often much more fluid and complex in nature than those in hard fields. Disciplinary knowledge in these fields is often established through more sophisticated argumentation and knowledge claim negotiation based on more various complex theoretical perspectives – often without agreed shared background knowledge and standard methodological procedures among community members. The establishment of new knowledge in soft fields is thus much more persuasive and dialogic in nature; it does not show the same liner developmental patterns as in hard fields. This plus the less hierarchical power structure among community members might explain why

professors in the two soft divisions tend to use questions to engage their students and to manage teaching flow more often than those in the hard fields in PS. Further, among the three divisions, the social construction of new knowledge (i.e., scholars competing for the acceptance for their knowledge claims through negotiation with other community members) is especially apparently exhibited in SS professors' use of question to solicit their students' agreement.

To conclude, this study not only reveals the interdisciplinary similarities and differences in the use of question in academic lectures. The major question functions along with their major associated linguistic forms found in this study could also be used by EAP teachers as supplementary information to guide their students to improve their English lecture comprehension.

REFERENCES

- Aschner, M. J. (1961). Asking questions to trigger thinking. *NEA Journal*, 50, 44-46.
- Athanasidou, A. (1991). The discourse function of questions. *Pragmatics*, 1(1), 107-122.
- Bamford, J. (2005). Interactivity in academic lectures: The role of questions and answers. In J. Bamford, & Bondi, M. (Ed.), *Dialogue within Discourse Communities: Metadiscursive Perspectives on Academic Genres* (pp. 123-145). Tübingen: Max Niemeyer.
- Crawford Camiciottoli, B. (2003). Interacting with the audience: modality in business lectures delivered in a cross-cultural context. In R. Facchinetti & F. Palmer (Eds.), *English modality in perspective: Genre analysis and contrastive studies* (pp. 27-43). Frankfurt: Peter Lang.
- Crawford Camiciottoli, B. (2004). Interactive discourse structuring in L2 guest lectures: some insights from a comparative corpus-based study. *Journal of English for Academic Purposes*, 3, 39-54.
- Crawford Camiciottoli, B. (2008). Interaction in academic lectures vs. written text materials: The case of questions. *Journal of Pragmatics*, 40(7), 1216-1231.
- Csomay, E. (2002). Variation in academic lectures: Interactivity and level of instruction. In R. S. Reppen, M. Fitmaurice, & D. Biber (Eds.), *Using corpora to explore linguistic variation* (pp.205-224). Amsterdam: John Benjamins.
- DeCarrico, J.S., & Nattinger, J. R. (1988). Lexical phrases for the comprehension of academic lectures. *English for Specific Purposes*, 7(2), 91-102.
- Dudley-Evans, T., & Johns, M. J. (1981). A team teaching approach to lecture comprehension for overseas students. In *The Teaching of Listening Comprehension (ELT Documents Special)*(pp. 30-46). London: British Council
- Flowerdew, J. (Ed.). (1994). *Academic Listening: Research Perspectives*. Cambridge: Cambridge University Press.
- Fortanet, I. (2004a). Enhancing the speaker- audience relationship in academic lectures. In P. Garces, R. Gómez, L. Fernandez,& M. Padilla (Ed.), *Current Trends in Intercultural, Cognitive and Social Pragmatics* (pp. 83-96). Kronos, S.A. : Sevilla.
- Fortanet, I. (2004b). The use of 'we' in university lectures: reference and function. *English for Specific Purposes*, 23, 45-66.
- Gall, M. (1970). The use of questions in teaching. *Review of Educational Research*, 40(5), 707.
- Goody, E. N. (1978). Towards a theory of questions. In E. N. Goody (Ed.), *Questions and Politeness: Strategies in social interaction* (pp. 17-43). Cambridge: Cambridge University Press.
- Goffman, E. (1981). The lecture. In E. Goffman (Ed.), *Forms of talk* (pp.162-195). Philadelphia, PA: University of Philadelphia Press.
- Hansen, C., & Jensen, C. (1994). Evaluating lecture comprehension. In J. Flowerdew (Ed.), *Academic listening* (pp. 241-268). Cambridge: Cambridge University Press.

- Hyland, K. (2000). *Disciplinary Discourse: Social Interactions in Academic Writing*. Harlow: Longman.
- Launspach, S. L. (1998). Interactional strategies and the role of questions in the acquisition of academic discourse. *Dissertation Abstracts International, A: The Humanities and Social Sciences*, 59(5), Nov, 1547-A-1548-A.
- Liu, W. L., & Chang, Y.Y. (2009, April). Professors' use of questions in small lectures: Are there gender differences? Presented at the 2009 International Conference on Applied Linguistics & Language Teaching. Taipei: National Taiwan University of Science and Technology.
- Long, M. H. & Sato, C. J. (1983). Classroom foreigner talk discourse: forms and functions of teachers' questions. In H. W. Seliger & M.H. Long (Eds.), *Classroom Oriented Research in Second Language Acquisition*. Rowley, MA: Newbury House, 146-165.
- Morell, T. (2001). The role of discourse markers and personal pronouns in lecture discourse. In A. Moreno & Y. V. Cowell (Eds.), *Recent perspectives on discourse*. Universidad de León. AESLA: León.
- Morell, T. (2004). Interactive lecture discourse for university EFL students. *English for Specific Purposes*, 23(3), 325-338.
- Northcott, J. (2001). Towards an ethnography of the MBA classroom: A consideration of the role of interactive lecturing styles within the context of one MBA programme. *English for Specific Purposes*, 20, 15-37.
- Rounds, P. (1987a). Multifunctional personal pronoun use in educational setting. *English for Specific Purposes*, 6(1), 13-29.
- Rounds, P. (1987b). Characterizing successful classroom discourse for NNS teaching assistant training. *TESOL Quarterly*, 21(4), 643-671.
- Swales, J. M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge, U.K.: Cambridge University Press.
- Thompson, S. (1998). Why Ask Questions in Monologue? Language choice at work in scientific and linguistic talk. In S. Hunston (Ed.), *Language at work*. (pp. 137-150). Clevedon: University of Birmingham.

Rime Familiarity Effect on Spoken Vocabulary Acquisition

Aleck Shih-wei Chen (陳世威), Patti Yi-ping Shih (施宜評)

National Cheng Kung University
University of Wisconsin, Madison

aleck@mail.ncku.edu.tw

Unlike its Chinese counterpart, an English syllable can be very complex and an English word can be multisyllabic. This implies that each word could be represented by a unique combination of speech sounds or, more specifically, there could be as many sound combinations as there are words, which can be in tens of thousands. Familiarity with the invariant units in similar-sounding words such as rimes thus appears critical to efficient lexical processing, as it presumably could greatly reduce the learner's cognitive load. This should be true especially with speakers learning English as a foreign language (EFL), as the poverty in exposure to English could prove unfavorable to the fostering of such familiarity. The present study examines the potential contribution of rime familiarity to the success in EFL spoken word learning. It is hypothesized that, given that spoken word learning is basically a practice of associating a concept with a phonological form, familiarity with English rimes should facilitate word learning. Seventy-six college freshmen were tested for, among others, their (1) familiarity with nonwords rhyming with those given in the word learning task (target rime familiarity) and (2) that with non-rhyming words (non-target rime familiarity). As the results showed, familiarity with the target rimes explains a statistically significant proportion of variance in spoken word learning, even after short-term memory, spoken vocabulary, and phoneme awareness have been accounted for. In contrast, familiarity with non-target rimes does not predict success in spoken word learning. Pedagogical implications will be discussed.

INTRODUCTION

Word learning is basically a continual effort in associating concepts with linguistics forms, written and/or spoken (Hu, 2003). For native speakers, this process usually involves familiar spoken forms. Native speakers of English, as an example, already possess "thousands of words already present in their spoken lexicon" when they begin learning to read as a child (Ziegler & Goswami, 2005, p.3). This familiarity enables them to focus their cognitive effort on the task of association alone, thus greatly reducing their cognitive load in word learning. For foreign language learners, however, word learning involves not only association but an additional effort to familiarize themselves with the target language's phonological forms. This is true especially with learners of English as a foreign language (EFL).

English has a complex syllable structure. An English syllable, for example, can have as many as three prevocalic consonants (e.g., *script*) and up to four postvocalic consonants (e.g., *texts*). An English word, moreover, can be multisyllabic. Taken together, these imply that, in theory, each word in English can be represented by a unique combination of speech sounds. One potential problem with such a huge number of linguistics forms is the heavy cognitive load demanded of word learning. That is, if each spoken word is learned as a holistic unit on its own, the number of phonological representations to be stored would be as numerous as the number of words acquired, which can be in tens of thousands. Fortunately,

the number can be significantly reduced if similar patterns can be found among syllables. That is, with subdivision into recurring chunks such as rimes, the number is greatly reduced to a cognitively affordable one. For example, for the most frequent 3,000 monosyllabic English words, there are only 400 rimes (Ziegler & Goswami, 2005, p.19). This suggests that familiarity with subsyllabic units such as rimes may lighten a learner's cognitive load, hence improving spoken words learning.

METHOD

If familiarity with subsyllabic units indeed facilitates spoken word learning, one should expect a significant correlation between EFL word learning and familiarity with English rimes, even after potentially confounds such as short-term memory (e.g., Avons, Wragg, Cupples, & Lovegrove, 1998; Cheung, 1996; Gathercole, Hitch, Service, & Martin, 1997; Michas & Henry, 1994; Service & Craik, 1993; Service & Kohonen, 1995), existent vocabulary and phoneme awareness (e.g., Hu, 2003) have been accounted for. To examine this hypothesis, seventy-six college freshmen from a national university were recruited to test for their performance in spoken nonword learning, the criterion variable. They were also tested for their familiarity with (1) nonwords rhyming with (for target rime familiarity) and (2) those not rhyming with (for non-target rime familiarity) the nonwords tested in the learning task, as well as their performance on the three potential confounds mentioned above. It is hypothesized that, if rime familiarity indeed has an effect, it should be able to predict the participants' success in learning spoken nonword with all the said confounds statistically controlled. In contrast, familiarity with non-target rimes should not.

Measures

Among the five variables, receptive vocabulary was measured using the standardized receptive vocabulary test Peacock Picture Vocabulary Test, Revised Version (PPVT-R). The phoneme awareness (PA) tasks consist of two tests—phoneme deletion, where the participant was asked to repeat a nonword without saying a designated phoneme in it; and phoneme isolation, where the participant was asked to say a designated sound of a heard nonword. The two PA test were adapted from Chen (2006). Short-term memory was measured with a memory span task the authors had created. The participant listened to pairs of strings of nonwords that were identical in every aspect except for their order, and were then asked to judge whether they are identical or not. The pairs were varied in length from 3 to 8 nonwords. There were two main reasons for employing this matching task instead of the more traditionally used serial recall task. For one, testing participants speaking the target language as a foreign language with serial recall risk the confounding of the FL proficiency factor. For another, series recall has been argued to be a less accurate measure than other measures such as (immediate) nonword repetition (e.g., Gathercole & Baddeley, 1993).

Even nonword repetition as a purer measure is not without a problem, however.

In fact, it is difficult to rule out long-term memory contribution to immediate memory performance even with nonwords. In a detailed item analysis of the nonword repetition test, we found that children's repetition accuracy was highly associated with the rated wordlikeness of the nonword (Gathercole, Willis, Emslie, & Baddeley, 1991). For this reason, we suggest that where close *phonological analogies to nonwords* are available in long-term memory, subjects may use them to support working memory representations of nonwords (Gathercole & Baddeley, 1993; p.48; the italics are ours). For the very same reason, nonword repetition is used in this study to measure rime familiarity, which is assumed to be stored in the long-term memory. As have been shown in the literature, nonword repetition reflects "an additional availability of phonological specifications in long-term memory" (ditto; p.48). As such phonological specifications are based on wordlikeness and the wordlikeness in the items created for the present study is limited to the

rime portion, repetition of nonwords sharing the same rimes should reflect, at least in part, the rime familiarity we are interested in. Moreover, as the items are limited to monosyllabic nonwords only, whose repetition is well within most adults', even children's STM capacity, they are unlikely to reflect individual differences in short-term memory, as in Gathercole and Baddeley's studies (ditto), where the nonwords vary in both the number of syllables and syllable complexity.

Rhyme familiarity in the present study was thus measured using nonword repetition, in which the participants were asked to repeat a monosyllabic nonword as fast and as correctly as they can. Rhyme familiarity items were divided into two categories, including (1) the target items, i.e., items sharing rimes with those tested in the word learning task and (2) non-target items, i.e., those not sharing rimes. Finally, as the criterion measure, the nonword learning task requires the participant to learn to associate three pictures with their names. The participants were first taught the three nonwords by playing their pronunciation, one at a time, accompanied by their corresponding pictures (i.e., the concept). They were then asked to name each picture without the recordings. The recording was played again only if they failed to correctly name the picture. A trial consists of naming all three pictures, with or without success. The same process went on until all three pictures were correctly named for two consecutive trials.

RESULTS AND DISCUSSION

Two sets of stepwise multiple regression analysis were performed with the collected data, with number of trials in the word learning task as the criterion variable. In both sets, scores on memory span, receptive vocabulary, and phoneme awareness were entered at respectively step one, two, and three to control for their potential impact. The two sets differ in the order of entry of the two rime familiarity categories. In one, familiarity with the target rimes is entered at step 4 and familiarity with non-target rimes, at step 5. The order of the two was reversed in the other set. Among the control variables, as the results show (see Table 1), only phoneme awareness predicts success in word learning, but not memory span. This is consistent with findings in earlier literature (e.g., Gathercole & Baddeley, 1993; Hu, 2003). Essentially, while short-term memory is helpful with word learning for young children, it is long-term memory that plays a more significant role with older children, adults alike. It is a bit surprising, though, that existent receptive vocabulary is marginally non-significant. This probably reflects the limited exposure to the target language in an EFL context.

Indeed, as shown in the differential predictive power of the two rime familiarity categories, the effect of rime familiarity is limited to words sharing similar rimes (R^2 change = .052; $F(1, 69) = 4.99$, $p < .05$ when entered at step 5; and R^2 change = .059; $F(1, 70) = 5.72$, $p < .05$ when entered at step 4). Familiarity with rimes not used in the word learning task does not contribute to success in nonword learning even when entered at step 4. This suggests that the effect of rime familiarity is word-specific. More specifically, new word learning is facilitated only when the learner is familiar with its rime. Such familiarity is possibly derived from exposure to words sharing the same rimes. Pedagogically, this implies that an increase in the learner's rime inventory would be helpful in EFL word learning. To increase the rime inventory, furthermore, sufficient exposure to rhyming words will be necessary. In EFL vocabulary teaching, it is therefore desirable to integrate rhyming training to a word learning syllabus.

Table 1 *Predictors of English nonword learning*

Order of Entry	Variables	R^2	Adjusted R^2	R^2 change	F change	p -value
Step 1	Memory Span	.025	.011	.025	1.864	.176
Step 2	Receptive Vocabulary	.070	.044	.045	3.548	.064
Step 3	Phoneme Awareness	.225	.181	.155	7.090	.002*
Step 4	Target rime familiarity	.283	.232	.059	5.720	.019*
Step 5	Non-target rime familiarity	.284	.221	.001	0.055	.816
Step 4	Non-target rime familiarity	.232	.177	.007	.668	.416
Step 5	Target rime familiarity	.284	.221	.052	4.991	.029*

* Significant at $p < .05$.

REFERENCES

- Avons, S. E., Wragg, C. A, Cupples, L., & Lovegrove, W. J. (1998). Measures of phonological short-term memory and their relationship to vocabulary development. *Applied Psycholinguistics*, 19, 5843-5601.
- Cheung, H. (1996). Nonword span as a unique predictor of second-language vocabulary learning. *Developmental Psychology*, 32, 867-873.
- Gathercole, S. E. & Baddeley, A. D. (1993). *Working memory and language*. Hove/New York: Psychology Press.
- Gathercole, S. E., Hitch, G., Service, E., & Martin, A. J. (1997). Phonological short-term memory and new word learning in children. *Developmental Psychology*, 33,966-979.
- Hu, C. F. (2003). Phonological memory, phonological awareness, and foreign language word learning. *Language Learning*, 53, 429-462.
- Service, E., & Craik, F. I. M. (1993). Differences between young and older adults in learning a foreign vocabulary. *Journal of Memory and Language*, 32,608-623.
- Service, E., & Kohonen, V. (1995). Is the relation between phonological memory and foreign language learning accounted for by vocabulary acquisition? *Applied psycholinguistics*, 16, 155-172.

Literature Discussions in an EFL Reading Class

Hsiu-chuan Chen (陳琇娟)

Kang Ning Junior College of Medical Care and Management

hcchen@knjc.edu.tw

The use of student-led literature discussions helps students become more actively engaged in meaning making through exchange of interpretations of texts rather than the meaning controlled by teachers. The study describes the meaning making process of the EFL learners who participated in reading, writing, and talking about books in a literature-based reading class. Participants included two focal groups with eight second-year students from the five-year program at a junior college. Field notes, semi-structured interviews, student's reading logs, and transcribed audiotapes of 16 discussions were collected; 8 representative discussions were selected for intensive analysis to understand how the EFL learners constructed meanings in their literature discussion groups. Analysis of student discourse revealed different categories of utterances including literal comprehension, personal involvement, interpretation, and evaluation, etc. The findings indicated that students journeyed from the literal, interpretive, to applied level of reading comprehension and that they found their group reading experiences meaningful and enjoyable. Moreover, they developed and practiced their cognitive skills through four complementary stances toward reading: (a) an efferent stance, (b) an aesthetic stance, (c) a critical stance, and (d) a reflective stance. In conclusion, pedagogical implications are addressed.

INTRODUCTION

Most of English teachers in Taiwan spend almost the entire class time breaking the reading process into smaller, more manageable units (Lin, 2002; Huang, 2004). When students read in English, they think they need to know all the words in a text to understand it. Reading word-for-word is in itself an inefficient and wearisome experience. Students tend to get the notion that they read primarily for the purpose of answering teacher's comprehension questions. As they fail to find pleasure in reading, reading becomes meaningless (Hsu, 2002). The use of student-led literature discussions helps students become more actively engaged in meaning making through exchange of interpretations of texts rather than the meaning controlled by teachers. The study describes the meaning making process of the EFL learners who participated in peer-led group discussions of literature in an EFL reading class.

LITERATURE REVIEW

Rosenblatt's (1994) transactional theory and Vygotsky's (1978) sociocultural theory provide the underlying theories for literature discussion groups. The transactional theory views the reading process as a transaction in which the reader makes a deal with the writer by means of the prints. How much one gets from a reading transaction is influenced by the reader's life experiences. Therefore, different readers may have different interpretations of the same text. Rosenblatt also proposes peer reading to foster growth in students' reading ability. Such perspectives are in line with Vygotsky's (1978) sociocultural theory which views cognitive development as a primarily social act.

Vygotsky (1978) maintains that learning is closely tied to social interactions within the *zone of proximal development* (ZPD). It is the distance between the actual development level in which a child can perform a task on his own and the level of potential development in which a child can perform a task with more knowledgeable others. Student collaboration in

discussion groups is viewed as ZPD where they can construct meaning through both a social process and an individualization process. In other words, groups provide opportunities for students to interact with others. Therefore, students are more apt to organize and clarify the text in new ways so that their interpretations can be easily understood (Alley, 2005; Ghaith & El-Malak, 2004).

Due to its structural complexity, literature is almost impossible for all except advanced ESL/EFL learners. Simplification matches the language competence of the ESL/EFL learners with a limited vocabulary. Simplified texts such as graded readers have attracted criticism of inauthenticity. Recently, the corpus-based studies comparing the original and simplified texts have indicated that language such as word frequency, lexical chunks, and structure patterns are similar in both versions (Allan, 2009; Claridge, 2005). Therefore, short stories from graded readers were used as literary texts for EFL learners with limited proficiency to conduct peer-led group discussions in the present study.

METHODOLOGY

This study took place in a second-year EFL reading classroom of 44 female and 4 male students from the five-year program at a junior college. There were 12 heterogeneous teams in the class and each team consisted of one high-achiever, two middle-achievers, and one low-achieving student. The sampling criterion of focus groups was characteristics of discussion proficiency identified by Almasi et al. (2001). The group with the highest frequency of presence of the characteristics and the group with the lowest frequency of presence of the characteristics were selected to fully display diverse perspectives (Creswell, 1998).

The EFL learners read the eight short stories selected from three graded readers in the later half of spring semester of 2008, from the end of April to the end of June. The level of these graded readers is elementary because all of the other textbooks for the English courses such as English conversation they took were elementary level. The instructional procedures were as follows: (1) Read-aloud of short story by me, (2) Peer-led discussions of literature, (3) Whole-class debriefing sessions of discussion, and (4) Writing in reading logs. Field notes, semi-structured interviews, student's reading logs, and transcribed audiotapes of 16 discussions were collected; 8 representative discussions were selected for intensive analysis, using a categorization system modified and expanded from Eeds and Wells's (1989) to understand how the EFL learners constructed meanings in their literature discussion groups. Other data sources were put into content analysis to probe how participants responded to group reading.

RESULTS AND DISCUSSION

Analysis of student discourse revealed eight different categories including literal comprehension, interpretation, evaluation, personal involvement, conversational maintenance, discussion of illustration, log-related talk, and off-task chatting. In this paper, I discuss the four categories generated from discussions of verbal texts: literal comprehension, interpretation, evaluation, and personal involvement

Literal comprehension included talk dealing with word meaning and recounting and questioning fact regarding setting, characters, and plot. Interpretation reflected students' abilities to make hypothesis and predictions. On the whole, students supported their inferences and predictions with textual information and background knowledge. Evaluation included statements in which the essence appeared to make judgments about literary quality such as story development.

Personal involvement included talk in which students connected the story to their own lives, gave themselves agency in stories, and shared their reactions to the literary work.

Discussions evolving from reactions to the story also inspired students to reflect on real-life issues. An example comes from the group with low discussion proficiency discussing *The Speckled Band* which tells a story of a stepfather who murdered his stepdaughters with a snake. Every night, he put the snake through the air-vent which went down the bell-rope to his stepdaughter's bed. The story ends with the stepfather getting killed with the same technique he used against his stepdaughter. The discussion began with Amy's⁴ opinion question which shifted the textual talk from factual level to applied level:

- Amy: *So, he had made the plan to murder his stepdaughter for years.*⁵
Gina: *Indeed, but he ended up getting killed.*
Amy: *Well...Do you think money can change people?*
Gina: *Yes.*
Judy: *Yes.*
Debbie: *Of course.*
Amy: *Wow!*
Judy: *Money is evil.*
Debbie: *There are many people like that, right? Because of...the money he got, a kind-hearted person ended up becoming a bad guy.*
Judy: *Something like fight over property among family members.*
Amy: *But do people murder their genetic children?*
Gina: *They do, I guess.*
...⁶
Amy: *I've got the final question. How to be a contented person?*
Debbie: *Yeah, how to be a contented person?*
Amy: *It's a little bit difficult to be a contented person.*
Debbie: *Yeah, it is.*
Judy: *To be a contented [⁷ person.*
Amy: *[Human desires are endless.*
Debbie: *Yeah.*
Judy: *Because human desires are endless, we need to be contented with what we have.*
Gina: *I believe that isn't applicable to a miser.*
Amy: *They probably need to experience frustration and difficulties.*
Gina: *Such experiences could cause one to understand the misery of others.*

In this vignette, after retelling the story event of how the stepfather died, students explored their own worldviews about money and then reflected on being contented with what they have. This excerpt also indicated that in group discussion, students journeyed from the literal, interpretive, to applied level of reading comprehension. They found their group reading experiences meaningful and enjoyable as one student remarked: *What you have never thought about the text is probably what other people have to say about it. And you will gain more answers, perspectives or something. Group reading is fun* (1st interview, Amy, 2008/05/07). Moreover, students developed and practiced their cognitive skills through four complementary stances toward reading: (a) an efferent stance, (b) an aesthetic stance, (c) a critical stance, and (d) a reflective stance. In developing an efferent stance, they recounted story content. In developing an aesthetic stance, they related the story to their personal lives. The personal engagement is what Rosenblatt (1978) calls "aesthetic reading" in which

⁴ All participants were assigned a pseudonym.

⁵ Words in italics are my English translations of participants' words in Chinese.

⁶ Segment of transcript deleted

⁷ Overlapping talk

students had a lived-through experience of the story. They practiced a critical stance when questioning texts and making evaluative statements about the literary quality. Students practiced a reflective stance when discussing substantial, real-life issues such as money and sharing their own value system.

CONCLUSION

In this study, the literature discussions have helped the EFL learners become more actively engaged in meaning making through exchange of interpretations of texts. Students were able to move beyond reading as decoding and to relate text to personal experiences. This small investigation has implications for instruction. First, we EFL teachers need to guide students to move beyond superficial story retellings. We can ask students to prepare several questions for group discussions and record them in their reading logs. In subsequent classes, we share a few excellent student-generated questions to offer a springboard for literature discussions as well as a model of how engaged readers “dialogue” with text. Second, for students who don’t get any formal training in English writing, reading logs may be a demanding cognitive challenge to them. We need to provide them with a guide to reading log to assist them in organizing ideas and expressing their thoughts more clearly. Such teacher support helps students move from literal understanding to aesthetic appreciation of literature.

REFERENCES

- Allan, R. (2009). Can a graded reader corpus provide ‘authentic’ input? *ELT Journal*, 63(1), 23-32.
- Alley, D. C. (2005). A study of Spanish II high school students’ discourse during group work. *Foreign Language Annals*, 38 (2), 250-258.
- Almasi, J. F., O’Flahavan, J. F., & Arya, R. (2001). A comparative analysis of student and teacher development in more and less proficient discussions of literature. *Reading Research Quarterly*, 36 (2), 96-120.
- Claridge, G. (2005). Simplification in graded readers: Measuring the authenticity of graded texts. *Reading in a Foreign Language*, 17 (2), 144-158.
- Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks: Sage.
- Eeds, M., & Wells, D. (1989). Grand conversations: An exploration of meaning construction in literature study groups. *Research in the Teaching of English*, 23 (1), 4-29.
- Lin, C. L. (2002). *Constructivism and second language learning: A web-based reading-writing activity*. Unpublished master’s thesis, National Taiwan Normal University.
- Hsu, M. L. (2002). *The effects of pair think-aloud procedures on Taiwanese senior high school students’ EFL reading comprehension and metacognitive awareness*.
- Huang, C. Y. (2004). *Think to win: An inquiry-based approach via collaborative strategic reading technique to teach English reading in a senior high EFL classroom*. Unpublished master’s thesis, National Taiwan Normal University.
- Ghaith, G. M. & El-Malak, M.A. (2004). Effects of Jigsaw II method on EFL reading comprehension. *Educational Research and Evaluation*, 10 (2), 105-115.
- Rosenblatt, L. M. (1978). *The reader, the text, the poem: The transactional theory of literary work*. Carbondale, IL: Southern Illinois Press.
- Rosenblatt, L.M. (1994). The transactional theory of reading and writing. In R. Ruddell, M. Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading* (4th ed.) (pp. 1057-1092). Newark, DE: International Reading Association.
- Vygotsky, L. (1978). *Mind in society: The development of higher psychological process*. (M. Cole et al. Eds. and Trans.). Cambridge, MA: Harvard Univ. Press.

Examining the Scoring Accuracy and Feedback Quality of an Automated Essay Evaluation Program

Howard Hao-jan Chen (陳浩然)

National Taiwan Normal University

hjchen@ntnu.edu.tw

Feedback is an important research issue in second language writing. Writing teachers expect that students will use the feedback provided to them to improve their writing. As computers become more affordable and powerful, new automated writing evaluation (AWE) systems have been developed to provide feedback to students. Several commercial products (e.g., My Access) are now widely available to ESL learners. Some empirical studies have been conducted to investigate the effectiveness of these AWE systems; however, the findings of these evaluation studies have been contradictory. To better understand the effectiveness of these programs, this study specifically examined two core services in My Access- the accuracy of scoring and feedback on errors. The scores and feedback provided by My Access were compared with the scores and feedback provided by human raters. The results showed that My Access tended to assign higher scores to students' essays, though there was a significant correlation between machine scores and human scores. Moreover, the feedback on student's grammar errors is neither comprehensive nor accurate. Based on these findings, it is suggested that teachers and students should be careful in interpreting the scores and using the feedback provided by My Access. Currently, the AWE systems can only provide limited help for ESL writers.

INTRODUCTION

Many second language learners find it very challenging to become proficient writers of the target language. These learners must devote much time and effort to practicing writing before they can produce texts that are fluent and accurate. In addition to sufficient writing practice, many second language teachers and researchers believe that learners need to receive feedback on their writing. If learners only keep on writing and do not receive adequate feedback, they will not be able to make progress quickly. Other researchers have questioned the assumptions (Ferris, 1999; Truscott, 1996, 1999). Even though the role of corrective feedback in second language learning remains controversial, many teachers and learners firmly believe that feedback plays an important role in second language writing.

Although it is widely accepted that second language writers need feedback (on both form and content), it is often difficult for writing teachers to provide detailed feedback. In many ESL/EFL settings, teachers have 40 or 50 students in their classes. Reading and commenting on students' essays thus becomes a great burden for many teachers.

The rapid development of computer technology, particularly natural language processing (NLP) technologies, has provided a brand new type of computer application for second language writing: automated writing evaluation programs. Several companies have begun to sell these programs. The companies selling these products claim that these automated writing evaluation programs can provide immediate feedback on learners' grammatical errors and an evaluation of the content and organizational aspects of essay writing. In addition, these systems can provide convenient writing tools (writing guides, spelling checker, and thesaurus) to facilitate students' writing processes. With the help of these new automated writing evaluation programs, writing teachers can reduce time spent correcting and

commenting on students' compositions. Given that automated writing evaluation programs can help teachers provide assessment and feedback, these programs have been increasingly used as an alternative tool for assessing students' writing skills.

Two of the most well-known programs of this type are Criterion (developed by the Educational Testing Services) and My Access (developed by Vantage Learning). By subscribing to these programs, students can choose from a wide range of practice essays topics to write multiple drafts and receive immediate feedback in the form of both holistic scores and diagnostic comments on grammar, organization, style and usage.

These automated writing evaluation systems try to emulate the human rater's behavior in grading students' essays. NLP researchers and statistical experts first collected many essays scored by human raters and tried to identify the textual features for each different proficiency level. The features of an essay include the length of the article, lexical complexity, syntactic variety, topical content, and grammatical errors. These features were identified and classified with the help of NLP technologies. Automated writing evaluation systems in fact do not understand the meaning of an essay, but they can be trained to identify the key features in an essay rapidly. After many trials and adjustments, the automated writing evaluation systems were gradually able to make a human-like judgment and assign proper scores based on the calculation of weights of many different textual features.

The agreement rate on holistic scores between the Criterion e-rater and human expert graders is typically 97% (Burstein, Chodorow, & Leacock, 2003). According to Elliot (2001), My Access agrees with the expert grader scores within one point 99% of the time, and exactly with the expert grader scores 73% of the time.

Conflicting Findings about the Validity of Automated Writing Evaluation Systems

Given the power and convenience of these automated writing evaluation systems, several studies on English native speakers have been conducted to assess the effectiveness of Criterion and My Access. Higgins et al (2004) examined Criterion for its responsiveness to coherence in 989 twelfth-grade students' essays. They found that Criterion is capable of identifying four aspects of coherence, including a) relationship to the topic of the essay, b) relationship of other discourse elements, c) relevance with discourse elements, and d) errors in grammar, mechanics, and usages. Moreover, Attali (2004) evaluated Criterion for its feedback and revision features. Using large-scale field data from approximately 9,000 drafts and revised compositions of sixth to twelfth grade students, he found that the students understood the feedback given by the program and utilized the feedback to improve their revised versions.

Other researchers have been interested in examining the efficacy of My Access. Based on four separate studies, Elliot & Mikulas (2004) reported that the writing skills of fifth to eleventh grade students, as measured by student performance on statewide writing assessments, were significantly improved by using My Access. In a survey administered to 94 eighth-grade students using this program, over 80% described the automated feedback on their essays as both helpful and accurate.

Although the aforementioned studies above have found some positive results, other studies reported some negative outcomes. Herrington (2001) found that the scoring engine of My Access can be fooled by the writer, partly because it awards scores for essay length, even when the content is less than adequate. One way of fooling the system is to copy and paste the same paragraph many times. The writer can thus devise means of beating the program, rather than making a genuine effort to improve his or her writing. He also noted that the AWE system failed to recognize nuances such as sarcasm, idioms and clichés used in students' essays.

Grimes and Warschauer (2006) used interviews, surveys, and classroom observations to study teachers and students using My Access and Criterion. Although teachers and students

valued the automated feedback as an aid for revision, teachers scheduled little time for revising, and students made little use of the feedback except to correct spelling errors.

In addition to the studies focusing on native writers, a number of empirical studies have been conducted to investigate the effectiveness of My Access in EFL contexts. Yu & Yeh (2003) examined the use of the computerized feedback of My Access in enhancing Taiwanese college students' writing. They found that most students considered the feedback from My Access useful for writing their first revisions, but they soon felt the program was not helpful after receiving the same feedback several times. Some students also pointed out that the computerized feedback was not specific enough to help them revise their writing.

Yang (2004) investigated the effectiveness of using My Access in three different contexts. Her study involved approximately 300 subjects from freshman English classes, English composition classes, and one class from a self study program. She found that, although a majority of the students felt positive toward the automated essay grading tool, her students also pointed out that the fixed, repeated feedback became useless to them. This finding is similar to the findings in Yu and Yeh (2003). Moreover, only 13% of the students thought that the scores they received from My Access were adequate, whereas more than half felt uncertain about the scores.

Chen and Cheng (2006) also investigated 68 college students' responses to My Access. Their findings showed that only 55% of the students felt this program either moderately or slightly helpful to them in improving their writing skills. Students were much less satisfied with the scores assigned by the electronic rater. Some of them felt that the My Access program gave them improper scores. Their dissatisfaction was mainly because this program failed to give specific feedback in the content and rhetorical aspects of their writings. This study also found that how instructors used the writing tool might influence students' attitudes toward the program.

Wang and Brown (2007) specifically examined the validity of automated essay scoring (AES) by comparing group mean scores assigned by an AES tool, IntelliMetric™, and by human raters. Data collection included two steps. First, they administered the Texas version of the WritePlacer Plus test. Second, they then obtained scores assigned by both IntelliMetric™ and by human raters. A One-Way Repeated-Measures ANOVA was conducted to examine the differences between the AES mean score and human raters' mean score. Statistical results indicated that the mean score assigned by the IntelliMetric™ was significantly higher than the faculty human raters' mean score on the WriterPlacer Plus test. This finding did not support previous studies that reported non-significant mean score differences between AES and human scoring.

The findings of these empirical studies are very similar. Students either were not satisfied with machine scores or did not trust the scores assigned by the electronic raters. Students also pointed out that the feedback from My Access was vague and repetitive so that they cannot make good use of it. In addition to the automatic scoring, the feedback provided by the AWE is not without problems. Several negative comments were made concerning the macro-level feedback were noted in the aforementioned studies.

To better understand why many EFL college students do not trust the automated scoring mechanisms and had complaints or difficulties in utilizing computerized feedback, the purpose of this study was to further investigate the accuracy of scoring and the quality of feedback of AWE systems. Previous studies did not carefully examine the correlation between human scoring and AWE scoring and they also did not compare the grammar feedback provided by AWE and human instructors. Given that the AWE systems are widely used by many ESL and EFL students around the world, it is essential to carefully investigate the scoring accuracy and feedback quality of these systems. For the automated

scoring, it is necessary need to examine the correlations between the human raters and the AWE and ascertain the differences between the AWE programs and human raters. To investigate the quality of the feedback, it is important to examine the macro-level feedback (content and structure) and in particular the micro-level grammar feedback provided by AWE systems and to determine the strengths and limitations associated with these systems.

METHODOLOGY

To answer the research questions proposed above, the AWE product My Access, created by Vantage Company, was investigated. To obtain the scores and feedback from My Access, students' essays should be submitted to My Access first. To ensure the essays were written by students at different proficiency levels, several groups of students (college freshmen, English majors, and master/doctoral students in a TESOL program) were invited to submit one or two essays to My Access. Students could select a topic they liked from a list of available topics. They first composed their writing in Microsoft Word and then copied and pasted their essays into My Access. A total of 124 essays were collected.

Before describing how the data was analyzed to answer the research questions stated above, the scoring and feedback mechanisms used in the My Access system will be briefly introduced. When students submit their essays to My Access, the writing system provides an automated holistic score on a 6-point scale along with immediate diagnostic feedback in terms of the following five categories (focus and meaning, content and development, organization, language use and style, and mechanics and conventions). Hence, students might find out their strengths and weaknesses and focus more on particular domains in which they receive lower scores.

In addition to providing scores on each of the five major categories, the My Access program also provides learners with written feedback. The feedback includes some written comments and suggestions on each of the five categories mentioned above.

In addition to these different types of general feedback and comments, My Access also provides detailed grammar feedback. When students click a writing tool called My Editor, they are provided with feedback on various grammar errors. This function is similar to the grammar checker included in Microsoft Word. One example of errors identified by My Editor is shown in Figure 1.

After- School Job

In nowadays, people<Misused words>???<Punctuation errors> attitude toward after-school job<Missing articles> has changed a lot. Before, most teachers and parents would think that a student???<Punctuation errors> responsibility is studying hard. They disliked this idea of after-school job<Missing articles> because they were worried about students??? academic<Punctuation errors> performance. But<Clause errors>, now more and more people think after-school job<Missing articles> is very important in students<Misused words>??? lives<Punctuation errors>.

Why does after-school job<Missing articles> become so important? The main reason is that working experience<Similar words> has become a crucial factor when companies want to recruit people. Studying is still important but only studying without any working experience in society is not enough for students to deal with the working field. So, that???<Punctuation errors> why there are so many students asking for internship during vacations. Although they are unpaid, this chance<Similar words> can provide them good working experience<Missing prepositions (ESL)><Similar words> they never learn in school. Besides, more and more students acknowledge that studying without any practical<Possible word confusions> experience is useless. Through practical<Possible word confusions> experience, they can improve or revise what they learn in school.

Apart from earning working experience<Similar words>, students will have additional income to make use from their after-school job. If a person wants to buy something, he must make money by himself. In this way, students will feel that making money is<Errors in word form> a hard work. They will become considerate toward their parents. For some students, their incomes are all from their after-school jobs for their parents can???<Punctuation errors> afford their expenses. For example, many students of this kind will find tutoring as their jobs. This job may be the most suitable one for them due to its regular working hours and good pay. They can take care of<Similar words> both their studies and jobs.

In fact, students are part of the<Unnecessary articles> society. They are included in society. After-school job<Missing articles> can also provide them effective<Possible word confusions> learning<Missing prepositions (ESL)>.

Figure 1. Grammar Errors Identified by My Editor

As stated in previous sections, the goals of this study were to further examine the accuracy of automated scoring and the feedback quality of My Access. To understand if there is a strong correlation between the scores assigned by My Access and human raters, it is necessary to compare the scores assigned by human raters and My Access.

These students' essays were submitted to My Access and copies of these essays were also sent to two human raters. These two raters are both senior high school English teachers in Taiwan. They were asked to give a holistic score based on the same 6-point rating scale used in My Access. With the two different sets of scores, the correlation between the scores assigned by My Access and the human raters could then be calculated.

As for the quality of the feedback, it is necessary to closely examine the macro-level feedback (content and structure) and the micro-level grammar feedback provided by My Access. Through close examinations of these different types of feedback, the strengths and limitations of AWE systems could be determined.

RESULTS

The Correlations between Human Rater Scores and My Access Scores

There were 124 essays submitted to the My Access system; however, five of the essays were rated as "off-topic" by My Access and did not receive any score. This left a total of 119 essays. The mean score assigned by MY Access for 119 essays was 4.12. The highest score was 6 and the lowest was 3.

The scores assigned by the human raters are summarized below in Table 1. The average score assigned by the human raters was 2.99. The highest was 6 and the lowest was only 1.

Table 1. Average Scores Assigned by Human Raters

	Mean	Std. Deviation
Rater 1	3.38	1.05
Rater 2	2.60	.938
Average	2.99	.916

Based on the raw scores, it is clear that My Access consistently assigned higher scores to this group of EFL students. This happened more often with the lower level EFL writers. They would score a 1 or 2 from the human raters, but they would get a score of 3 from the My Access system.

In addition to the mean scores, it is also important to test the correlations of scores assigned by the human raters and My Access. The correlation between the two human raters was high ($r=.688$ **, $P<.01$). However, the correlation between the human raters and My Access was not that strong. There was a significant but weak correlation ($r=.260$ **, $P<.01$). The statistical results were not surprising because the mean scores already indicated clearly that the My Access system tended to assign higher scores to students.

In addition to the "overestimation" problem, another more serious problem is that sometimes My Access would also assign improper scores to students' essays. Some essays received 4 points and 6 points from My Access. However, the human raters only assigned 1 point and 2 points to these essays. Although there were not many cases in our study, these cases should be treated carefully. These improper scores might confuse some ESL students.

Another interesting observation concerns the correlation between the length of an essay and its score. According to Herrington (2001) and Chen & Cheng (2006), longer essays are more likely to receive higher scores from My Access. We also had similar impressions in reading through the students' essays.

The Quality of Macro-level and Micro-level Feedback

As introduced in previous sections, My Access offers both macro and micro feedback. The macro-level includes the following types of feedback: focus and meaning, content and development, organization, language use and style, and mechanics and conventions. Besides, students are also were provided with micro-level grammar feedback on grammar, spelling, and punctuation. This feedback is provided by a tool called My Editor

For the macro-level feedback, the feedback was largely the same for learners who obtained the same scores from My Access. The system mostly gave some general guidelines and reminded the writers that they had to pay attention to some writing rules. My Editor can provide many different types of error messages. Based on the 119 essays graded by My Access, a small learner corpus was compiled. The learner corpus helped identify the most common errors provided by My Access. The 30 error types arranged by their frequencies were provided in Table 2. This list below is not a comprehensive list of all the error types provided by My Access, but it represents the majority of the errors identified by My Access.

Table 2. Error Messages Provided by My Access

Error Types (Frequencies)
1. Punctuation errors (723)
2. Spelling errors (318)
3. Similar words (289)
4. Clause errors (270)
5. Subject -Verb agreement (168)
6. Missing articles (168)
7. Pronoun errors (ESL) (83)
8. Misused words (64)
9. Punctuation errors (ESL) (50)
10. Possible word confusions (40)
11. Preposition errors (1) (28)
12. Noun phrase consistency errors (23)
13. Nouns: mass or count? (21)
14. Unnecessary prepositions (21)
15. infinitive or -ing form (17)
16. Unnecessary prepositions (ESL) (16)
17. Capitalization errors (16)
18. Unnecessary articles (14)
19. Formatting errors (13)
20. Double words (10)
21. Missing prepositions (ESL) (9)
22. Contractions (9)
23. Adverbs or adjectives (8)
24. Errors in word form (7)
25. Redundant Expressions (6)
26. Passive voice (6)
27. Verb Tense (5)
28. A vs. An (5)
29. Word Choice: Nouns (5)
30. Wordy expressions (1)

Since we could not find enough examples for some types of errors, we only focused on the top ten error messages in the corpus. Two ESL teachers were asked to review the error messages and make comments about the quality of the feedback. The results for each of these ten types of error messages are discussed below.

1. Punctuation errors (723): The inappropriate types of feedback were shown in Table 3. The accuracy level of this error message was low. Errors of extra space or omission of a space can be detected. However, abbreviations or apostrophe “s” in writing would be flagged as punctuation error. This actually provides a huge number of “false alarms”.

Table 3. Examples of Inappropriate Feedback on Punctuation

- | |
|---|
| <ol style="list-style-type: none"> 1. It<Punctuation errors> is only[<Punctuation errors> sewing] that I am greedy. 2. I am clumsy. because<Punctuation errors> I like sewing very much. 3. That??s<Punctuation errors> right 4. it??s<Punctuation errors> almost impossible to imagine a life without electricity. |
|---|

Note: the ??? symbols are in fact ‘s in students’ essays

2. Spelling errors (318): The inappropriate types of feedback were shown in Table 4. The accuracy of this message was acceptable for most cases. However, the system failed to detect proper nouns. Most proper nouns were not accepted and were highlighted as errors. Contractions would be flagged as spelling plus punctuation errors.

Table 4. Examples of Inappropriate Feedback on Spelling

- | |
|--|
| <ol style="list-style-type: none"> 1. I didn<Spelling errors>??t<Punctuation errors> know what else I could do. 2. the king of invents<Possible word confusions> Thomas Alva<Spelling errors> Edison<Spelling errors>! 3. she just sleeps<Spelling errors> and it will be chaotic if waking her up. |
|--|

3. Similar words (289): The accuracy of this message was very low. This error message was often redundant and confusing for ESL students because it failed to provide useful information. It seems that an expression repeatedly occurring in an essay would be flagged as a “similar word”. However, students actually do not need to modify their writing in most cases. Some examples of inappropriate feedback are shown in Table 5.

Table 5. Examples of Inappropriate Feedback on Similar Words

- | |
|---|
| <ol style="list-style-type: none"> 1. People are wandering on the street, looking very frightened<Similar words>. 2. I can learn the experiences<Similar words> from you so I know you more! 3. Many parents are not aware of<Similar words> the problems that young children watch too much TV. 4. When sitting in front of<Similar words> a TV set and fixing their eyes on the screen for several hours, |
|---|

4. Clause errors (270): This error message is similar to “fragment” in Microsoft Word. The accuracy of this message is low. This error message can be useful for incomplete sentences; however, this message was observed for any sentence beginning with AND or BUT. In addition, some prepositional phrases were highlighted if they appeared at the very beginning of a sentence. Some examples of inappropriate feedback are shown in Table 6.

Table 6. Examples of Improper Feedback on Clause Errors

1. As<Clause errors> a result, tides and waves are intermittent sources of energy
2. In<Clause errors> 1879, he changed to use the silk of carbon to make the lamp shines
3. One<Clause errors> day, my mother known this situation.
4. Next<Clause errors> day, (what??/)
5. On<Clause errors> the other hand, if parents give their children an allowance without requiring them to earn any spending money,

5. Subject-Verb agreement (168): This error message could be useful to remind students about the agreement between the subject and the verb. However, it was not accurate in some cases, as is shown in Table 7.

Table 7. Examples Inappropriate Feedback on Subject-Verb Agreement

1. Of course it is not his will that go<Subject-verb agreement errors> to war.
2. There<Punctuation errors> are<Subject-verb agreement errors> two sports class a week on general in the school of Taiwan.
3. Second<Subject-verb agreement errors>, if I were give<Subject-verb agreement errors> up, I would become a person to stop half way<Missing articles>.

6. Missing articles (168): This error message appears when students do not add in plural marker when using a noun phrase. The more appropriate feedback should be “use the plural form”.

Table 8. Examples of Inappropriate Feedback on Missing Articles

1. I thought I could finally get ride of these boring exams<Missing articles> in junior high,
2. I think that the most help thing I learn was I have learned keeping to learn new thing<Missing articles>
3. Then spiritually, I lack financial resource<Missing articles> to fulfill my thirst.

7. Pronoun errors (ESL) (83): For this error message, it was used to show the problems with pronouns. Some proper nouns were highlighted but learners might have difficulties understanding why these pronouns were highlighted. Some inappropriate feedback was shown in Table 9.

Table 9. Examples of Inappropriate Feedback on Pronoun Errors

1. One<Pronoun errors (ESL)> has to work very hard in order to achieve their goals.
2. What a beautiful world we enjoy<Pronoun errors (ESL)> because TV!
3. I<Pronoun errors (ESL)> couldn't image life without electricity.
4. my husband always claimed that we<Pronoun errors (ESL)> didn't want to have babies without mentioning our difficulty in conceiving a baby.
5. Before that day, I<Pronoun errors (ESL)> didn't hear her speaking any heavy

8. Misused words (64): This error message show up when the possessive forms or some unusual word strings are used. However, the underlying rules for these errors were not clear.

Table 10. Examples of Inappropriate Feedback on Misused Words

1. I have seed the dusk with seven colors<Misused words> sky
2. Parents can draw children<Misused words>???'s attention to the coverage on the results of poor money management.
3. In nowadays, people<Misused words>???'s attitude toward after-school job has changed a lot.
4. However, the most serious problem is that it will divide young people<Misused words>???'s time to get along with their family.
5. Many parents hold different attitudes toward the ways to giving children<Misused words> allowances.

9. Punctuation errors ESL (50): The error message often shows up when learners use the expression “such as”. However, it also shows up in some other sentences. The underlying rules were not clear.⁸

Table 11. Examples of Inappropriate Feedback on Punctuation errors ESL

1. Furthermore, if kids know the hardship of earning money, they will appreciate what they own instead of requesting too many unnecessary utilities,<Punctuation errors (ESL)> such as toys or luxuries.
2. When I was sad , she will comfort me, taking toilet paper for me,<Punctuation errors (ESL)> when I catch a cold
3. Many of my classmates started to have different kinds of teaching experiences,<Punctuation errors (ESL)> such as teaching teenagers or teaching children, in sophomore year.
4. Although people have more powers and ideas to control the nature,<Punctuation errors (ESL)> such as natural disasters.

10. Possible word confusion (40): This error message appears when some words are not used properly. In addition, the message also pops up when some easily confused words (affect vs. effect) were used by ESL students. It shows up to remind the learners that they might use the word incorrectly. It could be useful as a reminder; however, it could be very confusing if learners believe that each error message should be corrected.

Table 12. Examples of Inappropriate Feedback on Possible Word Confusion

1. Education is not to make everyone the same but to discover<Possible word confusions> the differences and respect the development of every individual.
2. Placing a limit on the amount of homework<Similar words> is only to ensure<Possible word confusions> that the students???' have<Punctuation errors> time to explore their own interests.
3. Winner of the Nobel Peace<Missing articles> Prize in 1991, she leads without armies, media manipulation or economic<Possible word confusions> might.
4. Besides, more and more students acknowledge that studying without any practical<Possible word confusions> experience is useless.

Based on the analysis above, it is obvious that My Access provided too many false alarms to the students in this study. Many error messages for spelling, punctuation, similar

⁸ The My Access system was first designed for native speakers, and later it was also developed to accommodate the needs of EFL learners. The system has some categories of errors which are marked with ESL.

words, and word confusion generated by My Access were not errors. Students, however, needed to read through this confusing feedback and occasionally find some useful feedback in these categories. Generally speaking, the feedback function generated too many unnecessary error messages. The following types of error messages were often false alarms: punctuation errors, spelling errors, clause errors, similar words, pronoun errors, possible word confusion. When receiving error messages in these categories, ESL students would need to carefully assess the correctness of these feedback messages. For lower level learners, it would be challenging to distinguish between real errors and false alarms.

Although My Access was able to provide about 30 or more types of error messages to the learners, a review of the essays showed that My Access still failed to capture some common errors in the student's essays. To further investigate the quality of grammar feedback provided by My Access, fifteen essays were randomly selected from the 119 essays graded by My Access, and the error tags were manually removed from these essays. Two experienced English teachers were asked to help grade these fifteen essays and they were asked to provide thorough feedback on these essays. The feedback from My Access and these two human raters was then analyzed and compared.

The two teachers identified 134 errors that had not been found and treated in My Access. These errors are summarized in Table 13 below. To further classify these errors, different error labels were assigned in Table 13 below. Only errors which appeared 3 times or more are shown.

Table 13. The Errors Not Found by My Access Program

Confusing sentences	we could not <u>cut our time very clearly</u>	9
	Since it was already <u>very late night</u>	
	I talk to myself <u>whether happen anything or do anything job</u>	
	<u>we possible nothing to making life</u>	
	<u>make us keep a good mind</u>	
Verbs	in order <u>to making life</u>	9
	Instead of <u>stuff</u> knowledge into students	
	To <u>look</u> my favorite character every hour on the hour	
Determiner	it seems that I am <u>an actress to do the performances</u>	8
	I practice <u>paint</u>	
	They are included <u>in (the) society</u>	
Missing words	The amount of homework for students <u>in (an) English department</u>	8
	... <u>sports are important</u> part of school life	
	<u>they not use</u> so large place	
Modifiers	The baby <u>wouldn't come out (until)</u> two weeks later,	8
	<u>they not use</u> so large place	
	The traditional <u>Taiwan</u> puppet show	
Spelling	<u>the nature</u> power	7
	dancing is <u>a very health</u> sport	
	I think <u>I well</u> continue	
Connector	when I <u>sited</u> on the bicycle	7
	And he <u>kelp</u> me company	
	When you watch, <u>you well cry, laugh with story</u>	
Prepositions	We saw the Lan-Yan plain, <u>thousands lights</u> came out from each house	6
	<u>kelp me company</u> to overcome difficulties. <u>That I didn't fell</u> alone	
	<u>In</u> nowadays	
Prepositions	<u>on</u> general	6
	<u>get on well</u> with others	

	But when <u>the clock had been invented.</u>	
Tense	I <u>also learn playing the piano (happened in the past)</u> But when the clock had been invented.	6
	We want to run with time <u>but we do not</u>	
Modal	Having spent so much time on TV programs <u>must result in</u> some bad effects	5
	<u>a good competent</u>	
Part of Speech	Thus, <u>discovery the beauty of Taiwan</u> is my biggest leisure time activity <u>difficult to breath</u> gave me <u>many pleasant memory</u>	5
	<u>a lot of skill</u>	
Countable	I learn <u>different thing</u> <u>many informations</u>	4
	<u>provide them (with) effective learning</u>	
Missing Prepositions	especially (on) <u>the east coast</u> The major portion of <u>the products selling from Pili</u>	3
	<u>we exposure to</u>	
Voice	<u>Society was change too quickly</u>	3

In addition to these two major problems discussed above, My Access failed to provide feedback on the same type of errors in the same essay. Some types of errors (e.g., subject-verb agreement) were correctly detected and marked the first time the error was encountered. However, when the error appeared again later, it was ignored.

DISCUSSION

Since automated writing evaluation (AWE) systems have become popular in ESL teaching, it is necessary to examine the scoring accuracy and feedback quality more closely and carefully. For automated scoring, the My Access system often assigned higher scores to the students' essays in this study. On average, the score for My Access was about 4.11, but the score for the human rater was 2.99. The findings were similar to those of Wang and Brown (2007), who also found that My Access tended to assign scores higher than those assigned by human graders. A six point grading scale, a one point difference could have significant impact on the students who wrote the essays. If students can obtain a 4 or 5 from My Access, then they are very likely to consider their writing above average. However, from a human rater's perspective, students who achieved a 3 or 4 might still need to polish their writing.

Moreover, the correlations between My Access and the human raters were significant but weak ($r=.026^{**}$, $P < .01$) because human raters assigned lower scores than My Access system. Previous studies have investigated the agreement between AWE systems and human raters when the subjects' first language was English. Burstein et al. (2003) showed that the agreement rate on holistic scores between the Criterion e-rater and human expert graders is typically 97%. According to Elliot (2001), My Access agrees with the expert scores within one point 99% of the time, and exactly with the expert grader scores 73% of the time. Cohen et al. (2003) pointed out that My Access can reach about 0.82 correlations. The findings of the present show an obvious difference between native English writers and EFL writers. Is it because My Access is mainly designed for grading essays written by native speakers? More research must be conducted to explore the differences found by these empirical studies.

The weak correlation found in this study also supported the findings from previous studies on ESL/EFL learners. Earlier studies found that the scores assigned by the My Access system were inadequate or doubtful. Only 13% of the students in Yang (2004)

thought that the scores they received from My Access were adequate, whereas more than a half felt uncertain about the scores. Chen and Cheng (2006) reported that students were much less satisfied with the scores assigned by the electronic rater. Some of them felt that the My Access program gave them improper scores. In Chen and Cheng's (2006) study, students reported that the machine scores they received from My Access were different from what they received from teachers.

From the following two comments made by students in Chen and Cheng's (2006) study, the observation that My Access tended to assign higher scores received strong support.

"The grading program was too kind to me! I mean sometimes I still can get high scores without adding a conclusion in my essays. I wrote only the introduction and the body, and the program gave me a score of 5. How come an essay without a conclusion can get such a high score?"

"When I submitted the same piece of writing to both my writing teacher and My Access, I found that there was a big gap between the grades given by my teacher and those given by the software, so I don't really trust the program's grading function."

Based on the findings of these different studies, the My Access program has been too generous to ESL/EFL students, and the high scores might lead these students to believe that their writing is excellent. Writing teachers might need to remind students about the possibility that the My Access system might overestimate their writing abilities

The Feedback at the Macro-level

The problems of the Macro-level feedback were noted in several previous studies. This feedback is repetitive and not useful. Yu & Yeh (2003) found that most students considered feedback from My Access useful for writing their first revisions, but they soon felt the program was not helpful because of the repetitive nature of the when they submitted subsequent revisions. Some students also pointed out that the computerized feedback was not specific enough to help them revise their essays. Students in Yang's (2004) study also pointed out that the fixed, repeated feedback became useless to them. Chen and Cheng (2006) found that My Access failed to give specific feedback in the content and rhetorical aspects of students' writing. The students also pointed out that the feedback from My Access was vague and repetitive so that they could not make good use of the feedback.

The findings of the present study showed that the macro-level feedback provided by My Access was inadequate and not helpful. The results are not surprising because the state of the art NLP systems cannot easily process the meaning in learners' essays. The system cannot understand what the ESL writers want to express in their writing. Thus, it is very difficult for the AWE systems to offer good feedback on content and organization.

The Feedback at the Micro-level

As for the quality of the grammar feedback, we found that the grammar feedback provided by My Access might give inadequate feedback to ESL learners. Several items are problems were noted.

A. Misdiagnosis- The system highlighted some words and sentences but in fact the learners did not need to make corrections. The system also treated some correct sentences as incorrect ones.

B. Several types of errors such as run-on sentences, incorrect usage and structures were ignored.

C. Inconsistency of error correction: The same type of error in an essay was not marked consistently.

Many ESL students would closely follow the suggestions made by the automated essay evaluation system. The following five types of error messages: punctuation, spelling, clause errors, similar words showed up so many times in the mini-corpus. Many of these messages in fact were not necessary or useful. The students thus could not benefit much from these

various messages.

Other categories of error messages (subject-verb agreement; missing articles; misused words; Nouns: mass or count; unnecessary prepositions) might be more useful for ESL learners. However, when these error messages appear with the “false alarms”, learners might choose to ignore them or might not use them properly.

Another serious problem of the My Access checker is that it still fails to deal with the many errors produced by ESL learners. It should be noted that there are a wide variety of errors in ESL students’ essays. For instance, there were about 61 types of major errors found in CLEC corpus (Gui & Yang, 2002). Some of these errors were transfer errors, and students’ mother tongues often have a great influence on their L2 writing. The checker in My Access was originally designed for native English writers, and then expanded to process the writing of ESL learners.

It is not surprising that the My Access program has these problems. One of the best grammar checkers, the one included in Microsoft Word, has similar problems. It is rather complex for computer programs to deal with second language writers’ grammar errors. Second language learners’ error patterns are different from those of native. My Access was originally designed to process and grade the writing of native speakers. When the program was used to grade the writing produced by ESL learners, the level of difficulty was higher and the complexities greater, and thus the standards became different. In an article reviewing various computer assisted testing tools, Carr (2006) pointed out that it is hardly surprising that automated essay scoring is not yet ready for operational testing in a high-stakes second language context.

However, there are several possible ways My Access can make its products more user friendly and useful for ESL learners around the world. First, they can modify the scoring mechanism and make it stricter. Second, the error messages of the following categories (punctuation, spelling, clause errors, and similar words) should be reduced. Third, the checker should be able to detect more errors produced by ESL learners with different L1 backgrounds.

CONCLUSION

The demands for a good automated essay scoring system are very strong in many ESL settings. Several universities and college in Asia are also using the My Access program and the ETS Criterion program. The findings of this study remind teachers and students about the limitations and problems of My Access. Teachers and students who use My Access should be more careful in interpreting the scores assigned by this system. When students receive 5-6 points from My Access, it does not mean that they have reached the standard set by TOEFL or other high-stake tests. It is recommended that writing teachers still check the essays written by students if students have doubts about the scores and the feedback provided by My Access. Moreover, even though it is possible to use My Access in some testing situations, using My Access in high-stakes tests in ESL/EFL contexts is very likely to encounter problems.

As for detailed grammar feedback from My Access, it can be used but should be used with caution. Students should constantly consult writing teachers and tutors if they are not sure about the feedback they received. It might take a long time before My Access can offer accurate and comprehensive feedback on ESL learners’ writing errors. Many errors still can not be detected or corrected by the automated scoring system. The holistic scores assigned by My Access might be used as a reference to probe ESL learners’ writing competence, but it is not very helpful for ESL students to use the grammar feedback generated by the automated scoring system. Because of heavy teaching loads, writing teachers might expect that programs like My Access and Criterion can help them accurately grade and correct students’

writing assignments. However, these AWE programs at this point might not be able to meet ESL writing teachers' high expectations.

Although these automated writing evaluation systems are far from perfect at this stage, it does not mean these AWE programs are not useful for second language writers. These programs can still be used in college ESL writing programs and they can help teachers provide some useful feedback. However, before they can benefit more ESL learners, the accuracy of the scoring and the quality of the feedback should be greatly improved. The program should be more careful in assigning high scores to ESL learners because these improper scores might lead learners to believe that they have reached an acceptable level. The feedback mechanism should try to address more errors made by ESL learners. A more robust grammar checker should be designed to identify and correct learners' errors more accurately. The feedback component is clearly crucial if writing teachers want to use My Access as a self-access learning tool.

Some ESL writing teachers worry that these automated writing evaluation systems might replace them. However, various studies on these systems indicate that in fact these systems are not as smart as teachers. Many developers of these programs suggest that the automated essay grading tool should be used only as a supplement to classroom instruction (Burstein, Chodorow, & Leacock, 2003; Burstein & Marcu, 2003). Language teachers should not assume that AWE systems can, or will, replace human teachers. At this stage, these AWE programs can only detect the surface errors in ESL students' writing. Many lexical and syntactic errors like collocation errors and sentence fragments cannot be detected or corrected by these writing tools. More importantly, these programs still fail in examining the content or rhetorical aspects of students' writings. Even if these AWE programs can detect more errors in the future, the programs still will not be able to show students how they can improve the content and organization of their writing. For these and other more sophisticated writing problems, only writing teachers and tutors can provide valuable suggestions to individual students. For the foreseeable future, an automated essay evaluation system will not be able to provide this type of learning environment.

REFERENCES

- Attali, Y. (2004). Exploring the feedback and revision features of Criterion. Paper presented at the National Council on Measurement in Education (NCME) in San Diego, CA. April, 2004.
- Burstein, J., & Marcu, D. (2003). Developing technology for automated evaluation of discourse structure in student essays. In M. D. Shermis & J. Burstein (Eds), *Automated essay scoring: A cross-disciplinary perspective*. Hillsdale, NJ: Lawrence Erlbaum.
- Burstein, J., Chodorow, M., & Leacock, C. (2003). Criterion: Online essay evaluation: An application for automated evaluation of student essays. *Proceeding of the Fifteenth Annual conference on Innovative Application of Artificially Intelligence*, Acapulco, Mexico.
- Carr, N. (2006). Computer-based testing: prospects for innovative assessment. In L., Ducate & N. Arnold (Eds), *Calling on CALL: From Theory and Research to New Directions in Foreign Language Teaching* (pp. 298-312). San Marcos, TX: CALICO.
- Chen, C-F. & Cheng, W-Y. (2008). Practices and Perceived Learning Effectiveness in EFL Writing Classes. *Language Learning & Technology*, 2008, 12, 2, June, 94-112
- Chen, C-F. & Cheng, W-Y. (2006). The Use of a Computer-based Writing Program: Facilitation or Frustration? Paper presented at the 23rd International Conference on English Teaching and Learning in the R.O.C. May 27, 2006. Kaohsiung: Wenzao Ursuline College of Languages.
- Chodorow, M. & Burstein, J. (2004). Beyond essay length: Evaluating e-rater's performance

- on TOEFL essays. Educational Testing Service Research Report 73. Educational Testing Services.
- Elliot, S. (2001). IntelliMetric: From here to validity. Paper presented at the Annual Meeting of the American Educational Research Association (AERA), Seattle, WA. April, 2001.
- Elliot, S., & Mikulua, C. (2004). The impact of My Access! use on student writing performance: A technology overview and four studies. Paper presented at the Annual Meeting of the American Educational Research Association (AERA) in San Diego, CA. April, 2004.
- Ferris, D. (1999). The Case for Grammar Correction in L2 Writing Classes: A Response to Truscott. *Journal of Second Language Writing*, 8(1), 1-11.
- Grimes, D., & Warschauer, M. (2006). Automated Essay Scoring in the Classroom. Paper presented at the American Educational Research Association (AERA) Annual Conference, San Francisco, CA.
- Gui, S. and Yang, H. (2002). *Chinese Learner English Corpus (CLEC)*. Shanghai: Shanghai Foreign. Language Education Press.
- Herrington, A. (2001). What happens when machines read our students' writing? *College English*, 63 (4), 480-499.
- Higgins, D; J. Burstein; D. Marcu; C. Gentile, (2004). Evaluating multiple aspects of coherence in student essays. Retrieved May, 20, 2006, from: http://www.ets.org/Media/Research/pdf/erater_higgins_dis_coh.pdf
- Truscott, J. (1999). The case for "The case against grammar correction in L2 writing classes": A response to Ferris. *Journal of Second Language Writing*, 8(2), 111-122.
- Ware, P., and Warshauer, M. (2005). Electronic feedback and second language writing. In K.Hyland and F. Hyland (Eds.) *Feedback and second language writing* (pp.1-29). Cambridge: Cambridge University Press.
- Wang, J., & Brown, M.S. (2007). Automated essay scoring versus human scoring: A comparative study. *Journal of Technology, Learning, and Assessment*, 6(2). Retrieved [date] from <http://www.jtla.org>.
- Yang, N. D. (2004). Using My access in EFL writing. *The proceedings of 2004 International Conference and Workshop on TEFL & Applied Linguistics* (pp. 550-564). Taipei, Ming Chuan University.
- Yu, Y. T., & Yeh, Y. L. (2003). Computerized feedback and bilingual concordancer for EFL college students' writing. *Proceedings of 2003 International Conference on English Teaching and Learning in Republic of China* (pp. 35-48). Taipei, Crane.

Conversational Narrative Inquiry into Learning to Teach English

Jin-shan Chen (陳錦珊)

Chihlee Institute of Technology

jschen@mail.chihlee.edu.tw

This paper reports partial results of an in-depth analysis on the conversational narratives of an inquiry group of prospective TESOL teachers, who collaboratively sought for the ways of teaching English through a campus-based practicum teaching. The discussions highlight the process of learning to teach, especially the dilemmas and challenges, which the prospective teachers encountered throughout the practicum teaching. Adopting a holistic approach, the assembled group discussions were analyzed through identifying the topic shifting in the conversational discourse on the attempt of discovering meaningful patterns in the prospective teachers' conversations. As a result, a three-level inquiry was identified in the conversational narratives – *inquiry into the embedded context*, *inquiry into pedagogical knowledge*, and *inquiry into personal experiences*. Conversational narrative inquiry into embedded context takes a global scope concerning the issues relevant to the external factors, such as the institutional contexts, distribution of social goods, as well as the interactional dynamics, power and voices, in the inquiry group. Conversational narrative inquiry into pedagogical knowledge draws more attention to practical and technical considerations in learning to teach English. Conversational narrative inquiry to personal experiences deals with the issues of the prospective teachers' self-conception and self-determination, as well as the connections and contradictions of the past and present experiences when engaging in the teaching activity. Significantly, the findings suggest that prospective TESOL teachers employ a cross-level analysis in the process of learning to teach. The results indicate that the prospective teachers, like more experienced successful teachers, actively and continuously construct and reconstruct their knowledge and beliefs regarding teaching and learning English through practical teaching in contexts.

INTRODUCTION

Bruner (1986) explained that people construct themselves through narratives and make sense of their lives by telling stories. Knowledge-making, learning and development are recognized as an active, creative and interpretive process through which the telling and retelling of one's stories provide the framework for the construction of the world.

Bell (2000) further addressed that human beings make sense of random experiences by the imposition of story structures. We select elements of story structures and pattern those chosen elements in ways that reflect the story structures available to us. That is, although the notion of stories is common to every society, one of the defining features of a culture is the story structures it uses to construct the world. The shape of our stories, the roles of our stories, the development of the story lines, and the possible story structured are all shaped by the cultural contexts within which we were raised. Accordingly, a key to understanding a person from another cultural group is to examine his or her stories and become aware of the underlying assumptions of those stories. Narrative Inquiry involves working with people's consciously told stories, recognizing there are always deeper stories of which people are often unaware. When research enters the domain of thinking and knowing, inquiry goes into the examinations of the interactive systems among participants, tools and means, and

communities or societies.

According to Freeman (1996), in the research tradition of language teacher education, teaching is mainly studied through examining students' learning and achievement. Little attention has been paid to the process of how they come to know what they know and do what they do as teachers. Following the paradigm of Behaviorism, teaching is examined through observable behaviors and actions which could be quantified and assessed via learning outcomes. The mental processes of teachers' thinking and knowing are of little concern.

However, around the mid-1970s with the development of Cognitivism, and more lately Socio-cultural theories, teaching has been recognized as a complex cognitive activity, and teachers have to be viewed as individuals who learn, shape and are shaped by the activity of teaching. This development has both led to the new conceptual understanding of teaching and diversification of research methodologies in the research of language teachers. Traditional research methodologies which mainly focus on the behavioral observations, product-based orientation, and cause-effect reasoning and interpretation are no longer adequate.

Freeman (1996) further points out that when the research targets at examining the invisible mental processes, all the data are rendered in languages. Languages provide the linkage in the research between the unseen private worlds of teachers and public worlds of data collection processes. Languages are, hence, always used to express – and represent – thoughts. From this perspective, Narrative Inquiry empowers teachers by giving voices to what they know, enabling them to articulate how they know, and providing a link among their lived experiences, practical knowledge, and thinking and believing. Narrative Inquiry as a research approach also enables researchers to work with teachers collaboratively shifting from traditionally outsiders' viewpoints to insiders' interpretations and analysis of the cognitive activity of teaching.

Following this research paradigm, this paper reports partial results of an in-depth analysis on the conversational narratives of an inquiry group of prospective TESOL teachers who collaboratively sought for the ways of teaching English through an alternative form of practicum teaching, a campus-based supplemental English instruction program. The discussions highlight the process of learning and teach, especially the dilemmas and challenges which the prospective TESOL teachers encountered throughout the practicum teaching.

LITERATURE REVIEW

Since the 1980s, teachers' practice became even more complex as they were situated in personal and situational histories. This raises the specific question in teacher education of what the role of educational institutes can be in learning to teach. Teacher educators and researchers are now asking how schools as sociocultural environments mediate and transform what and how teachers learn. Unfortunately, the vast majority of preservice and inservice teacher education programs ignore these questions. They continue to operate with knowledge transmission perspective, applying a set of prescriptive top-down teacher training packages (Freeman & Johnson, 1998). The question of how teachers learn to teach requires more explorations and inquiries.

Warford (2003) conducted a qualitative study in order to understand the preconceptions novice TESOL teachers might have about teaching English language. Participants were nine student enrolled in one of the two courses offered in a TESOL teacher education program. Findings suggest that novice teachers, like their more experienced mentor teachers, have a system of metaphors to conceptualize teaching. The apprenticeship of observation, being language learners at school setting for years, has a great impact on the teachers' thinking and believing. Warford suggests that the preliminary evidence of metaphors, apprenticeship of observation, and adaptations of expert systems constructs a preliminary framework for

describing novice TESOL teacher thinking and knowledge construction.

Rust (1999) conducted a qualitative research on a conversation group formed by preservice, first-year, and second-year teachers to explore factors in teacher education and schools that support and hinder new teachers' work in urban schools. Research findings indicated that the preservice teachers' stories and concerns were qualitatively different from beginning teachers. The most distinctive difference was that the preservice teachers did not talk about their students in the manner that they owned their students, in a way that the first- and second-year teachers did. The preservice teachers paid more attention to instructional strategies, curriculum development and classroom management; their questions remained on technical level in more general terms. In discussion, Rust cautioned about that the lack of involvement with students, and lack of encountering with conflicts and contradictions of real messiness in school settings (e.g., worrying about tomorrow or who would work with the group of students next year) would make the practicum teaching simply survival for course requirement.

Besides, comparing the stories told by successful and tough first-year teachers, Rust found that the tough first-year teachers complained about their students and never spoke of their students in possessive manner like what the successful first-year teachers did. And the tough first-year teachers focused on the discussions of classroom management, while the successful first-year teachers were keen on their learning how to negotiate with the external system on behalf of their students. Taking the perspective that learning is a socially constructed process and community of learners are shaped by the inhabiting culture, Rust proposed that teachers need supported opportunities to reflect upon their own knowledge, explore their attitudes and beliefs, and extend the repertoire of skills and strategies of teaching. On the basis of the study, Rust suggested a conversation group as an effective mediated tool to promote and extend the support of teacher education program.

Rust's work brings up the valuable insights of the different concerns and stories of teachers with different degree of teaching expertise, ranging from student teachers to second-year teachers. More importantly, four implications could be drawn from this study. First, it seems clear enough now that the ownership of students and control over curriculum directly correlate to the happening of successful teaching. Second, similar to Johnson's (1996) contention, conflicts and dissonance in practical teaching may be the driving force for teaching upgrade. Furthermore, successful first-year teachers' stories reveal the concept that learning to teach also means learning to negotiate within and across the teaching activity. And the last, teacher education program should structure supportive resources, such as a conversation group, for encouraging teachers to tell their teaching stories and promoting reflection on critical issues relating to their teaching practices. In this work, Rust clear depicted the differences between novice and experienced teachers, and tough and successful teachers; however, the developmental process of how a teacher transformed herself/himself into a successful experienced teacher still remained unknown.

METHOD

The Setting

A campus-wide supplemental English instruction program was designed as an alternative form of practicum teaching at a vocational college in northern Taiwan for two consecutive semesters. The overall goals of the program were twofold: to assist students with low-achievement in English to improve their English learning, and to help prospective TESOL teachers develop their teaching profession. The supervising scheme mainly focused on encouraging the prospective teachers to develop themselves as reflective practitioners through personal and collaborative inquiry. In practice, all the prospective teachers were required to attend a weekly-three-hour inquiry group. The agency, voice and power, and

ownership regarding their teaching were the vital concerns throughout the inquiry process.

The Participants

At the first semester, eight prospective TESOL teachers participated in the study with one male and seven female at the age of 22 and 23. They were all the seniors in the Department of Applied English at the same college. However, at the second semester, two prospective teachers left the study because of some personal reasons. As a result, the number of participants reduced to be six female prospective teachers at the second semester. Nevertheless, since the group discussions were collaborative efforts, the talks of every prospective teachers were taken into accounts as valuable field texts regardless the time of departing the program.

Data Collection

Data collection started at the beginning till the end of the practicum teaching of the prospective teachers. The discussions of the weekly inquiry group were video-taped, and then transformed into texts words by words. The transcribed texts were returned to the participants for member checking. To ensure data triangulation, other types of data were also collected, including surveys, lesson plans and teaching materials provided by the participants, interview transcriptions, personal reflective journals, field notes and researcher memos.

Data Analysis

The data analysis of this narrative inquiry is inductive, generative, constructive, and subjective. Drawing on the perspective of holistic-content mode of reading, this inquiry primarily searched for the meaning of a story in its entirety through recursive movement between emerging themes of the whole story instead of fragment parts, like separated coding and categorization as widely used in traditional qualitative research (Lieblich, Tuval-Mashiach & Zilber, 1998). The analysis procedure included: (1) reading and familiarizing; (2) identifying important concepts to look for; (3) identify critical events; (4) identify significant others, and (5) identify personal beliefs.

RESULTS AND DISCUSSIONS

Results of the data analysis contain a substantial body of interactional dialogues, social dynamics and stories exchanged and shared among the participants all along the way of their carrying out teaching practices. A thick description is compiled with salient themes emerging from the conversation narratives collected from the weekly inquiry group discussions, accompanied with excerpts from other documentations. The assembled conversational narratives can be discussed as inquiries at three levels – *inquiry into the embedded context*, *inquiry into pedagogical knowledge*, and *inquiry into personal experiences*. Conversational narrative inquiry to the embedded context concerns the issues relevant to the external factors, such as the institutional contexts, critical and political issues, as well as the interactional dynamics, power and voices, in the inquiry group. Conversational narrative inquiry to pedagogical knowledge draws attention to practical and technical considerations on learning to teach English. Conversational narrative inquiry to personal experiences deals with the prospective teachers' self-conception and self-determination, as well as the connections and contradictions of the past and present experiences (Table 1).

Table 1 A Three-Level Inquiry into Learning to Teach English

Levels of Inquiry	Themes	Sub-themes
Inquiry into Embedded Context	Discussing the distribution of social goods	Identifying the learners' characters, identifying the goals of the program, identifying the learning problems
	Negotiating the social norms	Defining the program, defining the inquiry group, defining the curriculum
	Asserting the agency in teaching	Relying on the supervisor, arguing with the supervisor, participating in the policy making of the program
Inquiry into Pedagogical Knowledge	Knowing about the students → knowing the students	Identifying the students' needs and interests, identifying the students' attitude, indicating the students' duties in learning
	Knowing about the curriculum → knowing the curriculum	Content selection, curriculum coherence, meaningful learning to individual students
	Knowing about the teaching → knowing the teaching	Classroom interactions, teaching techniques, and alternative assessment
Inquiry into Personal Experiences	My English learning experiences	English learning strategies, meaningfulness of learning English, and the opportunities of exposure to English and of using English for communication
	My teaching style	Realizing the uniqueness and individuality in learning and teaching English for both parties, namely students and teachers
	My roles	Realizing the co-existing of multiple roles in teaching, role conflicting and role reconciliation

The three-level inquiries present themselves in chronological order, following the sequences of their becoming salience in the group discussions although most of the topics emerged repeatedly at different stages. It shows that the concern-dynamics of the group inquiry moved from the global to the local, the broader to the narrower, the external to the internal, and the universal to the relative. For example, at the first few weeks, the topics being discussed were mainly related to the framing of the situational contexts of teaching practices while a great proportion of talks concerned about the development of pedagogical knowledge and gave voices to the uniqueness of individuality in teaching and learning by the end of the program.

CONCLUSIONS

Conventional educational research embraced a behavioral traditional which devaluated the role of teachers by recognizing them only as technicians who simply transferred static facts and truths to students. The notion of teaching in this tradition is regarded as the transmission of knowledge. Teachers' performance is evaluated in accordance with the learning outcomes of students. The critical lack of empirical information concerning proper ways of preparing prospective teachers urged me to seek for possible interpretations on the

nature of, and the process of learning to teach English. This paper illustrates partial results of a series of research studies on learning to teach English through practicum teaching. The discussions included in this paper highlight the concern dynamics of the inquiry group formed by prospective TESOL teachers, as well as the dilemmas and challenges they encountered throughout the practicum teaching. Results indicate that the prospective TESOL teachers, like experienced successful teachers, actively and continuously analyze their teaching with caring, applying a holistic approach. More studies embracing multiple perspectives from different research paradigms should be carried out in order to answer the fundamental questions, like how language teaching is learned, and how teacher education resources and interventions can be structured to support the teachers' learning in a more appropriate and meaningful way.

REFERENCES

- Bell, J. S. (2002). Narrative inquiry: More than just telling stories. *TESOL Quarterly*, 36(2): 207-218
- Bruner, J. (1985). *Actual minds, possible worlds*. Cambridge: Harvard University Press.
- Freeman, D. (1996). The "unstudied problem": Research on teacher learning in language teaching. In D. Freeman & J. C. Richards (Ed.), *Teacher learning in language teaching* (pp. 351-378). Cambridge: Cambridge University.
- Freeman, D. (1996). The "unstudied problem": Research on teacher learning in language teaching. In D. Freeman & J. C. Richards (Ed.), *Teacher learning in language teaching* (pp. 351-378). Cambridge: Cambridge University.
- Lieblich, A., Tuval-Mashiach, R. & Zilber, T. (1998). *Narrative research: reading, analysis, and interpretation*. CA: Sage.
- Rust, F. O. (1999). Professional conversations: new teachers explore teaching through conversation, story, and narrative. *Teaching and Teacher Education*, 19-367-380.
- Warford, M. K. & Reeve, J. (2003). Falling into it: novice TESOL teacher thinking. *Teachers and Teaching: theory and practice*, 9(2): 47-65.

Psychosocial Sources of Anxiety in Academic Oral Presentations

Lei-yi Chen (陳蕾伊), Yu-mei Chen (陳玉美)

National Chung Cheng University

psyc222@gmail.com

The present study aims to assess the levels of anxiety and investigate the psychosocial sources of anxiety in academic oral presentations. The participants consisted of 18 graduate students in a TEFL program of a university in southern Taiwan. They were required to read and present selected chapters orally twice in a testing course. Their oral presentation performances were videotaped. After class, all the presenters were invited to review their presentation videos and interviewed regarding their feelings and thoughts about their oral presentations. A questionnaire was also administered to each presenter to assess levels and sources of anxiety. More specifically, Personal Report of Public Speaking Anxiety (PRPSA) was adopted to measure anxiety. The interview recordings were transcribed and analysed by the method of content analysis.

Preliminary findings show that the students were moderately anxious, suggesting that the level of anxiety was not so severe that the students could not cope with it; and the interplay of factors such as classmates' response and behaviours, audience familiarity, accuracy of pronunciation, self-perceived English oral proficiency and personality may have played a role on presenters' levels of anxiety.

INTRODUCTION

The English language is a powerful tool for communication. It is a global language that people of different languages use to communicate and engage in business. It also provides access to much of the world's knowledge (UNESCO, 1995). Thus, having a good command of the language is likely to bring social and economical advantages because it allows one to communicate with people from around the world and have first-hand access to the latest knowledge. To equip their people with this powerful skill, many countries have carried out educational reforms in recent years by switching the medium of instruction from their first languages to English, a second or in most cases, a foreign language in these countries. In Taiwan, some courses at undergraduate and postgraduate levels are now offered in English. For instance, the Electrical Engineering Department of National Tsing Hua University has included 15 requisite courses (47 credits) that are taught in English at the freshman level (China Times, 2004). At National Taiwan Normal University, many graduate departments also offer courses in English, such as the Department of History, Physics and Chemistry. In English-speaking countries such as the USA, UK, Australia and Canada, the number of foreign students enrolled in tertiary education has increased from four hundred thousand in 1980 to six hundred thousand in 1992. Moreover, about 1,700 universities from around the world have added about 3,300 courses taught in English from 2004 to 2007 (Lin, 2007).

Therefore, EFL or ESL students in tertiary education from around the world have an increasing opportunity to use English as their working language. This may bring many challenges for the learners. One of these challenges is related to speaking ability. A recent study on the oral communication needs of East Asian international students in the US indicated that students find leading class discussions the most difficult, followed by

whole-class discussion, small-group discussion, then fourthly, formal oral presentations (Kim, 2006). The challenges with academic speaking are not only linguistic in nature, rather, it also involves learners' sociocultural and psychological characteristics (Morita, 2000). These results are not surprising given the fact that speaking activities have long been identified as an anxiety-provoking activity in both foreign language classroom (Young, 1990); Hilleson, 1996) and various public speaking contexts (Jackson & Latance, 1981). Considering the dual task of learning English and using it to express verbally, such a task may bring considerable stress at least for some EFL/ESL speakers.

Previous studies suggest that in both contexts – language classrooms and public speaking situations, speaking activities are likely to arouse anxiety. However, there are two problems with the applicability of the past anxiety studies to the context of graduate seminar with EFL speakers. First, previous studies on foreign language speaking anxiety have been focused on language learning anxiety (Cheng, Horwitz & Schallert, 1999; Saito & Samimy, 1996). The speaking tasks in these studies were typically conversations or speeches, little information was found on academic speaking anxiety. Secondly, many anxiety studies in communication research only focus on presentation or speech anxiety of native English-speakers. It has been shown that language anxiety tends to correlate with performance in second language but not in first language (MacIntyre & Gardner, 1994). This means that a more anxious student tends to perform worse in a second language but not necessarily in his first language. Therefore, EFL speakers are likely to experience anxiety differently with English first language speakers commonly examined in communication research.

The purpose of this study was to assess EFL graduate students' level of anxiety and investigate their sources of anxiety on academic oral presentations. More specifically, this study attempted to explore the factors that are likely to contribute to the levels of anxiety during the course of preparation and performance for academic oral presentations. It was also hoped that through participating this study, the students would have an opportunity to reflect and self-evaluate on their learning process.

LITERATURE REVIEW

Oral presentation anxiety of EFL speakers relates to two types of anxiety – public speaking anxiety and foreign language classroom anxiety. Public speaking anxiety is a type of context-based communication anxiety that is relatively enduring and is only aroused in one type of context. Oral presentation is a form of public speaking because the both the speakers face an audience. However, research on public speaking anxiety typically focused on first language speakers. For foreign language speakers, it is like that they experience anxiety related to language proficiency, as in Morita's (2000) study. Therefore, research on foreign language classroom anxiety may also enrich our understanding of oral presentation anxiety especially with issues regarding to utilising a foreign language to communicate. Foreign language classroom anxiety was first defined by Horwitz, Horwitz and Cope (1986) as a type of state anxiety or specific anxiety because it is limited to language learning situations.

Based on previous research (Deffenbacher, 1977; Deffenbacher, 1978; Deffenbacher & Hazaleus, 1985), three components were found to predict anxiety. These components are worry, emotionality and task-generated interference. Worry refers to cognitive concerns about performance, consequences of failure, and self-perception of ability relate to others or confidence. Emotionality was defined as self-perception of physiological arousal such as increasing heart rate and sweaty palm (as cited in Strelau, 1992). Task-generated interference was derived from drive theories of anxiety (Spielberger, 1966, as cited in Deffenbacher, 1978). Task-generated interference suggests that under high anxiety, individuals are more likely to be distracted from task parameters that are irrelevant. An example of task generated

interference may be preoccupation with an unsolved problem. Hence, this study takes on a three-component conceptualisation of anxiety and views anxiety as a cognitive and physiological construct.

Anxiety has been found to correlate negatively with performance variables such as language classroom oral performance (Cheng, Horwitz & Schallert, 1999; Woodrow, 2006), grades (Aida, 1994; Saito & Samimy, 1996) and efficiency with tasks (Deffenbacher, 1978). These suggest that anxiety may have a negative effect on language performance.

Such a negative relationship may be due to the negative effect of anxiety on one's cognition. It has been found that anxiety correlates negatively with cognitive variables such as analogical reasoning (Leon & Revelle, 1985), retrieval of L2 vocabulary (MacIntyre & Gardner, 1994), generation of elaborate argument (Zohar, Livne & Fine, 2003), working memory capacity (Darke, 1988) and contributes to more negative cognitions such as self-evaluations and off-task thoughts (Zats & Chassin, 1985). Moreover, among the three components of anxiety, the cognitive components, worry and task-generated interference were found to predict performance better. Worry was found to be the primary contributor to anxiety, compared to emotionality (Deffenbacher, 1977; Deffenbacher, 1978; Deffenbacher and Hazaleus, 1985) and that the levels of task-generated interference were not significantly different from those of worry (Deffenbacher, 1978; Deffenbacher & Hazaleus, 1985).

These findings suggest that cognition plays an important role in affecting one's anxiety. Therefore, to understand anxiety, one has to examine the perceptions that each individual hold towards an experience. This study adopts the Perun's (1992) expectancy value theory of anxiety (EVTA) as its theoretical perspective. This theory postulates that anxiety arises as a result of expectancies and values and that anxiety disrupts performance by impairing the cognitive processing systems. The main assumption of the expectancy value theory of anxiety (EVTA) is that one's expectations about and subjective values attached to a task are the central determinants of one's anxiety. More specifically, anxiety involves a process of appraisal – evaluating the threatening situation, the actions needed, the feasibility and effectiveness of the action, the importance of the task and the importance of the consequences of failing the task. Therefore, by examining the participants' expectancies and values about the oral presentation experiences, it is hoped that we may understand the meanings attached and hence their sources of anxiety.

RESEARCH QUESTIONS

There are two research questions addressed in the present study:

1. How anxious are students while performing academic oral presentations?
2. What psychosocial factors play a role on students' anxiety during oral presentations?

METHODOLOGY

Setting and Participant

This study was conducted from the end of September 2008 to the beginning of December 2008, in a university in southern Taiwan. The participants were eighteen TEFL master's students who took the English Testing and Statistics course at the time of the study. Six of the participants were male, twelve female. Six of them were second year students and twelve of them first year students. The average age was twenty-four. Sixteen of the participants are native Mandarin speakers and have been educated in Taiwan. Two of the participants had received education in English-speaking countries with one speaks English as his L1, Chinese an L2, and the other speaks Chinese as his L1 and English an L2. The majority of the participants had majored in English or in English-related subjects.

Procedures

Informed consent was obtained before the data collection procedures of the study. There

were two main stages in the data collection procedure. During the first stage, the researcher videotaped participants' first oral presentation, which yielded eighteen presentations in total. Each participant was then invited to review, reflect and discuss on his/her own presentation with the researcher. At the second stage, the participants first filled out a questionnaire which inquired about participants' English learning background and assesses their levels of anxiety using the Personal Report of Public Speaking Anxiety (PRPSA: McCroskey, 1970, as cited in Richmond & McCroskey, 1995) (See Appendix); then, they were videotaped and interviewed for their experiences with their section presentation. A summary of the data collected is in Table 1.

Table 1. *Summary of Data*

Methods	Data collection schedule	Data collected
Classroom observation	18 participants × 2 = 36 recordings (3 × 10 weeks = 30 hours)	36 six filed notes
Interview	18 participants × 2 = 36 interviews	36 interviews (17 hours)
Questionnaire	18 participants	18 questionnaires

Data Analyses

The findings of this study were generated by triangulating multiple sources of data. Both qualitative and quantitative analyses were undertaken for the purpose of this study. More specifically, data obtained from classroom observation and interviews were analysed by the method of content analysis; data gathered from the questionnaire were analysed statistically. The PRPSA scores were analysed according to its categorisation and corresponding interpretations of the instrument. The interpretation scheme is presented in Table 2

Table 2. *Interpretation of PRPSA Scores*

PRPSA Scores	Category	Interpretation
34-84	Low	Very few public speaking situations would produce anxiety.
85-92	Moderately Low	While some public speaking situations would produce anxiety, most situations would not arouse anxiety.
93-110	Moderate	Exhibit moderate level of anxiety in most public speaking situations, but the level of anxiety is not likely to so severe that the individual could not cope with it.
111-119	Moderately High	Tend to avoid public speaking because it usually arouses a fairly high level of anxiety.
120-170	Very High	Exhibit very high anxiety in most public speaking situations and are likely to go to considerable lengths to avoid them.

(Richmond & McCroskey, 1995)

FINDINGS

In this section, the preliminary results of the study will be presented according to the two

research questions.

How anxious are students while performing academic oral presentations?

Table 3. *A Comparison of PRPSA Scores between the Participants in the current study and U.S. College Students*

Category		Current Study		U.S. College Students
Description	Scores	Number of People	Percentage (%)	Percentage (%)
Low	34-84	3	17	5
Moderately low	85-92	2	11	5
Moderate	93-110	6	33	20
Moderately High	111-119	4	22	30
Very High	120-170	3	17	40
Total		18	100	100

(Richmond & McCroskey, p.45)

The results of the PRPSA scores are shown in Table 3. The average PRPSA score was 103.6, which is within the moderate range (Table 2). This is also the range where most people (33%) fell in among the five categories of the scale (Table 3). This suggests that a majority of the class did experience some anxiety during oral presentations, and will experience anxiety in most public speaking situations, but the anxiety is not so severe that the person could not cope with it.

However, the analysis of the interview data has demonstrated a slightly different pattern. Table 4 shows a comparison on levels of anxiety between PRPSA scale and verbal report. Three categories were used to describe the degree and frequency of anxiety reported in the interview. People who were categorised as having a *low anxiety* reported low levels of anxiety and anxious episodes only occurred in a few occasions and lasted for a short amount of time. They typically remained calm throughout the oral presentation activity. People who were said to have a *medium* level of anxiety when they reported stronger level of anxiety and reported more anxious episodes and worrisome thoughts. Participants classified as having a *high* level of anxiety when they reported being highly anxious and expressed more worrisome thoughts.

Table 4. *A Comparison on Levels of Anxiety between PRPSA and Verbal report*

PRPSA			Interview	
Category	Percentage (%)		Category	Percentage (%)
Low	17	28	Low	33
Moderately Low	11			
Moderate	33		Medium	50
Moderately High	22	39	High	17
Very High	17			

By comparing the categories between those of PRPSA and those established for the interview data, the distribution of participants was different in these two sources of data. Looking at Table 4, it is noticed that in the PRPSA, most participating students fell within the

moderately-high and very-high groups, generating a combined percentage of 39%, with the moderate group comes second (33%) and finally the low and moderately low anxiety group (28%). The distribution is different in interview data. For the interview data, the high level group decreased (17%) and both the medium and low level groups increased (50% and 33% respectively). A close examination between PRPSA scores and the interview data revealed that out of the seven participants who had a discrepancy between the PRPSA score and interview category, six showed more anxiety on PRPSA scale than had revealed in the interviews. More specifically, among the six participants, four reported experiencing medium level of anxiety although they had a PRPSA score of moderately high or above; and two participants who had a moderate level of anxiety only reported experiencing little stress. An explanation for this could be the human tendency to avoid revealing personal weaknesses. That is, interviewees might have the tendency to under-report their true feelings.

However, one participant did not show the same tendency and reported more anxious feelings than shown on the PRPSA score. This may be due to the fact that PRPSA scale puts different emphasis on different aspects of anxiety compared to the ones that the participant actually experienced. An analysis of the PRPSA scale suggests that two were proposed to measure task-generated inference, nine measured worry and twenty-three measured emotionality. Base on the participant's verbal account, the participant was stressed because of the time limit, a task-generated interference as the participant reported,

“Before I went up (the stage), I felt a bit anxious, it was because I was the last presenter and there was only forty-five minutes left and I had about forty slides to go through. I also noticed that my classmates were already tired, and I was pressurised because of the time and amount of stuff to go through. I felt a bit dizzy before I went up...” (Interview, Nov 19, 2008).

Moreover, this participant has not shown anxiety over her ability to perform, her personal confidence as a public speaker did not suffer. Therefore, the PRPSA, a scale designed to measure *personal confidence* as a speaker, is likely to generate a low score for this individual whereas the interview data has revealed a source of anxiety undetected by PRPSA.

Comparing the PRPSA scores obtained for the current study to the U.S. college students, the current group of students exhibited lower levels of anxiety than the general population. On Table 3, we can see the distribution of U.S. college students' public speaking anxiety on PRPSA scale (Richmond & McCroskey, p.45). Several thousand U.S. college students took the PRPSA and it was found that a seventy percent of the students experienced at least a moderately high level of anxiety for public speaking. However, only thirty-nine percent of the students in the present study showed the same level of anxiety for public speaking. Therefore, it may be said that it is normal to experience some anxiety for public speaking and that the participating students were more at ease than 'normal'.

What psychosocial factors play a role on students' anxiety during oral presentations?

From the first interview with each of the participants regarding the sources of anxiety during oral presentations, ten major sources of anxiety were found. These ten sources were included in the questionnaire where the participants had to choose five which affect them the most in the second interview. An 'others' option with a space was provided for sources of anxiety not listed, if any. The most frequently chosen source of anxiety was classmates' reactions and behaviours, with accuracy of pronunciation as the second, and oral proficiency, personality, and audience familiarity as the third. Table 5 shows the sources of anxiety ranked from the most chosen to the least.

Classmates' Reactions and Behaviours

This was the most chosen factor and hence the most prevalent among participants. Base on participants' verbal account, twelve participants reported that they felt anxious when the

audience looked confused, showed no response to his/her oral presentation and other forms of disregard. Here are some common reactions:

S08: If I saw my classmates talking or laughing, I would wonder if I had done something and they were laughing at me. I would think I must have said something wrong and they are laughing at me...I would feel uneasy (Interview, Dec 1, 2008)

S09: If my audience looked confused I would be really nervous, it felt like I have not explained clearly (Interview, Dec 1, 2008).

S13: I wondered if some classmates were in a bad mood or not... they looked as if they were annoyed. So I thought to myself...not to look at them...because I would be more relieved if I don't look at them...but I wasn't sure if that's because I have some problems with my oral proficiency (Interview, Dec 10, 2008).

It can be seen that the audience reactions may contribute to anxiety by raising presenter's self-doubt about the quality of explanation in the topic and oral proficiency.

Table 5. *Ranking of Sources of Presentation Anxiety*

Rank	Source of Anxiety	Frequency	Percentage
1	My classmates' reactions and behaviours during my oral presentation.	15	83.3%
2	Accuracy of my pronunciation	14	77.8%
3	Oral proficiency	11	61.1%
3	My personality	11	61.1%
3	Audience familiarity	11	61.1%
6	Professor's reactions and behaviours during my oral presentation.	10	55.6%
6	Number of rehearsal	10	55.6%
8	Classroom settings	4	22.2%
9	Others	2	11.1%
10	Handling of software	1	5.6%
10	Professor's personality	1	5.6%

Accuracy of Pronunciation

Although accuracy of pronunciation was the second highest chosen source of anxiety in the questionnaire and fourteen people chosen it as one of their top five, only six participants talked about worrying their pronunciation in the interviews. This suggests that accuracy of pronunciation may have caused some stress in most participants, however, there may well be other sources of anxiety that affect participants' emotion more. This lower level of anxiety for pronunciation is evident when participants talked about that they do not normally worry about their pronunciations prior to giving a presentation, but only discover that they are not sure how to pronounce some of the words on the slides and got anxious on stage. For example,

S05: ...because I didn't really rehearse, I just read over it (the slides), and I thought I knew how to say them, so sometimes I might not give accurate pronunciation (during the performance) (Interview, Nov 13, 2008).

Interviewer: Were you nervous when you struggled to pronounce the word?

S12: A little bit, but then I thought I would throw this problem back to the audience, so I won't feel that anxious... I mean... I care about my pronunciation whether if it's accurate, but I won't worry about it before the presentation, I tend to worry about it after the presentation (Interview, Dec 10, 2008).

Self-perceived Oral Proficiency

Self-perceived oral proficiency was the third most chosen source of anxiety. The participants were also asked to rate their own English oral proficiency as beginner, intermediate, high intermediate, or advanced. Statistical analysis of the data revealed that English oral proficiency and the PRPSA score were positively correlated ($r = .49, p = 0.039$). This suggests that the higher self-perceived oral proficiency, the less anxious the participant will be. However, this relationship is rather weak and students who perceived themselves as having higher self-perceived English oral proficiency does not always feel less anxious than the ones who perceived themselves as having lower English oral proficiency. This is evident in a student who rated his/her oral proficiency as advanced but obtained a PRPSA score in the moderately high level, which was higher than many of the other students who rated themselves as having lower English oral proficiency. This participant explained his/her anxiety during the interview,

S08: ...actually, the day before my presentation, or when I am about to go on stage, I always feel out of breath, although I have presented many times during the undergraduate courses and have also done some stage performances, but I just couldn't...(help feeling anxious) (Interview, Oct 15, 2008).

Self-perceived Personality

Many participants regarded personality as one of the important qualities of a good presenter. In answering question about the elements of a good presentation performance, besides good content and clear explanation, presenter's personality was also recognised as an important feature. Many participants believed that a lively and energetic personal style makes the presentation performance more attractive and interesting to listen, and that their own personality inhibits them from doing so.

S12: I've always felt that my presentation is kind of boring, and it's boring because of my personality...at least when compared to S06's (a classmate) presentation style. I hope I could present like he/she does, to create a relaxed atmosphere but at the same time to explain the content, but I think this has got something to do with one's personality (Interview, Dec 10, 2008).

S13: I think I'm not that active (personality), so my presentation is not attractive for the audience (Interview, Dec 10, 2008).

Audience Familiarity

All of the eleven participants who regarded audience familiarity as one of their top five sources of anxiety felt that an unfamiliar audience would increase their anxiety. They felt that if they make any mistakes, they would be less embarrassed in front of the people they know,

S01: I would not be so afraid of losing face in front of the people I know, and I won't worry that much for the audience's comment...if they say they could not understand what I'm talking about (presentation). I won't if I am familiar with the audience (Interview, Nov 20, 2008).

S17: If I present in front of people that I don't know, I would get very nervous (Interview, Dec 3, 2008).

Professor's Reactions and Behaviours vs. Professor's Personality

An interesting finding for the study is that although only one participant indicated *professor's personality* as a source of anxiety, over half of the participants (55.6%) indicated that the instructor's reactions and behaviours during the oral presentation was a source of anxiety for them. These two pieces of results suggest that although many participants perceive the instructor's personality as non-threatening, instructor's response and behaviours while listening to the oral presentation plays an important role in affecting their emotion, as many participants reported,

S11: She is a very nice teacher. I don't feel that much pressure for the class (Interview, Oct 15, 2008).

S11: I seldom looked at the teacher because I was scared to. It makes me more nervous if I look at her, but I think, if the teacher smiled, I would feel assured (Interview, Oct 15, 2008).

S17: She is very nice (personality), if I make a mistake (during the presentation), I think... she won't correct me right away. She would wait until I finish the section and elaborate for me. I cannot multi-task. I can only do a few things at a time. Because the teacher is very nice, I wasn't that anxious

for the presentation (Interview, Oct 29, 2008).

S17: ...every time the professor turned to look at the book I would think maybe I have said something wrong, but as I finished saying something I forgot what I have said...so I'm really afraid maybe I have made some mistakes, I would get nervous for her responses (Interview, Oct 29, 2008).

CONCLUSION

The findings of this study suggest that thirty-three percent of the participating students experienced a moderate level of anxiety and thirty-nine percent experiencing anxiety at higher levels. Therefore, this group of EFL speakers were comparatively more relaxed than the U.S. college students. The sources of anxiety may be considered as coming from two clusters, that of peers' evaluation and self-evaluation. Peers' response, the most chosen source of anxiety, plays an important role in affecting the participants' levels of anxiety during oral presentations. The participating students attached importance to how their peers regard of them, but such influence diminished when the students got to know their classmates. The second cluster of sources of anxiety came from evaluation of accuracy of pronunciation, oral proficiency, and personality. As can be seen, the relationship between level of anxiety and these factors were found to be less straightforward than the first cluster. An implication of the study is that a supportive and encouraging classroom culture, especially among the classmates is likely to alleviate presenter's presentation anxiety, and that the more proficient EFL/ESL speakers are not always immune to stage fright.

REFERENCES

- Aida, Y. (1994). Examination of Horwitz, Horwitz and Cope's construction of foreign language anxiety: The case of students of Japanese. *The Modern Language Journal*, 78, 155-168.
- Cheng, Y. S., Horwitz, E. K., & Schallert, D. L. (1999). Language anxiety: Differentiating writing and speaking components. *Language Learning*, 49(3), 417-446.
- Darke, S. (1988). Anxiety and working memory capacity. *Cognition and Emotion*, 2(2), 145-154.
- Deffenbacher, J. L. (1977). Relationship of worry and emotionality to performance on the Miller Analogies Test. *Journal of Educational Psychology*, 69, 191-195.
- Deffenbacher, J. L. (1978). Worry, emotionality, and task-generated interference in test anxiety: An empirical test of attentional theory. *Journal of Educational Psychology*, 70, 248-254.
- Deffenbacher, J. L. & Hazaleus, S. L. (1985). Cognitive, emotional, and physiological components of test anxiety. *Cognitive Therapy and Research*, 9(2), 169-180.
- Hilleson, M. (1996). "I want to talk with them, but I don't want them to hear": An introspective study of second language anxiety in an English-medium school. In K.M. Bailey & D. Nunan (Eds.), *Voices from the language classroom: Qualitative research in second language education*. Cambridge: Cambridge University Press.
- Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language anxiety. *The Modern Language Journal*, 70(2), 125-132.
- Jackson, J., & Latane, B. (1981). All alone in front of all those people: Stage fright as a function of number and type of co-performers and audience. *Journal of Personality and Social Psychology*, 40(1), 73-85.
- Kim, S. (2006). Academic oral communication needs of East Asian international graduate students in non-science and non-engineering fields. *English for Specific Purposes*, 25, 479-489.
- Leon, M. R. & Revelle, W. (1985). Effects of anxiety on analogical reasoning: A test of three theoretical models. *Journal of Personality and Social Psychology*, 49(5), 1302-1315.
- MacIntyre, P. D. & Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive

- processing in the second language. *Language Learning*, 44(2), 283-305.
- Morita, N. (2000). Discourse socialization through oral classroom activities in a TESL graduate program. *TESOL Quarterly*, 34, 279-310.
- Richmond, V. P. & McCroskey, J. C. (1995). *Communication: Apprehension, avoidance, and effectiveness* (4th ed.). Arizona: Gorsuch Scarisbrick Publishers.
- Saito, Y. & Samimy, K. K. (1996). Foreign language anxiety and language performance: A study of learner anxiety in beginning, intermediate, and advanced-level college students of Japanese. *Foreign Language Annals*, 29(2), 239-251.
- Strelau, J. (1992). Introduction: Current studies on anxiety from the perspective of research conducted during the last three decades. In D. G. Forgays, T. Sonsnowski, & K. Wrzesniewski (Ed.). (1992). *Anxiety: Recent development in cognitive, psychophysiological and health research* (pp.1-10). Washington: Hemisphere Publishing Corp.
- Woodrow, L. (2006). Anxiety and speaking English as a second language. *RELC*, 37(3), 308-328.
- Young, D. J. (1990). An investigation of students' perspectives on anxiety and speaking. *Foreign Language Annals*, 23(6), 539-553.
- Zatz, S. & Chassin, L. (1985). Cognitions of test-anxious children under naturalistic test-taking conditions. *Journal of Consulting and Clinical Psychology*, 53(3), 393-401.
- Zohar, D., Livne, Y., & Fine, J. (2003). The effect of anxiety on linguistic parameters of public speech: A verbal impairment model. *Anxiety, Stress, and Coping*, 16(3), 293-306.
- 國立大學首創清大電機系必修課全說英文 (2004 年 4 月 14 日)。中時電子報 (China times)。 <http://www.ieeuc.com.tw/news6-1-6.htm>
- 林琳 (Lin, L.) (2007 年 4 月 12 日)。《研究顯示英語課程成為全球高等教育趨勢》。中央社。 <http://www.english.com.tw/modules/news/article.php?storyid=2521>

APPENDIX – Personal Report of Public Speaking Anxiety (PRPSA)

This instrument is composed of thirty-four statements concerning feelings about oral presentations. Please reflect back to your *second oral presentation* and tick what most accurately describes you (SA: strongly agree; A: agree; N: are neutral; D: disagree; or SD: strongly disagree). Work quickly; record your first impression.

Item	Description	SA	A	N	D	SD
1.	While preparing the oral presentation, I felt tense and nervous					
2.	I felt tense when I saw the words <i>oral presentation</i> on the course outline.					
3.	My thoughts became confused and jumbled when I was giving the oral presentation.					
4.	Right after giving the oral presentation I felt that I had had a pleasant experience.					
5.	I got anxious when I thought about the oral presentation coming up.					
6.	I had no fear of giving the oral presentation.					
7.	Although I was nervous just before starting the oral presentation, I soon settled down after starting and felt clam and comfortable.					
8.	I looked forward to giving the oral presentation.					
9.	When the professor announced there would be oral presentation activities for the course, I felt myself getting tense.					
10.	My hands trembled when I was giving the oral presentation.					
11.	I felt relaxed while giving the oral presentation.					
12.	I enjoyed preparing for the oral presentation.					
13.	I was in constant fear of forgetting what I had prepared to say.					
14.	I would get anxious if someone asked me something about my topic that I did not know.					
15.	I faced the prospect of giving the oral presentation with confidence.					
16.	I felt that I was in complete possession of myself while giving the oral presentation.					
Item	Description	SA	A	N	D	SD
17.	My mind was clear when giving the oral presentation.					
18.	I did not dread giving the oral presentation.					
19.	I perspired just before starting the oral presentation.					
20.	My heart was beating very fast just as I started the oral presentation.					
21.	I experienced considerable anxiety while sitting in the room just before my oral presentation started.					
22.	Certain parts of my body felt very tense and rigid while giving the oral presentation.					

23.	Realizing that only a little time remained in the oral presentation made me very tense and anxious.					
24.	While giving the oral presentation I knew I could control my feelings of tension and stress.					
25.	I breathed faster just before starting the oral presentation.					
26.	I felt comfortable and relaxed in the hour or so just before giving the oral presentation.					
27.	I would do poorer on the oral presentations if I was anxious.					
28.	I felt anxious when the teacher announced the dates for oral presentations.					
29.	When I made a mistake while giving the oral presentation, I found it hard to concentrate on the parts that follow.					
30.	During the oral presentation, I experienced a feeling of helplessness building up inside me.					
31.	I had trouble falling asleep the night before the oral presentation.					
32.	My heart was beating very fast while I was presenting.					
33.	I felt anxious while waiting to give my oral presentation.					
34.	While giving the oral presentation, I got so nervous that I forgot the facts I really knew.					

Vocabulary Games as a Memory Enhancement Device

Margaret Chen (陳美華)

Wenzao Ursuline College of Languages
meihua@mail.wtuc.edu.tw

This paper reports a follow-up study to Chen's research (2009) which investigated vocabulary gains in a 20-hour vocabulary learning program using specially designed games in the form of Poker Cards, Chinese Chess and Gobang invented by the author (2008). Forty-six students from a language college in southern Taiwan participated in that study in the summer of 2008. Their English competence level was equivalent to A1-A2 of the Common European Framework Reference (CEFR). A 174-item vocabulary test written by the author was administered each time after a set of games was completed in addition to a pretest. The tested items were the words retrieved from the vocabulary games. Totally, the participants took four tests throughout the program. The result indicated significant gains after completing the programs. To investigate further if the designed games would help learners retain the learned vocabulary during the program, another test was conducted 45 days later which forms the main discussion of this paper. All 46 participants returned for the test. The results showed that vocabulary retention, compared to the first three vocabulary tests, was significant. That means the participants managed to maintain the vocabulary size they had gained during the learning program. Nevertheless, if compared to the result of the 4th test, the last test taken when the program had been completed, the scores were significantly different. That means that participants, though they did not forget all they had learned, somehow forgot some words after a 45 days' break. The research concludes that learning vocabulary through games contributes to the increase of vocabulary size, acceptable retention rate, and different approaches to learning vocabulary among EFL learners.

INTRODUCTION

The present study is a follow-up study to Chen's research (2009) in which the author reported her findings in vocabulary gains through a vocabulary learning program by using a series of designed vocabulary games. To prove if the designed games could facilitate learners' long-term memory regarding the words learned from the previous experiment, the author invited the original 46 participants to take the same test 45 days after the program had been completed. Therefore, the purpose of this paper is to report the result of the designed games in relation to their effectiveness in learning vocabulary and its retention.

Over the past few decades, prolific researches have revealed the fact that learning vocabulary is essential to language acquisition (Horwitz, 1988). Vocabulary knowledge, referring to "knowing a word" and vocabulary size, referring to the size of the words known, are crucial assets to a goal of effective language manipulation. Nation (2006) estimated that generally speaking, an ESL/EFL learner needed an 8,000 to 9,000 word-family vocabulary for unassisted reading comprehension and a 6,000 to 7,000 word-family vocabulary for comprehending spoken texts. In addition, a group of researchers have shown evidence that vocabulary size and knowledge are related to reading achievement (Francis & Simpson, 2003; Qian, 2002) and learning outcome (Gu & Johnson, 1996; Kojic-Sabo, & Lightbown, 1999). Without doubt, any language teacher or educator would agree that if learners do not have enough vocabulary size and knowledge, not only does their reading and listening comprehension lack fluency, but the acquisition of other language skills would be thwarted. As a result of such concerns, studies about vocabulary learning strategies have been compiled and vocabulary strategy instruction programs implemented (Gu & Johnson, 1996; Marefat & Shirazi, 2003; On-Lai, 1993, Rasekh, & Ranjbar, 2003; Sanaoui, 1995; Siyanova & Schmitt, 2007; Wesche, & Paribakht, 2000).

For instance, Schmitt and Schmitt (1993) conducted a research to identify the

strategies used by Japanese EFL students when learning vocabulary, and they also investigated their perception towards the usefulness of the vocabulary strategies used. Fourteen methods to learn new words and 28 strategies to memorize and retain words learned were reported by the 600 participants in this study. As for the perception of usefulness, participants favored writing and verbal repetition; usage of word lists and flash cards were deemed the least useful strategies.

Sanaoui (1995) did a similar study to classify vocabulary learning strategies among 50 ESL adults, and she also conducted four case studies to further investigate the strategies used. Her findings indicated that there were two approaches used by adult learners: structured and unstructured. She concluded that both beginning and advanced learners applied repetition, imagination and association as mnemonic strategies. Furthermore, Gu and Johnson (1996) probed into the relationship between the use of vocabulary learning strategies and learning outcomes among 850 non-English major university students in China. The result indicated that learners' self-initiation, elective attention, oral repetition and cognitive strategy use correlated with their vocabulary size and language proficiency. However, memory strategies such as imagery, visual and auditory associations were found correlated with vocabulary size only. Noting that a discrepancy existed between ESL and EFL learners, Kojic-Sabo and Lightbrown (1999) explored vocabulary learning strategies used in different language learning settings and found similarities and discrepancies among 47 ESL and 43 EFL learners. The similar learning characteristics the participants shared were: one of frequent use of note-taking strategies and the dictionary. On the other hand, disparity was found in that EFL learners reviewed what was learned more often than ESL learners. The above mentioned studies indicated the fact that for EFL learners, repetition is a commonly used practice when EFL learners learn to write and speak.

Attempts to reinforce learners' short-term and long-term memory have been practised by several Taiwanese researchers in forms of strategy instruction and programs. (Y. M. Chen, 2006; Chu, 2007; Chien, 2006; Dunbar, 2004; Hsu, 2004; Huang, 2004; Yek, 2006). Y. M. Chen investigated young learners' use of memory strategies by implementing keyword methods with mnemonic flash cards. Sixty-nine participants aged 11 were divided into three groups, and received keyword-received method, keyword-generated method and rote repetition method. The result indicated that students who were given keyword-received method performed better than students from the other groups in a one-week delayed test, two-week delayed test and a three-month delayed test. By the same token, Chu (2007) conducted memory strategy instruction in elementary students' vocabulary learning. One hundred and twenty six fourth-graders were divided into experimental groups and control groups. The result showed that the experimental group who received memory strategy training outperformed the control group in spelling accuracy and use of vocabulary learning strategies.

On the other hand, strategies such as imagery and visual repetition were mostly applied to students of higher levels. Huang (2004) applied drawing activities in vocabulary learning to the experimental group and traditional method to the control group. From the in-depth interview, participants in the experimental group reflected positive attitudes towards drawing to memorize vocabulary and claimed that the strategy helped with their recall of vocabulary. Hsu (2004) implemented an iconic-morphological approach to teach commonly-used roots in a senior high school. Ninety-one students underwent an iconic-morphological approach, a non-iconic-morphological approach and a traditional-definition-based method respectively, and the result showed that students who received the iconic-morphological approach performed better in awareness of morphology, short-term and long-term memory, as well as spelling ability than the students in the other groups.

Researches done in Taiwan have shown that the vocabulary learning outcome is

affected to a large extent by implementation of vocabulary learning strategies either in regular classroom teaching or activities outside classrooms. Despite the consensus on the need of vocabulary learning training, studies focusing on improving learners' gains in vocabulary size and memory strategy through specifically designed games were limited.

Vocabulary learning to most EFL students is a painful experience. Many researches attempt to find the best ways to help learners although results very often are found controversial. For example, some studies contended that learning words in context enhance reading ability, while other studies maintained that learning through reading word lists or isolated from context can also be transferred effectively to long term memory (Martin-Chang, Levy & O'Neil, 2006). Regardless of the dispute, if learning takes place under circumstances of a relaxing, pleasant and challenging atmosphere, it would be ideal for both teachers and students. Such a notion led to the coinage of the word "edutainment" in the early 1990s, and it has become a popular term in educational sectors (Hadfield, 1990; Lewis & Bedson, 1999). As the word suggests, the word "edutainment" comprises "education" and "entertainment." An "edutainment" activity should achieve the purposes of "educating" and "entertaining". This belief explains why educators emphasized that educational games have to have the elements of fun, pleasure, competition challenges, and goals (Crookall, Oxford & Saunders, 1987; Hadfield, 1990; Lewis & Bedson, 1999; Shie, 2004; Thornton & Cleveland, 1990).

Reviewing popular language games in EFL/ESL classrooms, repetition was found most frequently in games such as guessing games, memory games, drills, reconstruction, simulation...etc. As some previous studies suggested, the repetitiveness enables learners to practice the target language without a high-level of cognitive skills, and assists them to build competency, speed, and accuracy in language (Newby, Stepich, Lehman & Russell, 2000). Apart from repetition, if a game provides users adequate time to repeat the content of the game, their anxiety level will be alleviated (Macedonia, 2005).

In conclusion from the above-mentioned literature, though some methods proved to be effective, some proved to be controversial; learning vocabulary through games instead of through contextual reading or other devices seemed rare. The author has thus strived to incorporate vocabulary learning strategies into educational games in order to facilitate the EFL/ESL vocabulary learning process.

PURPOSE OF STUDY

The paper aims at assessing the participants' vocabulary gains and retention rate by employing a 20-hour-vocabulary-game learning program. The study was designed to find answers to the following questions:

1. Will the designed games help the participants gain vocabulary and retain their gain to a significant level after completion of the program?
2. Which type of game will help the participants gain vocabulary to a significant level?
3. Which proficiency group will gain and retain more vocabulary in this experiment?

LIMITATION

There were 46 participants involved in this study. They were categorized into three proficiency groups based on their English proficiency average scores and standard deviation. The numbers of participants in the lowest (1st Group) and highest groups (3rd Group) are relatively low in numbers thus the distribution of the groups is unequal. Also, due to the small number of participants the generalizability of the findings has to be interpreted carefully. Second, the retention test was scheduled at an interval of 45 days after the programs which were participated in during this study, and the test was taken in the first week and the 5th

week of the new semester after the summer vacation in 2008. Although the games and word list were not offered to the participants for review during the 45-day-break, they might have been influenced by other factors such as schoolwork, language classes, self-study...etc. What the participants did during the interval was unknown to the researcher. Lastly, the study mainly focuses on participants' retention rate after the program; therefore, the possible factors attributed to their performance were not discussed in this study.

THE STUDY

This study follows up the recommendation of Chen (2009), which suggested a retention test be assessed to validate if the invented games would facilitate the learners' long-term memory in vocabulary. The findings of the learning program indicated that a significant vocabulary gain was found among the 46 participants; therefore, a retention test, with an interval of 45 days after the program was scheduled. The methodology of this study used the vocabulary learning program as the framework and only discussed the participants' performance on vocabulary gains. Detailed information about the vocabulary learning program can be found in Chen's study (2009).

Participants

Forty-six students from a language college in Kaohsiung, Taiwan voluntarily registered for the vocabulary learning program in the summer, 2008. The participants' ages ranged from 17 to 31, and their average English proficiency score was 143.65 based on the test result of the CSEPT (College Students English Proficiency Test), administered in September, 2007, by the Language Training and Testing Center (LTCC) in Taipei, Taiwan. The participants' proficiency level was equivalent to the level of CEFR A1 (Common European Framework Reference). During the training program, the 46 participants (15 from the July program and 31 from the August program) were divided into groups based on the nature of the games and their proficiency scores.

Four staff members from the Language Diagnostic and Consulting Center (LDCC) where the activity took place were recruited to act as game dealers. All the four dealers were female holding a bachelor degree in English or Foreign Language Instruction. Three other student tutors recruited for conducting this task also majored in English or Foreign Language Instruction. One college student from another university in Southern Taiwan who was doing an internship in the LDCC also received training to be a game dealer.

Instrument

The instrument applied in this study involved four major parts. One was the designed vocabulary games per se, and the other was the learning program implemented to evaluate the effectiveness of the designed vocabulary games. Furthermore, a vocabulary test and a questionnaire were used in the vocabulary learning program.

The Vocabulary Games

This section briefly introduces the designed vocabulary games used in the learning program. Detailed descriptions such as design rationale and learning program have been included in Chen's study (2009).

The content of the vocabulary games used in the learning program was noun-noun compound words, and game rules were adapted from a regional game rule "Pickred" using Poker Cards, a popular game rule "An-qi" using Chinese Chess, and the typical game rule in Gobang. The rules of the above games were adapted in order to simplify their complications and help the players concentrate more on words rather than regulations. Coupled with the advantages of their popularity, the games can be adjusted easily according to the needs of the players and learning situations.

Poker Cards vocabulary game. Two sets of Poker Cards were used in the learning program. Each set of Poker Cards consists of 54 cards and can be played like a regular deck

of cards. Designed with one noun and an illustration demonstrating the meaning, the card also included the Chinese translation for reference in small print. The basic principle is to match noun with noun to form another compound noun.

As for the content of the card game, the first set of Poker Cards contained 42 basic words and second set used 45 words. By matching two cards together, the first set of cards can generate 163 compound nouns while the second can produce 148 compound nouns. In addition, some words were repeated in the two sets in order to reinforce learners' memory. The two sets can be mixed together, providing the possibility of 311 combinations, within which 61 compound nouns were repeated.

The game flows in the learning program follow the procedure below (*Figure 5*). In this study, the Poker game was played by three to four students. A dealer at each table helped hand out cards to each player and watched for correct matches. Whoever played first, started matching either from his/her hand or from the desk by saying the word out loud. The player needed to place the matched cards aside if there were any, and draw one card from the deck in the middle in order to match again on the same turn. During the game process, players were required to read out the combinations they have made so that all players can learn together. If no match can be formed, the player needed to throw down one unwanted card. The game proceeded as above until there was no card to be drawn and none of the players had a pair in hand. The winner was either the one who had matched the most pairs or who had completed the game first within the set time. The word list was provided after each round which participants could use for preview or review, or during the game round; dealers employed the word list when they needed to verify players' associations.

Figures below present real examples of the Poker Cards vocabulary games.



Figure 1. Invented Poker Cards (Set 1)



Figure 2. Invented Poker Cards (Set 2)

Chinese Chess vocabulary game (Xiangqi). The second type of game applied in the learning program is a variation of a Chinese Chess game (*Figure 3*) called Blind Chess (An-qi). Four sets of Chinese Chess vocabulary game were used in the program; each differs in its level of difficulty. There are eight red words and 24 black words in each set. Totally, 107 words were applied to compose the four sets of Chinese Chess in this learning program. Similar to the Poker Cards game design, each piece of Chinese Chess has one noun engraved on it. Players use the red pieces to capture the black pieces. Each red piece can be matched

with at least three black ones to form another three compound words. Approximately 34 to 58 compound nouns can be generated and if combined, 161 unrepeated words can be produced.

The participants spent about one and half hours to play three rounds for each set of Chinese Chess. Totally, the participants played 12 rounds for about six hours.



Figure 3. Invented Chinese Chess game

Gobang vocabulary game. One set of Gobang was used in this study. It is considered highly competitive in terms of requiring players' intense concentration, stable mentality, and sharp observation. The Chinese Gobang (*Figure 4*) invented in this vocabulary game comprised 60 noun words. In this research, it was played by two persons and each one had the same 60 words in hand but in two different colors. By matching the same 60 words together, a minimum of 222 compound nouns could be created.

The game rule is adapted from the traditional game called, "Gobang". The game is played by two sides. One holds black pieces, the other holds red pieces. Whoever plays first, places one piece in the middle of the board. The other player then tries to match the piece from any direction with one of the nearby pieces in order to make a compound noun. The pieces can be lined vertically, horizontally and diagonally. The first player who lines up five pieces with the same color wins. Sometimes, victory can be blocked if the matching is interrupted by the opponent who is able to match the word with his/her piece. In short, the players not only need to know the possibility of compound words to be matched by using their own pieces, but must pay attention to winning and defeating the opponent.

Before playing Gobang, the participants had already played Poker Cards and Chinese Chess for approximately 5 to 6 hours each. Among the 222 compound nouns in the Gobang, 120 compounds appeared in the previous games. Therefore, participants were assumed to have 54.1% of possibility to match pairs based on the previous known words. That means playing Gobang, also offered the participants chances to review learned vocabulary in previous games while continuing to learn further matched words in this new type of game.



Figure 4. Invented Gobang

Word counts in the games. The detailed word count and the repetition in each game are shown in *Table 1*. Based on the total number of the original words and compound nouns

in the Gobang, the matched compound words appearing in the Gobang were defined to be “repeated” in this study. For instance, among the 163 compound nouns which could be matched from the Poker Card set 1, only 91 compound nouns did not appear in the Gobang. It also means 72 compounds appeared in the Gobang. After playing all the games in the program, players were expected to learn 454 compound nouns excluding the matched pairs in different types of games. Most of the words chosen are between the beginning and lower intermediate levels, however, the combination would advance one to two levels.

Table 1 *The Word Counts and Repeating Words in Each Game*

Game type	Coding	Original word count	Unrepeated original word count	Compound noun count	Unrepeated compound noun count
Poker Cards	P1	42	18	163	91
	P2	45	19	148	64
Chinese Chess	C1*	31	7	53	64
	C2	31	4	58	7
	C3	31	14	38	16
	C4	32	18	34	26
	C5	32	20	33	28
Gobang	G	60	60	222	222
Total**		273	153	696	454
Total compound nouns introduced			607 words after these games		

Note: 1. *=C1 was not used in this study

2. **=The total did not include statistics in C1.

Vocabulary Learning Program

The participants did a vocabulary pretest which comprised 174 items at the outset of the program. Course syllabus and regulations, as well as rules were explained on the first day.

Participants followed the schedule of the program; each participant had about five to six hours to play at least three to four rounds following the order of two sets of Poker Cards, four sets of Chinese Chess and one set of Gobang. After each round, participants were given ten minutes to review the word list, and after completing one type of game, the participants took the vocabulary test. Totally, the participants wrote the vocabulary tests four times during the program. *Figure 5* illustrates the flow of the games in the learning program.

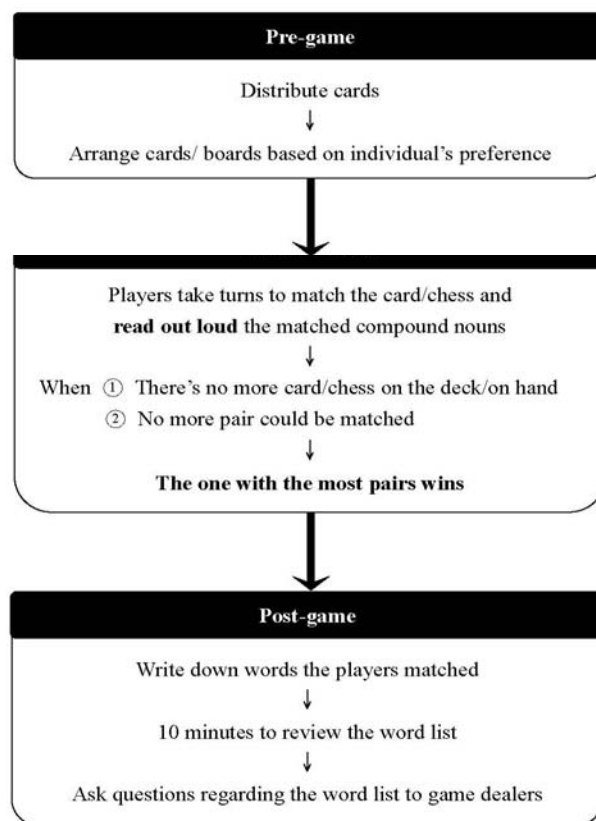


Figure 5. The procedure of the game in the program

Vocabulary Test

A 174-item Vocabulary Test designed by the author was used to assess the participants' vocabulary gains and their retention rate from the learning program. The 174 questions were written in Chinese and 348 English compound nouns were placed aside as choices for each question. The format of the test was adapted from Vocabulary Level Test developed by N. Schmitt, D. Schmitt and Clapham (2001). Participants chose the correct Chinese definition from the six English test items. All the words were taken from the three types of games. Among 348 English compound nouns, except for the 19 repeated test items, 187 words appeared from the two sets of Poker Cards, 76 words from the four sets of Chinese Chess and 66 from the Gobang set.

The test was administered four times during the program, namely pretest, post-test and twice on completion of sets of games. The order of the test items was altered once to prevent students from memorizing the answers. The same test was administered again after 45 days of the program, in order to assess learners' retention rate. The following example illustrates the format of the test, and the complete test can be obtained from the author upon request.

Table 2 *Excerpt of the Vocabulary Test*

1. lifetime 2. lifework 3. handline 4. website 5. workhand 6. worksite	<p>_____ 一生</p> <p>_____ 雇工</p> <p>_____ 畢生的工作</p>
1. horseback 2. horseman 3. horsepower 4. mankind 5. manpower 6. workhorse	<p>_____ 馬力</p> <p>_____ 人類</p> <p>_____ 吃苦耐勞的人</p>
1. doorkeeper 2. doorway 3. eyebath 4. sunbath 5. suntrap 6. trapdoor	<p>_____ 日光浴</p> <p>_____ 陷阱門</p> <p>_____ 守門人</p>

Questionnaire

A 24-item questionnaire with four open-ended questions was written by the author in both Chinese and English in order to help students understand the questions easily. It was a 6-scale Likert type questionnaire, ranging from strongly agree (SA), agree (A), undecided (U), disagree (D), to strongly disagree (SD), and “I’m not sure” (N). The questionnaire investigated the participants’ perceptions towards the following questions: the importance of memorizing vocabulary, feedback on learning new words through games, and players’ ways of increasing vocabulary size. Four open-ended questions were also included which aimed at finding out which types of games helped them learn more, and which they liked the most. Detailed analysis of the questionnaire can be found in Chen (2009).

Data Collection

Before the 45-day delayed test was administered, the previous four tests taken at each stage of the game were completed. The participants were informed about the retention test on the last day of the program. Participants in the July program took the retention test from September 8th to the 15th, and participants in the August program took the test from October 6th to the 13th. They could come to take the test anytime they wanted during the allotted time. Unlike the previous four tests which were administered together, this time the participants came during the given period of time due to the difficulties in scheduling a fixed time among the 46 participants.

All the five test scores as well as participants’ answers were computed and analyzed.

Data Analysis

The Statistics Package for Social Science (SPSS) was applied to calculate the mean scores and the standard deviations of the vocabulary tests. All the participants were grouped according to their proficiency and the group test score was further analyzed in a *one-way ANOVA* test. A *paired-sample-t-test* was used to investigate vocabulary gains between different tests.

RESULTS AND DISCUSSION

Vocabulary Tests Results

In response to the first research question, “Will the designed games help the participants’ gain vocabulary and retain their gain to a significant level after completion of the program?,” the average of correct answers increased from 132 items (1st test) to 161 items (4th test) out of 174 items after the program. The standard deviation diminished between each test which seemed to show that the game helped decrease the level gap among participants. If compared to the first test (pretest) and the fourth test (test after Gobang), a significant difference was found in Pair IV ($p=.000***$), which means players had extended their vocabulary gains to the highest degree after the program. The t -value in Pair IV indicates the effectiveness of the learning program. From the *pair-sample-t-test* (table 3), the scores between each pair of tests is significant ($p=.000***$). It can be implied that the games succeeded in helping the participants gain vocabulary.

The second research question asks which type of games helps the participants gain vocabulary to a significant level. The result showed that significant difference was found each time after the game was completed and the participants tested. Comparing the first four tests, the participants learned most of their new words from playing the Poker Cards and they seemed to make least progress in playing Chinese Chess. However, players increased their vocabulary in playing Gobang more than they did after playing Chinese Chess. Comparatively speaking, the result seemed to indicate that Gobang was more effective in promoting learners’ vocabulary gains compared to Chinese Chess. But the gradual improvement in gains might result from repeating playing half of the same content though in different games, or from other factors which were unknown in this study. During this experiment, the order of the games was that Poker Cards were played first, Chinese Chess followed, and Gobang completed the program. The order of the game arrangement might bring out different results. Also, the finding showed that each set of games helped the participants gain a significant amount of vocabulary. As to whether it is necessary to use three types of game or employ one instead and whether the results be the same or not are worth further investigation. Furthermore, if playing one type of game is equally effective in comparison to playing three types, either within same length of time or within an even shorter time range, then the program can be readjusted to last less than 20 hours. However, the above questions require more study. In this report, the author was unable to provide a reason why the participants gained more vocabulary by playing Poker Cards and by playing Gobang.

Table 3 *Comparisons of Vocabulary Tests Results between Each Test*

	Pair I		Pair II		Pair III		Pair IV		Pair V		Pair VI		Pair VII	
	T1	T2	T2	T3	T3	T4	T1	T4	T4	T5	T1	T5	T3	T5
M	131.7	148.0	148.0	154.8	154.8	161.1	131.7	161.1	161.1	154.7	131.7	154.7	154.8	154.7
SD	11.6	9.6	9.6	8.4	8.4	7.6	11.6	7.6	7.6	8.5	11.6	8.5	8.4	8.5
t	-12.458		-9.181		-10.466		-21.850		9.158		-18.680		.067	
p	.000***		.000***		.000***		.000***		.000***		.000***		.947	

Note: 1. *= $p<.05$, **= $p<.01$, ***= $p<.001$

2. N= 46

3. T1=pretest, T2=test after Poker, T3=test after Chess, T4=test after Gobang, T5=delayed test

The third research question attempts to find out which proficiency group will gain and retain more vocabulary in this experiment? The participants were grouped in the program according to their proficiency, thus a *one-way ANOVA* test was used to determine whether a significant difference existed among the three proficiency groups for all the five vocabulary tests taken during and after the program. All 46 participants were put into three proficiency groups based on their CSEPT scores. *Table4* presents the grouping information.

Table 4 *The Participants' Grouping Information*

Group	No.	CSEPT Range	Means	SD
Group 1	7	≤ 114	96.86	8.84
Group 2	30	115~172	142.43	17.37
Group 3	9	≥ 173	184.11	9.31
Overall	46	86~198	143.65	29.82

Note: 1. The total score of CSEPT is 360

The result in *Table 5* suggested that significant difference existed among the three groups for the first four tests, particularly the tests taken after players finished Poker Cards (*Test 2*) and Chinese Chess (*Test 3*). From the result of *post-hoc*, a slight difference was found between Group 1 and Group 3 in the pretest but no significance was found between Group 1 and Group 2. After finishing playing the Poker Card vocabulary game, the test result showed a significant difference between Group 1 and Group 2. On average, the three groups all improved about 15 items. The first group increased 16 items and the second group increased 17.6 items whereas the third group only increased 12.3 items (*Test 4*). The lowest proficiency group improved more items than those of the third group.

After finishing the Chess, a test was administered (*Test 3*) and a significant difference was found again between Groups 1 and 2. Group 2 increased to 157 items and Group 1 to 145.6 items. Group 1 improved 7.9 items, Group 2 improved 6.8 items, and Group 3 improved 5.6 items. Surprisingly the first group, the lower proficiency group improved more than the other two groups, and the second group surpassed the third, which was the highest proficiency group. When the program was completed, the last test result showed a slight difference among the three groups ($p=.032^*$). The gap between the three proficiency groups diminished.

However, the result of the delayed test which was administered 45 days after the program indicated that no significant difference was found among any of groups. The gap between the three proficiency groups disappeared which suggested proficiency no longer served as a factor in terms of players' long-term memory. On the other hand, the means of the retention test after 45 days decreased about 6 items, and the standard deviation increased from 7.6 to 8.5 (See *Table 3*, Pair V) which indicates that players' vocabulary size had reduced but somehow they maintained a considerable amount of vocabulary if comparing the retention score to that of the pretest (See *Table 3* Pair VI). Players were found to maintain the vocabulary gains similar to those they had achieved in *Test 3*. This proves that the game was able to help students remember the vocabulary they had acquired in the program to a certain degree. The result is also shown in *Table 5*.

Table 5 *Results of Three Groups' Performance on the Vocabulary Tests*

	<i>Test 1</i>		<i>Test 2</i>		<i>Test 3</i>		<i>Test 4</i>		<i>Test 5</i>	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Group 1	121.7	6.3	137.7	8.0	145.6	8.5	154.3	9.0	148.0	7.9
Group 2	132.6	11.6	150.2	8.2	157.0	7.4	162.3	7.2	155.8	8.4
Group 3	136.4	11.1	148.7	10.5	154.3	6.9	162.4	5.4	156.2	7.3
<i>p</i>	.029*		.005**		.003**		.032*		.072	
<i>f-ratio</i>	3.855		5.901		6.619		3.730		2.797	
<i>Post-hoc</i>	G3>G1*		G2>G1**		G2>G1**		G2>G1*			

Note: 1. * = $p < .05$, ** = $p < .01$, *** = $p < .001$

2. T1=pretest, T2=test after Poker, T3=test after Chess, T4=test after Gobang, T5=delayed test

With regard to each group's gain in each game and test, tables 6-8 showed similar results to those of the whole group. There were seven participants in the first group whose proficiency scores were the lowest among all participants (*Table 6*). When comparing their pretest result with the test after Gobang, the gain increased most significantly (Pair IV). Although the delayed test (Test 5) showed that they had lost some vocabulary, compared to the pretest (Test 1), their gain was still significant (Pair 6, $p=.001^{***}$). The p -value in Pair VII indicated that the participants' vocabulary gains decreased nearly to those of the third test; this phenomenon was found in all groups which led the researcher to ponder whether the game order served as a factor to players' vocabulary gains in the learning program. Another explanation could be that since 54.1% of compound nouns were repeated in Gobang. This made them stay within the comfort zone and thwarted them from memorizing new vocabulary. Therefore, even though they did learn a few new words in the gaming process of Gobang, they might forget them later.

Table 6 Results of First Group's Performance between Each Vocabulary Test

	Pair I		Pair II		Pair III		Pair IV		Pair V		Pair VI		Pair VII	
	T1	T2	T2	T3	T3	T4	T1	T4	T4	T5	T1	T5	T3	T5
M	121.7	137.7	137.7	145.6	145.6	154.3	121.7	154.3	154.3	148.0	121.7	148.0	145.6	148.0
SD	6.3	8.0	8.0	8.5	8.5	9.0	6.3	9.0	9.0	7.9	6.3	7.9	8.5	7.9
t	-4.024		-3.442		-4.620		-6.953		3.167		-5.966		-1.415	
p	.007**		.014*		.004**		.000***		.019*		.001***		.207	

Note: 1. *= $p<.05$, **= $p<.01$, ***= $p<.001$

2. T1=pretest, T2=test after Poker, T3=test after Chess, T4=test after Gobang, T5=delayed test

There were 30 participants in the second group whose CSEPT score was around the average among all participants (*Table 7*). Except for the pretest, the standard deviation among all the other tests denoted that the gap among the members was the most stable one compared to the other groups. If comparing all the tests, significant differences were found among all pairs (Pair I to Pair VI) which meant that participants continued expanding their vocabulary size. However, the p -value from Pair V also indicated that participants in the second group seemed to forget the vocabulary they had learned to a significant level, comparing Test 1 to Test 4 when it was taken on the completion of the program. On average, they lost around 7 test items. When comparing the pretest to the fourth (Pair IV) and the fifth tests (Pair VI), a significant difference still existed, meaning that in spite of some forgetfulness, participants did improve their vocabulary size throughout the program.

Table 7. Comparison of Second Group's Performance between Each Vocabulary Test

	Pair I		Pair II		Pair III		Pair IV		Pair V		Pair VI		Pair VII	
	T1	T2	T2	T3	T3	T4	T1	T4	T4	T5	T1	T5	T3	T5
M	132.6	150.2	150.2	157.0	157.0	162.3	132.6	162.3	162.3	155.8	132.6	155.8	157.0	155.8
SD	11.6	8.2	8.2	7.4	7.4	7.2	11.6	7.2	7.2	8.4	11.6	8.4	7.4	8.4
t	-12.472		-8.428		-8.875		-18.962		7.285		-16.977		1.577	
p	.000***		.000***		.000***		.000***		.000***		.000***		.126	

Note: 1. *= $p<.05$, **= $p<.01$, ***= $p<.001$

T1=pretest, T2=test after Poker, T3=test after Chess, T4=test after Gobang, T5=delayed test

Group 3 comprised the nine members who had the highest CSEPT scores among all the participants (*Table 8*). When comparing the pretest with the test after Gobang (Pair IV) and the delayed test (Pair VI), vocabulary gains were found to show a significant increase ($p=.000^{***}$). Comparing the gains from the pretest to Test 3, the gains reached some significance ($p=.010^{**}$, $p=.027^*$). Not till the last test did the group gain significantly (Pair

III, $p=.001^{***}$, Pair IV, $p=.000^{***}$) Despite a certain loss in the delayed test, participants overall increased their vocabulary size during the program.

Table 8. Comparison of Third Group's Performance between Each Vocabulary Test

	Pair I		Pair II		Pair III		Pair IV		Pair V		Pair VI		Pair VII	
	T1	T2	T2	T3	T3	T4	T1	T4	T4	T5	T1	T5	T3	T5
M	136.4	148.7	148.7	154.3	154.3	162.4	136.4	162.4	162.4	156.2	136.4	156.2	154.3	156.2
SD	11.1	10.5	10.5	6.9	6.9	5.4	11.1	5.4	5.4	7.3	11.1	7.3	6.9	7.3
<i>t</i>	-3.369		-2.713		-4.785		-9.693		4.128		-7.420		-1.241	
<i>p</i>	.010**		.027*		.001***		.000***		.003**		.000***		.250	

Note: 1. *= $p<.05$, **= $p<.01$, ***= $p<.001$

2. T1=pretest, T2=test after Poker, T3=test after Chess, T4=test after Gobang, T5=delayed test

However, the numbers of participants in Group 1 ($N=7$) and Group 3 ($N=9$) are comparatively smaller than those in Group 2. Only Group 2 had as many as 30 participants. From the statistical point of view, the results of groups 1 and 3 might need to be analyzed more carefully.

As stated earlier, this paper aims at discussing the retention rate of the participants after the learning program. The findings of the described purposes are crucial with regard to the design of this series of vocabulary games and its effectiveness in promoting a different way of memorizing vocabulary, as well as its effects on promoting long-term memory. As reported, 45 days after the completion of the program, the participants generally retained their vocabulary gains. From the comparison between the test after Gobang and the delayed test, participants were found to retain the vocabulary gains to a limited degree. Significant difference of decrease was found among all the three groups. However, if the pretest is compared with the delayed test, all participants managed to expand their vocabulary size to a significant level by playing the designed games. Further studies on analyses of types of vocabulary from which users can benefit more regarding vocabulary gains and retention would enable the researcher to modify the program and the games for greater learning effectiveness.

RECOMMENDATION

The designed vocabulary games have not yet been tried separately either in a regular classroom, or used as assisted learning material for individual learners of different backgrounds except for this initial study. To validate their effectiveness, further researches are suggested as follows:

1. The number of participants in this study is not representative regarding vocabulary gains by using games. It is recommended that a greater number and a greater variety of participants to be investigated so as to strengthen the effectiveness of the game-based vocabulary learning method among different groups of learners.
2. The separate games in the series can be tested independently either with the same group or different groups of learners in order to specify the actual vocabulary gains from each type of game, and henceforth modify the vocabulary training program.
3. The games can be played in different orders so to investigate if features of the games: Poker Cards, Chinese Chess and Gobang influence the results.
4. Quasi-experiments can be conducted in comparing learning vocabulary through the designed vocabulary games in the training program with a control group learning vocabulary from traditional approaches such as, memorizing from a

word list or repeating the words verbally or in writing.

REFERENCES

- Chen, M. (2009). Vocabulary games invention and its learning effectiveness. *Journal of Instructional Technology and Media*, 87, 43-67 (in press).
- Chen, M. (2008). Taiwan. Patent Application No. 097101167. Taiwan: Intellectual Property Office.
- Chen, Y. M. (2006). *The effect of keyword method on English vocabulary long-term retention of elementary school students in Taiwan*. (Master Thesis, Southern Taiwan University, Tainan, Taiwan, 2006). Electronic Theses Dissertations System, 094STUT0741024.
- Chien, Y. C. (2006). *Impacts of English vocabulary learning strategy instruction to 5th grade elementary school student*. (Master thesis, National Taipei University of Education, Taipei, Taiwan, 2006). Electronic Theses Dissertations System, 094NTPTC694012.
- Chu, J. W. (2007). *The effects of memory strategy instruction on vocabulary learning: A study of elementary fourth-grade students*. (Master thesis, National Pingtung Institute of Commerce, Pingtung, Taiwan, 2007). Electronic Theses Dissertations System, 096NPC05615004.
- Council of Europe (1996). *The common European framework of reference for languages*. http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf
- Crookall, D., Oxford, R. L. & Saunders, D. (1987). Towards a reconceptualization of simulation: From representation to reality. *Simulation/Games for Learning*, 17(4), 147-171.
- Dunbar, A. (2004). *Using portfolios to understand learners' responses to vocabulary learning strategy instruction and use: A case study*. (Master thesis, National Tsing Hua University, Hsinchu, Taiwan, 2004). Electronic Theses Dissertations System 093NTHU5094001.
- Francis, M. A & Simpson, M. L. (2003). Using theory, our intuitions, and a research study to enhance students' vocabulary knowledge. *Journal of Adolescent & Adult Literacy*, 47(1), 66-78.
- Gu, Y. & Johnson, R. K. (1996). Vocabulary learning strategies and language learning outcomes. *Language Learning*, 46, 643-679.
- Hadfield, J. (1990). *Intermediate communication games*. Harlow, Essex: Pearson Education.
- Horwitz, E. K. (1988). The beliefs about language learning of beginning university foreign language students. *The Modern Language Journal*, 72(3), 283-294.
- Hsu, K. F. (2004). *An iconic-morphological approach via commonly-used roots to English vocabulary teaching: How to help Chinese senior high school students memorize English vocabulary*. (Master thesis, National Chengchi University, Taipei, Taiwan, 2004). Electronic Theses Dissertations System, 092NCCU5238008.
- Huang, H. C. (2004). *Teaching vocabulary through drawing: A case study*. (Master thesis, National Chengchi University, Taipei, Taiwan, 2004). Electronic Theses Dissertations System 092NCCU5238006.
- Kojic-Sabo, I. & Lightbown, P. M. (1999). Students' approach to vocabulary learning and their relationship to success. *The Modern Language Journal*, 83(2), 176- 192.
- Language Training and Testing Center. <http://www.lttc.ntu.edu.tw/main.htm>
- Lewis, G. & Bedson, G. (1999). *Games for children*. Oxford: Oxford University Press.
- Macedonia, M. (2005). Games and foreign language teaching. *Support for Learning*, 20(3), 135-140.
- Marefat, H. & Shirazi, M. A. (2003). The impact of teaching direct learning strategies on the retention of vocabulary by EFL learners. *The Reading Matrix*, 3(2), 47-62.

- Martin-Chang, S. L., Levy, B. A. & O'Neil, S. (2007). Word acquisition, retention, and transfer: Findings from contextual and isolated word training. *Journal of Experimental Child Psychology*, 96, 37-56.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, 63(1), 59-82.
- Newby, T. J., Stepich, D. A. Lehman, J. D. & Russell, J. D. (2000). *Instructional technology for teaching and learning designing instruction, integrating computers, and using media* (2nd Ed). Upper Saddle River, New Jersey, Columbus, Ohio: Merrill Prentice Hall.
- On-Lai, A.L. (1993). The variety and effectiveness of strategies employed in vocabulary explanations in EFL classroom in Hong Kong. *Papers presented at the Annual International Language in Education Conference*, Hong Kong.
- Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*, 52(3), 513-536.
- Rasekh, Z. E. & Ranjbar, R. (2003). Metacognitive strategy training for vocabulary learning. *TESL-EJ* 7(2). <http://www-writing.berkeley.edu/tesl-ej/ej26/a5.html>
- Rieber, L. P. (1996). Seriously considering play: Designing interactive learning environments based on the blending of micro-worlds, simulations, and games. *Educational Technology Research & Development*, 44(2), 43-58.
- Sanaoui, R. (1995). Adult learners' approaches to learning vocabulary in second languages. *The Modern Language Journal*, 79(1), 15-28.
- Schmitt, N. & Schmitt, D. (1993). Identifying and assessing vocabulary learning strategies. *Thai TESOL Bulletin*, 5(4), 27-33.
- Schmitt, N., Schmitt, D. & Clapham, C. (2001). Developing and exploring the behavior of two new versions of the Vocabulary Levels Test, *Language Testing*, 18(1), 55-88.
- Shie, J. S. (2004). Main features of a game and their implications for language teaching. *Journal of Da-Yeh university*, 13(2), 119-129.
- Siyanova, A. & Schmitt, N. (2007). Native and nonnative use of multi-word vs. one-word verbs. *IRAL, International Review of Applied Linguistics in Language Teaching*, 45(2), 119-139.
- Thornton, G. C., & Cleveland, J. N. (1990). Developing managerial talent through simulation. *American Psychologist*, 45, 190-199.
- Wesche, M. B. & Paribakht, T. S. (2000). Reading-based Exercises in second language vocabulary learning: An introspective study. *The Modern Language Journal*, 84(2), 196-213.
- Yek, S. M.(2006). *Memory enhancing vocabulary learning strategy instruction*. (Master thesis, Ming Chuan University, Taipei, Taiwan, 2006). Electronic Theses Dissertations System, 094MCU05741005.

Effects of English Proficiency and Gender on Reading Strategy Use

Mei-chen Chen (陳美貞), Meng-fang Cheng (鄭孟舫)

National Pingtung Institute of Commerce

meichen@npic.edu.tw
s95307009@student-mail.npic.edu.tw

Strategies generate a context of learning autonomy. To date, compared to the existing literature of language learning strategies, not that many studies have specifically discussed reading strategies. This study aimed to probe the overall and specific types of reading strategy use among EFL college students. The effects of English proficiency and gender on strategy use were further examined. Based on the strategy taxonomy of Oxford (SILL, 1990), a reading strategy questionnaire with five types of reading behaviors (i.e., memory, cognitive, compensation, metacognitive, and social-affective) was constructed to suit the purposes of the study. In this quantitatively-oriented empirical research, 217 college English majors (M = 40, F = 177) were surveyed on their perceived strategy use of reading. Important results were produced: (a) college students used reading strategies in a moderate degree; (b) the top two types of strategy choice were found to be compensation and cognitive, and the bottom one was social-affective; (c) there were no significant differences among high-, mid-, and low-achievers in their overall and specific types of strategy use; (d) significant proficiency effects were detected in the use of individual strategies; more proficient learners were found to employ more read-aloud, skimming, guessing from the context, and analyzing sentence patterns and grammar structures strategies than less proficient learners; (e) there were no significant differences between male and female learners in their overall and specific types of strategy use; (f) significant gender effects were found in the use of individual strategies; females significantly employed more underlining, highlighting and notation strategies than males while reading; (g) metacognitive and cognitive strategies were most correlated with overall strategy use; these two types best predicted college EFL students' strategic behaviors of reading. Pedagogical implications and suggestions for further research were provided.

INTRODUCTION

Reading is a primary means for developing any language competence in all educational contexts. It helps cultivate synthesis and evaluation skills for learning autonomy (Grabe & Stoller, 2002). Reading is also essential for the acquisition of a foreign language. Written texts provide an access to language learning, and reading practice is greatly associated with the outcomes of language acquisition (Brown, 2004). Strategies facilitate language learning (O'Malley & Chamot, 1990; Oxford, 1990). Language learning strategies are specific behaviors or tools that people employ to enhance their acquisition, storage, retrieval, and use of information so as to become more active learners. In particular, reading strategies are crucial to most people; learners could become effective readers when they employ such techniques to assist them in reading (Chen & Huang, 2007; Huang, Chern & Lin, 2006; Shang, 2007; Yeh, 2006). These strategies strengthen reading comprehension and help overcome reading difficulties. Although use of strategy is greatly associated with reading achievement (Shang, 2007; Wang, 2006), previous research has found that EFL students are moderate strategy users (e.g., Ho, 2007; Hsu, 2007; Wu, 2003; Wu, 2007; Yu, 2006). How to mold our students to be strategic readers has presented a great challenge for the nonnative foreign language teachers.

Oxford (1990) identified six types of learning strategies for ESL/EFL language learners. Memory strategies, sometimes called mnemonics, are techniques that help learners store and retrieve new information, such as grouping and applying imagery. Cognitive strategies are typically the most popular strategies with language learners. They enable learners to comprehend and produce new language by different means, such as summarizing or reasoning deductively. Compensation strategies allow learners to use the language despite their limitations in knowledge. Guessing is one example of such type. Metacognitive strategies are techniques utilized to help learners control their cognition and coordinate the learning process. Planning and evaluating are common behaviors for this type. Affective strategies assist learners to regulate their emotion, motivation, and attitude, and social strategies involve interaction with other people. Positive self-talk, self-rewarding, and cooperating with peers are examples of these social-affective strategies.

Several empirical studies have been conducted with the learning strategy inventory developed in Oxford (1990) to explore the strategic behaviors of language learners of differing grade levels: junior high (e.g., Yu, 2006), senior high (e.g., Liu, 2004), and college (e.g., Chang, 1992; Chang, 1997; Hsiao & Oxford, 2002; Huang, 2005; Lin, 2006; Su, 2005; Sy, 1995; Sy & Liou, 1994; Tsai, 2005; Wharton, 2000; Yang, 1994). Reading behaviors take up a crucial, fundamental part of language learning (Grabe & Stoller, 2002). The strategy taxonomy proposed in Oxford's study was of great concern in the present reading strategy research due to its well-defined, distinctive strategy classification and suitability in assessing learning behaviors in the EFL context of Taiwan. It would be served as the theoretical framework to examine strategy use, particularly, in the domain of EFL reading.

Previous literature has indicated that English proficiency is highly correlated with strategy use (e.g., Chang, 1992; Chang, 1997; Hsiao & Oxford, 2002; Huang, 2005; Lin, 2006; Liu, 2004; Shang, 2007; Sheorey & Mokhtari, 2001; Sy & Liou, 1994; Tsai, 2005; Wharton, 2000; Wu, 2003; Yang, 1994; Yu, 2006). Use of strategies significantly promotes English reading comprehension (e.g., Chen & Huang, 2007; Hsu, 2003; Hu, 2007; Hu, 2006; Huang, et al., 2006; Hung, 2005; Hung, Tsou & Wu, 2005; Kuo, 2003; Phakiti, 2003; Shang, 2007; Shen, 2003; Sheorey & Mokhtari, 2001; Wang, 2006; Yeh, 2006). Despite these facts, a question is raised here. Do high-achievers still use more strategies than low-achievers while reading, particularly when these learners are all recognized as advanced EFL learners (i.e., college English majors) in Taiwan? More research is required to clarify this issue. In addition to proficiency, gender is another influential factor on language learning (Chang, 1992; Chang, 1997; Hsu, 2007; Huang, 2005; Kuo, 2003; Oxford, 1990; Phakiti, 2003; Sheorey & Mokhtari, 2001; Sy, 1995; Sy & Liou, 1994; Yang, 1994; Yu, 2006). Previous studies have shown that females reported greater strategy use than males (e.g., Ehrman & Oxford, 1988; Green & Oxford, 1995; Oxford, 1993; Oxford & Nyikos, 1989; Ray, Garavalia & Gredler, 2003; Sheorey & Mokhtari, 2001; Sy, 1995; Sy & Liou, 1994; Wu, 2003; Yang, 1994; Yu, 2006). However, when it comes to a more specific task of reading, are females still more strategic than males? The present study intended to probe these questions with more proficient adult learners in the context of EFL.

Moreover, previous studies on strategies tend to examine learners' general language learning behaviors (e.g., Chang, 1992; Chang, 1997; Chen, 2005; Ehrman & Oxford, 1988; Green & Oxford, 1995; Hsiao & Oxford, 2002; Huang, 2005; Lin, 2006; Liu, 2004; Oxford, 1993; Oxford & Nyikos, 1989; Ray, et al., 2003; Su, 2005; Sy, 1995; Sy & Liou, 1994; Tsai, 2005; Wharton, 2000; Wu, 2003; Yang, 1994; Yu, 2006). Relatively, not that many studies have discussed the more specific strategic behaviors in reading (e.g., Chen & Huang, 2007; Huang, et al., 2006; Shang, 2007; Sheorey & Mokhtari, 2001). Needless to say, such reading strategy research involving the regional context of EFL is even less. To date, the effects of individual differences in proficiency and gender on the choice of reading strategies

among college EFL learners of technological orientation have not been well-explored. These two factors have been claimed as having significant impacts on strategy use among ESL/EFL college learners (Oxford & Nyikos, 1989). Particularly, gender differences have not received as much attention as proficiency factor in the reading strategy literature. There have been a very limited number of such studies involving adult EFL readers (i.e., Hsu, 2007; Phakiti, 2003). These studies have, unfortunately, produced an insignificant gender effect on the reading strategy use, which contradicts what has been detected in the learning strategy literature with a female superiority over male learners in strategy use as mentioned previously. This gender inconsistency found between the learning strategy literature and the limited reading strategy research calls for more studies to discern the issue of sex differences on the choice of reading strategies. Generally, this study attempted to explore both variables to shed light on the role of proficiency and gender in reading strategy use.

The study aimed to examine the frequency and type of reading strategy use among college English majors in southern Taiwan. The individual differences in ability and gender were probed underlying the strategy use. Basically, there were four main purposes: first, to explore general reading strategy use among adult EFL learners; second, to examine the effect of differing reading proficiency (i.e., high-, mid-, and low-level readers) on strategy use; third, to probe the effect of gender (i.e., males vs. females) on strategy use; fourth, to assess the relationship between type of strategy use and overall strategic behavior in reading. Several predictions were made according to the research purposes. Since EFL students were found to be medium learning strategy users (e.g., Chang, 1992; Liao, 2000; Peng, 2002; Wu, 2007; Yu, 2006), reading strategies were expected to be moderately used by the college learners in the study. As for the ability effect, English proficiency had generally been asserted as a significant factor to affect learning strategy use (e.g., Huang, 2005; Liu, 2004; Tsai, 2005; Wharton, 2000; Wu, 2003; Yang, 1994; Yu, 2006). It was, therefore, hypothesized that high-achievers would significantly employ more reading strategies than their low-achieving counterparts. When the gender effect was concerned, it was anticipated that female readers would use more strategies than their male counterparts because the former had generally been found as better strategy users (e.g., Ehrman & Oxford, 1988; Green & Oxford, 1995; Oxford & Nyikos, 1989; Ray, et al., 2003; Su, 2003; Sy & Liou, 1996; Wu, 2003; Yu, 2006). Lastly, metacognitive and cognitive strategies had been indicated having significant associations with reading achievement (e.g., Hsu, 2007; Shang, 2007; Wang, 2006); it was thus predicted that such strategies would produce stronger relationships with overall strategic behavior than other strategy types. It was hoped that the findings would add credence to the existing strategy literature, provide EFL teachers with more sophisticated knowledge of student strategy use, and infuse remedial reading programs with pertinent strategy training to achieve optimal teaching and learning.

LITERATURE REVIEW

Learning is a dynamic course in which learners employ strategic methods to process a variety of information (O'Malley & Chamot, 1990). Learning strategies are specific behaviors or thoughts that students engage in to facilitate their learning (Oxford, 1990; Wenden, 1987). To date, a few researchers have taken the initiative in classifying learning strategies in different taxonomies. Rubin (1981) proposed six direct strategies (clarification/verification, guessing/ inductive inferencing, deductive reasoning, practice, memorization, and monitoring), and two indirect strategies (creating opportunities for practice and production tricks). Brown and Palinscar (1982), and O'Malley and Chamot (1990) raised a similar taxonomy with three broad types of strategies: metacognitive strategies which involved planning, monitoring, and evaluating the learning process, cognitive strategies which involved meaningfully manipulating materials in specific learning

tasks, and social-affective strategies in which learners used their affective control and worked with others on a task. In addition, the six-factor strategy classification developed by Oxford (1990) has been claimed as the most consistent and comprehensive taxonomy for learners' strategy use (Hsiao & Oxford, 2002; Liu, 2004). It was this strategy categorization that was of great concern in this study.

Reading strategies facilitate language learning (Chen & Huang, 2007; Huang, et al., 2006; Shang, 2007; Wang, 2006). Basically, these strategies are thoughts or behaviors used to conceive texts, learn new ideas, and overcome difficulties (Hsu, 2007; Richards, Platt & Platt, 1992). Strategic reading intrinsically motivates students and allows them to be aware of their progress and involvement in reading (Brown, 2001; Hung, et al., 2005). With the strategic behaviors, meaningful and autonomous learning takes place. In addition, quite a few empirical studies have probed the strategy taxonomies developed by Taillefer and Pugh (1998) (e.g., Hu, 2006; Hung, 2001), and by Mokhtari and Sheorey (2002) (e.g., Hung, 2005). The two taxonomies are specifically associated with reading behaviors. Taillefer and Pugh (1998) categorized strategy use into general reading strategies as those regularly used while reading, and local problem-solving strategies used particularly when learners' comprehension was blocked. With such strategy framework, Hung (2001) found that there was no significant difference for Taiwanese senior high school students in their use of general reading strategies and local problem-solving strategies. In addition, Hu (2006) further studied the use of local problem-solving strategies for both junior high school and college students. It was found the former often repeatedly read difficult sentences while encountering a comprehension problem and the latter would refer to the discourse of a text when comprehension problems occurred. The college learners were better strategy users in reading than their counterparts of senior high school.

Mokhtari and Sheorey (2002) proposed three broad types: global strategies, problem solving strategies, and support strategies. Global strategies were well-planned behaviors in which people monitored or managed their reading with a purpose in mind. Problem solving strategies involved direct actions and procedures like adjusting the speed of reading, guessing, and rereading. Support strategies referred to basic support mechanisms to aid the learner in comprehending the text such as using a dictionary, taking notes, underlining, or highlighting textual information. With such strategy classifications, Hung (2005) reported that college freshmen were more familiar with problem solving strategies, and global strategies were less known to them. Besides, among four motivational aspects (i.e., motivational intensity, intrinsic motivation, extrinsic motivation, and requirement motivation), extrinsic motivation was positively correlated with the use of reading strategies. Particularly, those who were more extrinsically motivated would employ more support strategies.

Research has revealed that several factors are likely to affect choice of reading strategies for individual learners: language proficiency (e.g., Chen & Huang, 2007; Chiu, 2007; Hsu, 2003; Hsu, 2007; Hu, 2007; Hu, 2006; Hung, 2005; Hung, et al., 2005; Huang, et al., 2006; Kuo, 2003; Shang, 2007; Shen, 2003; Sheorey & Mokhtari, 2001), gender (e.g., Hsu, 2007; Hung, 2001; Kuo, 2003; Phakiti, 2003; Poole, 2005; Sheorey & Mokhtari, 2001), different college systems (e.g., Hsu, 2003), grade level (e.g., Hu, 2006), motivation (e.g., Hung, 2005), text types (e.g., Hung, 2001; Kuo, 2003; Shen, 2003), language backgrounds (e.g., Hsu, 2007; Sheorey & Mokhtari, 2001), and reading difficulties (e.g., Hu, 2006; Huang, et al., 2006; Hung, 2001; Kuo, 2003; Shen, 2003). In addition, a training intervention is applied to enhance readers' strategy use (e.g., Hung, et al., 2005; Shang, 2007; Wang, 2006), and comparisons in Chinese and English reading also have been made under the framework of strategy use (e.g., Hu, 2007; Hu, 2006). Furthermore, test-based, text-based, online and general reading strategy uses are common empirical orientations for reading strategy research. Note that, various approaches have been used to explore readers' use of strategies:

questionnaires (e.g., Chiu, 2007; Hsu, 2003; Hsu, 2007; Hu, 2007; Hu, 2006; Hung, 2005; Hung, et al., 2005; Huang, et al., 2006; Kuo, 2003; Phakiti, 2003; Shang, 2007; Shen, 2003; Sheorey & Mokhtari, 2001), a think-aloud method (e.g., Chen & Huang, 2007; Kuo, 2003; Shen, 2003), an empirical training (e.g., Hung, et al., 2005; Shang, 2007; Wang, 2006), students' journals (e.g., Hung, et al., 2005; Shang, 2007), and interview (e.g., Huang, 2007; Kuo, 2003; Shen, 2003). Obviously, administering questionnaires is a common practice in reading strategy research.

English proficiency is highly correlated with strategy use (e.g., Chen & Huang, 2007; Chiu, 2007; Hsu, 2003; Hsu, 2007; Hu, 2007; Hu, 2006; Hung, 2005; Hung, et al., 2005; Huang, et al., 2006; Kuo, 2003; Shang, 2007; Shen, 2003; Sheorey & Mokhtari, 2001). Some learners do approach language learning tasks more successfully than others because they have better access to strategy use while the inexperienced ones lack this self-instruction ability (Rubin, 1987; Rubin & Thompson, 1994). Shang (2007) also detected that effective EFL college readers used more metacognitive and testing strategies than ineffective ones. Chiu (2007) further detected that cognitive strategies were used most frequently among college English majors, and high-achievers tended to use more memory strategies than their low-achieving counterparts while reading. In addition, Huang, et al. (2006) examined EFL readers' online reading strategy use and found that support strategies were used most frequently, and problem-solving strategies were used the least. Students of different proficiency levels used different strategies; high-achievers used more global strategies, and low-achievers used more socio-affective strategies. Particularly, technological college students used more repair strategies than their regular college counterparts. Hsu (2007) indicated that students' reading performance was greatly associated with the frequency of reading strategy use; more proficient students used more strategies. Shen (2003) reported that translation strategies were most frequently used by EFL learners of junior high school. In particular, proficient readers could employ different types of strategies to overcome different reading difficulties, but less proficient ones depended chiefly on the strategies of skipping unknown words and questioning. Hu (2007) revealed that uses of Chinese and English test-based reading strategies were greatly associated with reading abilities. Senior high school high-achievers used more English reading strategies than their low-achieving counterparts. Besides, all learners frequently employed more English reading strategies than Chinese ones. Despite this superior strategy preference over English, Hu (2006) reported students used Chinese reading strategies more effectively than English ones.

Gender affects learning strategy use (e.g., Ehrman & Oxford, 1988; Green & Oxford, 1995; Oxford & Nyikos, 1989; Ray, et al., 2003; Sy, 1995; Sy & Liou, 1994; Yang, 1994; Yu, 2006). Past literature has generally indicated that females used more learning strategies than males (e.g., Ehrman & Oxford, 1988; Green & Oxford, 1995; Oxford & Nyikos, 1989; Ray, et al., 2003; Sy, 1995; Sy & Liou, 1994; Yu, 2006). However, this female superiority in strategy use is not that obvious in the reading strategy research. Kuo (2003) found that English proficiency played a more critical role than gender on junior high students' choice of reading strategies since no significant differences were detected between the male and female readers in their strategy use. Similar findings of insignificant gender effect were also produced in Sheorey and Mokhtari (2001) with ESL adult learners and native English speakers, Poole (2005) with advanced ESL college students, Phakiti (2003) with Thai college students, Hung (2001) with Taiwanese senior high school students, and Hsu (2007) with EFL technological college students in Taiwan. These studies consistently detected sex differences not in the overall use of reading strategies, but in a certain type or individual reading strategies. Generally, the apparently inconsistent findings between the learning strategy and limited reading strategy literatures call for more research to confirm the gender effect on reading strategy use.

METHODOLOGY

The purpose of the current research was to probe the effects of English proficiency and gender on reading strategy use among technological English majors in southern Taiwan. This study is a survey research with a self-reported questionnaire. Questionnaire survey is a dominant approach in strategy research and has been asserted as a neutral, nonthreatening, cost-effective measure to assess learners' strategy choice (Chamot, 2004; Oxford, 1990; Oxford & Burry-Stock, 1995).

Participants

The participants of this study were 217 students (M = 40, F = 177) from six technological colleges located in southern Taiwan, including Tainan, Kaohsiung and Pingtung. They were English majors and currently enrolled in seven different English reading courses for college sophomores and juniors. Most of them had received at least eight years of formal English education. In this research, the students were asked to fill out a questionnaire of reading strategies. Clear instruction was provided in each class to facilitate the surveying procedures. The completed questionnaires were then collected for further data analyses.

Instruments

Reading Achievement. Previous research has used academic achievement as an indicator of learners' ability levels on strategy use since it reflects learners' consistent academic performance (e.g., Hu, 2007; Wu, 2003; Wu, 2007). The present study therefore used the participants' semester grades of English reading to identify their English proficiency levels (i.e., high-, mid-, and low-reading achievement) to probe the effect of differing ability on strategy preferences. The top-one-third learners in each class were grouped as high-achievers, the middle-one-third, as mid-achievers, and the bottom-one-third, as low-achievers. All of the participants were not aware of such classifications.

Reading Strategy Questionnaire. Based on the learning-strategy taxonomy of Oxford (1990), a reading strategy questionnaire with five types of reading behaviors was developed. These types were memory, cognitive, compensation, metacognitive, and social-affective. The taxonomy was chosen as the theoretical framework due to its comprehensive categorizations and applicability in an EFL strategy context. With a reference to the questionnaires developed by Oxford (1990), Taillefer and Pugh (1998), Mokhtari and Sheorey (2002), and Liao (2000), items of the present questionnaire were constructed, and several new items were added to fit the EFL cultural orientations and the intention of this survey study. To achieve the content validity, the questionnaire was reviewed by three professors specialized in the field of English teaching and learning. It was then pretested in a pilot study with 14 technological college students of homogeneous backgrounds as those of the target learners. The reliability coefficient associated with the internal consistency of the questionnaire items was .88, which was considered very reliable for a survey test. The validity and reliability of the questionnaire were carefully examined with the pretest to enhance the generalization of any potential findings. The Chinese version of the questionnaire was administered to the pretest students and the target college participants to eliminate confusions or misunderstandings on some of the items due to limited English knowledge.

Basically, this questionnaire contained two parts: 3 items for student background information (i.e., student ID number, gender, and years of learning English) and 39 items for English reading strategies. A five-point Likert scale of five possible responses ranging from "strongly agree (5 points)" to "strongly disagree (1 point)" was used to display the participants' varying degrees of agreement or disagreement with each statement of strategy use. These items revealed five types of reading behaviors: memory (items 1 to 8), cognitive (items 9 to 18), compensation (items 19 to 24), metacognitive (items 25 to 31) and social-affective (items 32 to 39). Sample items of the five subgroups of reading strategies

are provided in Table 1.

Table 1 *Sample Items of Five Types of Reading Strategies*

Type	Sample item
Memory	<ul style="list-style-type: none"> • I read with imagining. For instance, when I read <i>The Old Man and the Sea</i>, I would imagine I am on the sea. • I use new English words in a sentence so I can remember them. • When I see important words, I recite them several times to remember.
Cognitive	<ul style="list-style-type: none"> • I skim first and then go back to read carefully. • I analyze sentence patterns and grammar structures while reading. • I notice headings, subheadings and topic sentences of an article.
Compensation	<ul style="list-style-type: none"> • I make guesses from the context to figure out unclear patterns or phrases. • I underline or make a notation for unclear words or phrases, for instance, with a question mark “?” or some other symbols. • I translate unclear English sentences into Chinese to help comprehend a text.
Metacognitive	<ul style="list-style-type: none"> • I notice whether I catch the main ideas of an article. • I slow down the reading pace and read for several times when I see difficult words or sentences. • I study articles that match my reading ability.
Social-affective	<ul style="list-style-type: none"> • I read interesting stories or articles to relax myself. • When my reading comprehension is blocked, I notice if I am tense or upset. • When I have questions in reading, I seek help from my teachers or classmates.

Besides, the Cronbach alpha procedure was implemented to determine the internal-consistency reliability of this survey study with the 217 student participants. It was found that the questionnaire of reading strategies had an overall reliability coefficient as .90, and more specifically, memory strategies as .68 (8 items), cognitive strategies as .72 (10 items), compensation strategies as .58 (6 items), metacognitive strategies as .72 (7 items) and social-affective strategies as .69 (8 items). Though the coefficients of the five types were relatively lower due to the limited number of test items involved in each type, the overall coefficient of .90 suggested this questionnaire was a reliable instrument to measure the college students' use of reading strategies.

Procedures

There were several stages to collect data for this study (see Table 2). First, the reading strategy questionnaire was developed through a careful review by three professors and a pretest on 14 college English majors. Then, permissions were gained from six English reading teachers of technological college whose classes were studied for this research. These teachers also agreed to provide their students' semester grades of English reading. These grades helped group the students into three different levels of proficiency (high-, mid-, and low-achievers). The students were unknown to such ability categorizations. When the survey study was conducted in class, the purpose of study was told in the first place. The students were also informed that (1) they had to answer the questions honestly and intuitively; (2) their individual responses would not affect the grades of their studies; (3) the results of the study were only meant for research use and should be kept confidential. When the instruction to fill out the questionnaire was provided, the students were given 15 minutes to

work with it. Meanwhile, the researcher and the English teacher walked around the classroom to welcome any questions from the students. Eventually, they checked for any missing items from each returned questionnaire on the spot. The entire surveying procedure took about 30 minutes to complete.

Table 2 *Data Collection Procedures*

Stage	Data collection	Time
1	Develop a questionnaire based on the Oxford's taxonomy.	
2	Have the questionnaire reviewed by three experts.	
3	Pretest the questionnaire.	
4	Gain research permissions from six English teachers.	
5	Conduct survey study on the student participants in class.	
	Provide orientation of the study.	5 minutes
	Administer the questionnaires.	15 minutes
	Collect the questionnaires on the spot and check for any missing items.	10 minutes
6	Collect the participants' English grades from the teachers.	

Data Analyses

The quantitative data collected from the questionnaire of reading strategies were analyzed by the SPSS (Statistical Packages for the Social Science). The descriptive statistics of frequencies, means, and standard deviations were reported to exhibit the participants' overall strategy use and to determine the priority of specific types of strategy use. In addition, the procedures of ANOVA (i.e., one-way analysis of variance) were executed to explore the effect of differing English proficiency (i.e., high-, mid-, and low-achievement) on overall and specific types of strategy use. The independent-samples t-tests were conducted to probe the effect of gender (i.e., males vs. females) on overall and specific types of strategy use. Lastly, Pearson product-moment correlation analyses were used to determine the relationships between the individual types of strategy use and overall strategic behavior. The level of statistical significance for all ANOVA tests and t-tests was set at .05, and for correlation tests, at .01.

RESULTS

The statistical results of the study are presented in this section. Overall strategy use and priority of specific types of reading strategy use among the subjects are addressed first. The effect of differing English proficiency on overall and specific types of strategy use is then revealed, followed by a description for the effect of gender. Lastly, the relationships between individual types of strategy use and overall strategic behavior are reported. Interpretations of the results are also provided along.

General Reading Strategy Use. The means for the use of five strategy types among the college participants were within the range of 3.35 to 3.89 (see Table 3). The grand mean of strategy use was 3.67, which suggested that college students used reading strategies in a moderate degree. The top two types of strategy choice were found to be compensation ($M = 3.89$, $SD = .48$), and cognitive ($M = 3.80$, $SD = .47$), and the bottom one, social-affective ($M = 3.35$, $SD = .55$). Generally, compensation and cognitive strategies were most-frequently used, and social-affective strategies were least-frequently employed among EFL college students.

Table 3 Overall and Specific Types of Reading Strategy Use

Type	M	SD	Ranking
Memory	3.75	.52	3
Cognitive	3.80	.47	2
Compensation	3.89	.48	1
Metacognitive	3.57	.51	4
Social-Affective	3.35	.55	5
Overall strategy	3.67	.41	

Effect of Differing Ability on Strategy Use. The descriptive statistics of strategy use on strategy type by proficiency level are provided in Table 4. The one-way ANOVA procedure revealed that there was no significant difference among the high-, mid-, and low-achievers in their overall strategy use ($F = 2.28, p = .11 > .05$). In addition, there was no significant mean difference among the three ability groups for each type of strategy use. The rankings of the reported strategy use were also the same among the three ability levels. Although significant proficiency effects were not found in overall and specific types of strategy use, they were detected in four individual strategies: (1) read-aloud strategy ($F = 3.61, p = .03 < .05$), (2) skimming strategy ($F = 3.39, p = .03 < .05$), (3) strategy of analyzing sentence patterns and grammar structures ($F = 4.12, p = .04 < .05$), and (4) guessing from the context ($F = 3.87, p = .01 < .05$). The high- and mid-achievers generally reported greater strategy use than their low-achieving counterparts for these strategies. In sum, significant proficiency effects were not detected in overall and specific types of strategy use, but in the use of four individual strategies.

Table 4 Descriptive Statistics for Strategy Type X Proficiency Level

Type	High-achievers (n=72)			Mid-achievers (n=74)			Low-achievers (n=71)		
	M	SD	Ranking	M	SD	Ranking	M	SD	Ranking
Memory	3.73	.45	3	3.82	.59	3	3.70	.51	3
Cognitive	3.76	.42	2	3.87	.47	2	3.75	.50	2
Compensation	3.90	.53	1	3.94	.48	1	3.82	.43	1
Metacognitive	3.56	.51	4	3.65	.52	4	3.50	.51	4
Social-Affective	3.33	.56	5	3.45	.56	5	3.26	.54	5
Overall strategy	3.65	.38		3.75	.43		3.60	.41	

Effect of Gender on Strategy use. Table 5 presents descriptive statistics for overall and specific types of strategy use by gender. The independent-samples t-test revealed that there was no significant mean difference between males and females in terms of overall strategy use ($t(215) = -.83, p = .41 > .05$). Furthermore, the rank orders of strategy usage for males and females were the same; both groups used compensation strategies most frequently, and social-affective strategies least frequently. There was also no significant mean difference between the two gender groups for each type of strategy use. Note that two individual strategies were detected to reach statistically significant differences in gender effect: (1) underlining or highlighting strategy for important words or sentences ($t(215) = -3.34, p = .002 < .05$), and (2) notation strategy for unclear words or phrases ($t(215) = -3.66, p = .001 < .05$). The results generally indicated that female college learners significantly used more underlining, highlighting and notation strategies than their male counterparts.

Table 5 Descriptive Statistics for Strategy Type X Gender

Type	Male (n=40)			Female (n=177)		
	M	Ranking	SD	M	Ranking	SD
Memory	3.68	3	.63	3.77	3	.49
Cognitive	3.76	2	.55	3.81	2	.45
Compensation	3.83	1	.60	3.90	1	.45
Metacognitive	3.47	4	.59	3.59	4	.49
Social-affective	3.31	5	.71	3.36	5	.52
Overall strategy use	3.61		.53	3.68		.38

Relationships between Type of Strategy use and Overall Strategic Behavior.

The Pearson product-moment correlation procedure was conducted to examine the relationships between the individual types of strategy use and overall strategic behavior. As a result, each type was significantly correlated with the overall strategy use (see Table 6). Particularly, metacognitive and cognitive strategies had stronger associations with overall strategic behavior than the other strategy types. The correlation coefficients were $r = .86$ and $r = .85$ respectively, with .01 significance level. These two types best predicted college students' reading behavior.

Table 6 Correlations between Type of Strategy Use and Overall Strategic Behavior

Type	Overall strategy	Ranking
Memory	.797**	3
Cognitive	.853**	2
Compensation	.692**	5
Metacognitive	.856**	1
Social-Affective	.796**	4

* $r < .01$, two-tailed.

DISCUSSION

General Reading Strategy Use. As what had been predicted earlier, the study found that reading strategies were moderately employed among technological college students. This finding was consistent with those reported in the previous strategy research (e.g., Ho, 2007; Hsu, 2007; Wu, 2003; Wu, 2007). When the type of strategy use was concerned, compensation and cognitive were found to be the most-frequently used strategies, and the least used type was social-affective strategies. Specifically, compensation strategies (e.g., guessing from the context) aid learners in overcoming language gaps despite their limitations in knowledge. This type of strategy is even more important for the less proficient learners who encounter difficulties more often than the skilled ones (Oxford, 1990). The dominant use of compensation strategies is also noted in the previous reading strategy studies in the context of EFL (e.g., Chiu, 2007; Ho, 2007; Hu, 2007; Huang, 2005; Hung, 2001). In addition, cognitive strategies enable learners to comprehend and produce new language via various means, such as skimming, and analyzing sentence patterns and grammar structures (Oxford, 1990). Chang (1992) specifically revealed that language learners practiced to reach their acceptable proficiency levels so that they employed cognitive strategies more often than any other type. The present study echoed the previous research in learners' preference for cognitive strategies (e.g., Chang, 1992; Chiu, 2007; Huang, 2005; Hung, 2005; Sheorey and Mokhtari, 2001).

Note that social-affective domain receiving the least strategy use might be culture-conditioned (Chang, 1992; Huang, 2005; Liao, 2000). Such strategies involve social

interaction and self-esteem; positive self-talk, self-rewarding, and cooperating with peers are associated with these strategic acts (Oxford, 1990). As Oxford (1990), Politzer and McGroarty (1985), and Chang (1992) stated, the social acts of asking for help and having a language partner were apparently more related to the Western students' learning behaviors rather than the Asian ones. In Taiwan, the instructional system is test-based and teacher-centered. The EFL reading classes often engage in mechanical learning and error correction (Chang, 1992; King, 2002). The students often take their learning anxiety for granted and blame their unsuccessful achievement for their inabilities (Chang, 1992; King, 2002). They rarely reward their hard work, face their learning anxiety, or share their feelings when they learn English (Huang, 2005; Liao, 2000). In general, EFL readers lack social and affective strategies, but they employ more compensation and cognitive strategies to compensate for the cultural limitations to facilitate their language acquisition.

Effect of Differing Ability on Strategy Use. The present study found that differing English proficiency did not affect readers' choice of strategies in term of overall and specific types of strategy use. The finding, unfortunately, contradicted what had been predicted earlier for a positive ability effect on the strategy use. Although English proficiency has generally been found a significant factor affecting language learning strategy use (e.g., Chang, 1992; Chang, 1997; Huang, 2005; Liu, 2004; Tsai, 2005; Wharton, 2000; Wu, 2003; Yang, 1994; Yu, 2006), Hung (2005) reported a slight relationship, though significantly correlated ($r = .17, p < .05$), between reading proficiency and strategy use. It seems when it comes to more specific learning tasks of reading, the relationship between ability and strategy use is not that apparent. More research is needed to confirm this inference.

Another plausible explanation for the insignificant ability effect in this study is that the study recruited more proficient adult EFL learners (i.e., 217 college sophomores and juniors majoring in English), who could have, more or less, received reading strategy instruction or naturally picked up certain strategies along their college studies of various English subjects. That is, reading strategy instruction may have been integrated into their regular English coursework. This learning experience could have moderated or diminished the differences between the good and poor college readers in their strategy use. Poole (2005) further pointed out that task demands and contextual motivation could be two other potential factors for insignificant ability effect on the use of reading strategies. Additionally, one more explanation for the insignificant ability effect is well-described in Chang (1992) that language learning strategies are problem-oriented to help learning, and those who employ them extensively would possibly become automatic and unconscious strategy users. The more proficient college subjects in the study could have employed reading strategies automatically without much awareness, so it was likely that they did not report using them habitually and frequently in the survey. Note that using a self-reported questionnaire to collect data has an unavoidable limitation; the subjects' self-perceived strategy use might somehow differ from their actual strategy preferences (Poole, 2005).

In addition, it deserves our attention that significant proficiency effects were detected in the use of individual strategies; more proficient learners were found to employ more read-aloud, skimming, guessing from the context, and analyzing sentence patterns and grammar structures strategies than less proficient ones. This piece of evidence suggests these strategies are the critical, effective strategies that differentiate successful and unsuccessful readers. Specifically, read-aloud strategy could be done with or without an understanding of the contents (Richards, et al., 1992). When learners read the text aloud, this pronounceable reading ability tends to bring about self-confidence, especially for proficient language learners (Ho, 2007). Chao (2006) and Lo (2007) further found that the read-aloud strategy promoted elementary school students' English learning abilities, particularly in the areas of listening, speaking and comprehending sentence patterns. These

young learners shared what they knew to their peers and helped each other to learn more new things. Skimming is a practical reading strategy to comprehend the text and control test time effectively, as suggested in Wu (2003), Hu (2007), and Yeh (2006). Guessing is essential for people to understand new information through a variety of clues without necessarily comprehending all the details of text. Good language learners make educated guesses when confronted with unknown expressions, but less adept ones often look up every unfamiliar word in the dictionary (Oxford, 1990). Chiu (2007) and Ho (2007) further reported that proficient readers frequently examined sentence patterns and grammar structures while studying English. In sum, these strategies are a potential key to promoting the proficiency of less proficient EFL readers.

Effect of Gender on Strategy Use. This study found that there were no significant differences between male and female learners in terms of overall and specific types of strategy use. The finding was somewhat congruent with what had been detected in the limited reading strategy research with adult ESL/EFL readers (e.g., Hsu, 2007; Phakiti, 2003; Poole, 2005; Sheorey & Mokhtari, 2001), EFL students of senior high school (e.g., Hung, 2001), and EFL learners of junior high school (e.g., Kuo, 2003). In these studies, significant sex differences were not detected on the overall use of reading strategies, but on certain types or individual reading strategies. The insignificant gender effect commonly reported in the reading strategy research has conflicted what has been documented in the learning strategy literature with significant female superiority in strategy use (e.g., Ehrman & Oxford, 1988; Green & Oxford, 1995; Oxford & Nyikos, 1989; Ray, et al., 2003; Su, 2003; Sy & Liou, 1996; Wu, 2003; Yu, 2006). It appears that the effect of sex differences is moderated or diluted when it comes to a more specific learning task, in this case, of reading. Such gender impact on reading strategy use is no longer as significant as the gender effect detected in the language learning strategy research with a composite mixture of the four language skills.

It is worth noting that significant gender effects were specifically found in the use of individual strategies; female learners significantly employed more underlining, highlighting and notation strategies than male ones. Elaborately, underlining strategy involves drawing a line under a word or expression, highlighting strategy, marking key words or sentences in a passage with the use of a colored marker, and notation strategy, drawing symbols (such as a star, a check, a question mark, etc.) as a sign of emphasis. These strategies are notetaking techniques students use to help them notice important or unclear messages while reading and to reduce the amount of information that they must remember for a subsequent review. They are commonly used among EFL students as reported in Chiu (2007), Ho (2007), and Hu (2007). They reflect the mental status of the readers in processing information while reading. More importantly, these techniques help tell if the readers succeed in getting important main ideas (Caverly, Orlando & Mullen, 2000; Nist & Holschuh, 2000; Oxford, 1990; Richards, et al., 1992).

In this study, it is interesting to find females use more of these notetaking strategies than males. Females have been claimed as better academic achievers due to a desire for good grades and a social approval of their academic success (Oxford & Nyikos, 1989; Yang, 1994). Previous literature has indicated that notetaking strategies promote reading performance (e.g., Ho, 2007; Hu, 2007; Nist & Holschuh, 2000; Pintrich, 1999). These strategies are also noted in Mokhtari and Sheorey (2002) as support strategies to aid readers in comprehending the textual information. Hung (2005) specifically indicated that those who were more extrinsically motivated in learning would employ more support strategies. Since female learners have been claimed more motivated by their academic achievement and the public recognition of such achievement, it might help explain the significant gender effect found in the study on the use of these individual strategies. It calls for more research on the notetaking strategies in relation to sex differences to confirm this gender effect.

Relationships between Type of Strategy use and Overall Strategic Behavior. In the present study, significant and positive correlations were found between overall and individual types of strategy use, which confirmed what had been predicted earlier. Particularly, metacognitive and cognitive strategies had stronger associations with overall reading strategic behavior than the other strategy types. This suggested the two strategy types best predicted college students' reading strategy use and were the most representative and influential factors on the choice of strategy. The strategy preference over these types was also noted in Chiu (2007), Hsu (2007), Hu (2007), Sheorey and Mokhtari (2001), and Wang (2006). This finding did not come up with a surprise in the study because students mainly counted on metacognitive and cognitive strategies to carry out language learning tasks (Nist & Holschuh, 2000; Richards, et al., 1992). In fact, metacognitive strategies are techniques utilized to help learners control their cognition and coordinate the learning process. Planning, monitoring, and evaluating are common behaviors for these strategies (Hu, 2007; Sheorey & Mokhtari, 2001; Wang, 2006). Wang (2006) specifically indicated that metacognitive reading strategy instruction successfully promoted EFL learners' reading comprehension and their awareness of strategy use. The role of metacognitive behaviors in enhancing reading comprehension should never be underestimated.

On the other hand, cognitive strategies are typically the most popular type of strategy with language users (Oxford, 1990). They enable learners to comprehend and produce new language by different means, such as summarizing or reasoning deductively (Oxford, 1990; Sheorey & Mokhtari, 2001; Wang, 2006). In addition to the global perspective of cognitive strategy use, culture plays a role in this strategy preference. As noted in Hu (2007), senior high school students in Taiwan faced much pressure from school tests, so they tended to use cognitive strategies in their studies to facilitate the testing process. Specific instances detected in Hu's study were that these students recognized their prior knowledge, reread, and adjusted their reading pace to comprehend a text. They also noticed the nature and requirement of a test and employed a skimming strategy to control their time well. Generally, the prevalent use of cognitive strategies is greatly linked to the test-oriented education in Taiwan (Hu, 2007). Such strategy popularity characterizes the local culture of educational beliefs and practices.

PEDAGOGICAL IMPLICATIONS

Use of strategies facilitates reading comprehension and the ultimate goal is to attain learner autonomy (Grabe & Stoller, 2002; Rubin & Thompson, 1994). The present study found that college EFL students employed reading strategies in a moderate degree; in particular, social-affective strategies were least used among all types. It is thus recommended that intensive strategy instruction be introduced to the regular reading classroom to induce more strategic behaviors in reading. In addition, the approaches of collaborative strategic reading (Klingner & Vaughn, 1998) and communicative language instruction that promote classroom interactions between the teacher and students or within the students themselves deserve more of our attention. These measures should be integrated into the strategy training process to elicit more strategic behaviors in the social-affective domain. Generally, positive self-rewarding, cooperating with peers and knowing more cultural contexts are typical examples of social-affective strategies. Language learning is a dynamic process which includes two main components: linguistic knowledge and the social and cultural expectations associated with the context. Compared to the conventional knowledge-based instruction in Taiwan, the strategy training with an emphasis on interpersonal skills and self-efficacy strengthens comprehension, helps overcome difficulties, and more importantly, prepare learners to be strategic readers when they employ such techniques to assist them in reading (Chen & Huang, 2007; Huang, et. al., 2006; Poole, 2005;

Richards, et al., 1992; Shang, 2007; Yeh, 2006).

In addition, significant proficiency and gender effects were not found in overall and specific types of strategy use, but in the use of individual strategies. Effective strategies such as (1) read-aloud, (2) skimming, (3) guessing from the context, and (4) analyzing sentence patterns and grammar structures, all of which discriminated the proficient and less proficient readers, are advised to be taught to low-achievers in the remedial instruction to promote their English reading proficiency. EFL teachers should also be aware of the gender differences on male and female students in their use of notetaking strategies. They should notice the triangular relationship between the use of reading strategies and the two concerned variables (Ho, 2007).

Furthermore, the finding that metacognitive and cognitive strategies were the best predictors of overall strategy use also yields pedagogical significance. Since these two types are most related to the general strategic behaviors of reading, they should be the main focus of the suggested strategy training. Generally, learning strategies are readily teachable (Oxford & Nyikos, 1989). It is the EFL teachers that are responsible for building such scaffolding on strategy use, so their students can eventually employ strategies independently (Nist & Holschuh, 2000). After all, becoming constructively responsive in reading is powerful in enhancing academic achievement (Sheorey & Mokhtari, 2001).

SUGGESTIONS FOR FURTHER RESEARCH

There are limitations in the study, and suggestions for further research on reading strategies are provided. First, the participants recruited for this study were technological English majors in southern Taiwan. The generalization of the findings should be restricted to the types of the learners of homogeneous backgrounds. Further studies interested in using similar participants may increase the surveying sample size, particularly that of male learners, to compare with the gender effect found in the study. Second, this study focused on probing two variables (i.e., English proficiency and gender differences) in the context of perceived reading strategy use. The measure of English learning motivation could be added to explore the effects of different motivational traits on the reading strategy use, as what was noted in Hung (2005). Third, as indicated in Yang (1994), the language-majored students employed strategies in a more versatile and frequent manner than their counterparts of other academic specializations. A deeper exploration into the differences of strategy use between English and non-English majors would bring more inspiring results. Additionally, Hsu (2003) indicated that regular college students and technological college students employed different reading strategies. The aspects that participants from different majors or educational systems deserve more of our attention since a limited number of strategy studies of this sort have been documented to date in the regional context of EFL. Fourth, in this study, a survey questionnaire was constructed to detect the subjects' strategic traits. Future studies may triangulate the methodology and employ additional qualitative measures such as think-aloud method, post-questionnaire interviews, and students' reflective journals to discern the processes involved in perceived strategy knowledge and learners' actual choices of reading strategies. Lastly, additional strategy studies may shift the research attention to the effect of reading strategy training. Factors such as prior knowledge, reading motivation and complexity of reading task are worth probing since they are directly linked to reading behaviors.

REFERENCES

- Brown, H. D. (2001). *Teaching by principles: An interactive approach to language pedagogy* (2nd ed.). NY: Longman.
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*.

- White Plains, NY: Pearson Education.
- Caverly, D. C., Orlando, V. P., & Mullen, J. L. (2000). Textbook study reading. In R. F. Flippo, & D. C. Caverly (Eds.), *Handbook of college reading and study strategy research* (pp. 105-148). Mahwah, NJ: Erlbaum.
- Chamot, A. U. (2004). Issues in Language Learning Strategy Research and Teaching. *Electronic Journal of Foreign Language Teaching*, 1(1), 14-26.
- Chang, S. J. (1992). A study of language learning strategies that overseas Chinese students in the United States employed and the relation of their strategy use to oral proficiency and other factors. *Proceedings of the Eighth International Conference on English Teaching and Learning in the Republic of China* (pp. 443-70). Taipei: The Crane Publishing Co.
- Chang, Y. J. (1997). *A study of English learning motivation and strategies in relation to the English learning achievement of junior high school students in Taiwan*. Unpublished master's thesis, National Kaohsiung Normal University, Taiwan.
- Chao, J. A. (2006). *Effects of reading vocabulary aloud on the English learning and attitudes of EFL elementary school low achievers*. Unpublished master's thesis, National Cheng Kung University, Taiwan.
- Chen, I. J., & Huang, S. H. (2007). High-intermediate English proficiency learners' reading strategy use in comprehending different passages: A case study. *Proceedings of the 24th International Conference on English Teaching and Learning in the Republic of China* (pp. 443-454). Taipei: Taiwan ELT Publishing Company.
- Chen, J. C. (2005). *Explicit instruction of reading strategies at senior high school in Taiwan*. Unpublished master's thesis, National Kaohsiung Normal University, Taiwan.
- Chiu, H. X. (2007). *Reading strategies used by high and low English proficiency students studying in the department of applied English*. Unpublished master's thesis, Southern Taiwan University, Taiwan.
- Ehrman, M., & Oxford, R. L. (1988). Effects of sex differences, career choice, and psychological type on adult language learning strategies. *Modern Language Journal*, 72(3), 253-265.
- Grabe, W., & Stoller, F. L. (2002). *Teaching and researching reading*. Harlow: Longman.
- Green, J. M., & Oxford, R. L. (1995). A closer look at learning strategies, L2 proficiency, and gender. *TESOL Quarterly*, 29(2), 261-295.
- Ho, P. Y. (2007). *A study of EFL reading strategies used by vocational high school students in Taiwan*. Unpublished master's thesis, National Taiwan University of Science and Technology, Taiwan.
- Hsiao, T. Y., & Oxford, R. L. (2002). Comparing theories of language learning strategies: A confirmatory factor analysis. *The Modern Language Journal*, 86(3), 368-383.
- Hsu, L. R. (2003). *A study of relationships between feeling of knowing about English reading strategy use and reading comprehension of Taiwanese college students*. Unpublished master's thesis, National Changhua University of Education, Taiwan.
- Hsu, S. C. (2007). Reading strategies used by EFL technical students. *English Teaching/Learning Resources Center in Northern Taiwan*. Retrieved May 5, 2008, from http://www.etlc.ntust.edu.tw/tepaper/epaper_m22.htm
- Hu, C. C. (2007). *A study on the differences between Chinese and English reading strategies, and their relation to the reading comprehension abilities for senior high school students*. Unpublished master's thesis, Tamkang University, Taiwan.
- Hu, M. C. (2006). *Hidden transfer: Indicators of English reading performance among*

- Taiwanese students*. Unpublished master's thesis, Da-Yeh University, Taiwan.
- Huang, C. W. (2007). *The differential viewpoints on reading difficulties and metacognitive strategies application toward academic English reading between non-English major graduates and instructors*. Unpublished master's thesis, National Pingtung Institute of Commerce, Taiwan.
- Huang, H. C., Chern, C. L., & Lin, C. C. (2006). EFL learners' online reading strategies: A comparison between high and low EFL proficient readers. *English Teaching & Learning*, 30(4S), 1-22.
- Huang, S. C. (2005). *A study of factors affecting applied English majors' use of language learning strategies*. Unpublished master's thesis, Southern Taiwan University, Taiwan.
- Hung, H. C. (2001). *The effects of sex and language material styles on reading strategies of senior high school students in Taiwan*. Unpublished master's thesis, National Kaohsiung Normal University, Taiwan.
- Hung, H. P. (2005). *An investigation of factors that influence EFL college students' reading strategy use*. Unpublished master's thesis, National Changhua University of Education, Taiwan.
- Hung, H., Tsou, W., & Wu, K. Y. (2005). The effect of reading strategy training on freshman English reading in Taiwan. *Journal of National University Tainan: Humanities & Social Science*, 39(1), 79-106.
- King, J. (2002). *Ying yu chiao hsueh yu hsueh sheng jen shou pu chueh ting chi tsai tse tse lueh chih hsiang kuan hsing [The correlations with English teaching and tolerance of ambiguity and guessing strategy among students]*. Retrieved December 22, 2008, from <http://ejee.ncu.edu.tw/teacherarticle/9-3scujane.htm>
- Klingner, J. K., & Vaughn, S. (1998). Using Collaborative Strategic Reading. *Teaching Exceptional Children*, 30(6), 32-37.
- Kuo, L. H. (2004). *The correlation between Chinese reading and English reading of senior high EFL students in Hsinchu and Miaoli areas*. Unpublished master's thesis, National Kaohsiung Normal University, Taiwan.
- Kuo, W. K. (2003). *The proficiency and gender differences in reading strategies used by junior high school students toward the Basic Competence Test*. Unpublished master's thesis, Providence University, Taiwan.
- Liao, Y. F. (2000). *A study of Taiwanese junior high school students' EFL learning motivation and learning strategies*. Unpublished master's thesis, National Changhua University of Education, Taiwan.
- Lin, T. C. (2006). *EFL technological university students' use of language learning strategies and beliefs about language learning*. Unpublished master's thesis, Southern Taiwan University, Taiwan.
- Liu, M. L. (2004). *A study of beliefs about language learning and language learning strategies of senior high school students*. Unpublished master's thesis, National Kaohsiung Normal University, Taiwan.
- Lo, M. L. (2007). *A case study of elementary school students learning to read and write with English informational books*. Unpublished master's thesis, Tamkang University, Taiwan.
- Mokhtari, K., & Sheorey, R. (2002). Measuring ESL students' awareness of reading strategies. *Journal of Developmental Education*, 25, 2-10.
- Nist, S. L., & Holschuh, J. L. (2000). Comprehension strategies at the college level. In R. F. Flippo, & D. C. Caverly (Eds.), *Handbook of college reading and study strategy research* (pp.75-104). Mahwah, NJ: Erlbaum.
- O'Malley, M., & Chamot, A. U. (1990). *Learning strategies in second language*

- acquisition*. New York: Cambridge University Press.
- Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. Boston: Heinle & Heinle.
- Oxford, R. L. (1993). Language learning strategies in a nutshell: Update and ESL suggestions. In J. C. Richards, & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 124-132). New York: Cambridge University Press.
- Oxford, R.L., & Burry-Stock, J.A. (1995). Assessing the use of language learning strategies worldwide with the ESL/EFL version of the Strategy Inventory for Language Learning (SILL). *System*, 23(1), 1-23.
- Oxford, R. L., & Nyikos, M. (1989). Variables affecting choice of language learning strategies by university students. *Modern Language Journal*, 73(3), 291-300.
- Peng, I. N. (2002). *EFL motivation and strategy use among Taiwanese senior high school learners*. Unpublished master's thesis, National Taiwan Normal University, Taiwan.
- Phakiti, A. (2003). A closer look at gender and strategy use in L2 reading. *Language Learning*, 53(4), 649-702.
- Pintrich, P. R. (1999). The role of motivation in promoting and sustaining self-regulated learning, *International Journal of Educational Research*, 31, 459-470.
- Poole, A. (2005). Gender differences in reading strategy use among advanced ESL college students. *Journal of College Reading and Learning*, 36(1), 7-20.
- Ray, M. W., Garavalia, L. S., & Gredler, M. E. (2003). Gender differences in self-regulated learning, task value, and achievement in developmental college students. *Paper presented at the Annual Meeting of the American Educational Research Association*. Abstract retrieved ERIC database.
- Richards, J., Platt, J., & Platt, H. (1992). *Dictionary of language teaching and applied linguistics*. UK: Longman.
- Rubin, J. (1987). Learner strategies: Theoretical assumptions, research history and typology. In A. Wenden, & J. Rubin (Eds.), *Learner strategies in language learning* (pp. 15-30). Englewood Cliffs, NJ: Prentice Hall.
- Rubin, J., & Thompson, I. (1994). *How to be a more successful language learner: Toward learner autonomy* (2nd ed.). Boston: Heinle & Heinle.
- Shang, H. F. (2007). Reading strategy training for the development of Chinese EFL reading comprehension. *Proceedings of the 24th International Conference on English Teaching and Learning in the Republic of China* (pp. 424-442). Taipei: Taiwan ELT Publishing Company.
- Shen, Y. S. (2003). *A study of the effects of two text structures on Taiwanese EFL junior high school students' strategy use*. Unpublished master's thesis, National Cheng Kung University, Taiwan.
- Sheorey, R., & Mokhtari, K. (2001). Differences in the metacognitive awareness of reading strategies among native and non-native readers. *System*, 29, 431-449.
- Su, M. H. (2005). A study of EFL technological and vocational college students' language learning strategies and their self-perceived English proficiency. *Electronic Journal of Foreign Language Teaching*, 2(1), 44-56.
- Sy, B. M. (1995). Gender differences, perceptions on foreign language learning and language learning strategies. *Proceedings of the twelfth International Conference on English Teaching and Learning in the Republic of China* (pp. 215-229). Taipei: The Crane Publishing Co.
- Sy, B. M., & Liou, S. C. (1994). Sex differences and language learning strategies.

- Proceedings of the Eleventh International Conference on English Teaching and Learning in the Republic of China* (pp.19-51). Taipei: The Crane Publishing Co.
- Taillefer, G., & Pugh, T. (1998). Strategies for professional reading in L1 and L2. *Journal of Research in Reading*, 21(2), 96-108.
- Tsai, L. Y. (2005). *The relationships among language learning strategies, teaching strategies, and English achievement*. Unpublished master's thesis, Southern Taiwan University, Taiwan.
- Wang, M. J. (2006). *Studies of web-based metacognitive reading strategy training in English reading and perceptions of web-based learning for senior high school students in southern Taiwan*. Unpublished doctoral dissertation, National Kaohsiung Normal University, Taiwan.
- Wharton, G. (2000). Language learning strategy use of bilingual foreign language learners in Singapore. *Language Learning*, 50(2), 203-243.
- Wenden, A. L. (1987). Conceptual background and utility. In A. Wenden, & J. Rubin (Eds.), *Learner strategies in language learning* (pp. 3-13). Englewood Cliffs, NJ: Prentice Hall.
- Wu, C. C. (2003). *A study of EFL learning strategies used by vocational high school students in Taiwan*. Unpublished master's thesis, National Yunlin University of Science and Technology, Taiwan.
- Wu, Y. W. (2007). *Writing strategies and writing difficulties among college students of differing English proficiency*. Unpublished master's thesis, National Pingtung Institute of Commerce, Taiwan.
- Yang, N. D. (1994). A study of factors affecting college EFL students' use of learning strategies. *Proceedings of the Eleventh International Conference on English Teaching and Learning in the Republic of China* (pp. 53-82). Taipei: The Crane Publishing Co.
- Yeh, I. C. (2006). *The effects of reading strategies instruction on junior high school students' reading comprehension in Taiwan*. Unpublished master's thesis, National Chengchi University, Taiwan.
- Yu, H. C. (2006). *Motivation and learning strategy use among junior high school students with different levels of academic achievement*. Unpublished master's thesis, National Pingtung Institute of Commerce, Taiwan.

Use of Vocabulary Strategies in Relation to Major and Gender

Mei-chen Chen (陳美貞), Sin-yi Huang (黃馨儀)

National Pingtung Institute of Commerce

Meichen@npic.edu.tw
pennylovelu@hotmail.com

Learning strategies facilitate the language learning process, and their ultimate goal is to achieve learner autonomy. Although tons of studies have probed general language learning strategies over the past decades, specific tasks on vocabulary learning have not received an equivalent amount of attention from the strategy research. Even so, not much research on vocabulary strategy use has explored major and gender differences in EFL settings and the results have, unfortunately, been mixed. Therefore, this study aimed to examine the frequency and type of vocabulary strategy use among Chinese EFL learners. The effects of major and gender were further probed in the context of strategy use. One hundred and eighty-seven college students were surveyed with a standardized questionnaire of vocabulary strategies. Important findings were as follows: (a) EFL college students used vocabulary strategies in a moderate degree, (b) determination strategies were reported as the most used type, and meta-cognitive and social strategies as the least used types, (c) no significant differences were found between English and non-English majors in their overall and specific types of vocabulary strategy use, (d) significant major effects were only detected in individual strategies in favor of English majors, (e) although no significant difference was detected on overall strategy use between male and female learners, females significantly employed more determination strategies than males, and males significantly used more meta-cognitive strategies than their female counterparts, (f) relatively, gender was a more potent factor than major to influence college students' choice of vocabulary strategies, and (g) memory and cognitive strategies were found to be most correlated with the overall strategy use; these two types best predicted adults' strategic behaviors of vocabulary learning. Based on these findings, pedagogical and research implications were provided to conclude this study.

INTRODUCTION

Vocabulary acquisition is a life-long and infinite learning task. Nowadays, vocabulary learning is no longer regarded as a neglected, discrete component, but a fundamental, indispensable part in native or foreign language learning (Barcroft, 2004; Chiang, 2005; Chu, 2008; Ghazal, 2007; Gu & Johnson, 1996; Lin, 2003; Nation, 2001; Schmitt, 2000; Wei, 2007; Zimmerman, 2000). It is especially marked by its close link to the acquisition of basic language skills. Vocabulary knowledge has been found as an effective predictor of language learning proficiency, particularly in relation to reading comprehension (e.g., Boote, 2006; Gu & Johnson, 1996; Khaldieh, 2001; Li, 2004; Lin, 2003; Min & Hsu, 2008). Previous studies have also claimed vocabulary richness as vital for the development of writing (e.g., Smith, 2003; Zimmerman, 2004). Goulden, Nation, and Read (1990) further indicated that vocabulary size was crucial in determining the quality, accuracy, and descriptiveness of writing a piece of composition. Moreover, successful listening performance relies heavily on lexical knowledge. Comprehension in listening could be impeded due to limited word knowledge (Goh, 1999; Miao, 2006; Rubin, 1994). Generally, without adequate vocabulary knowledge, it is difficult to develop other language skills.

In recent years, there has been a growing research interest in learners' beliefs about the

nature of vocabulary learning and the efficacy of the strategies they use (e.g., Catalán, 2003; Lee, 2007; Liao, 2004; Gu, 2002; Wei, 2007). Previous literature has indicated that use of strategies significantly promotes vocabulary learning (Brown & Perry, 1991; Chu, 2008; Green & Oxford, 1995; Lawson & Hogben, 1998; Yek, 2006). According to Schmitt (1997), vocabulary learning strategies (VLS) were the techniques used to determine the meaning of a new word or retain a known word in memory. Catalán (2003) defined such strategies as an intention or action to figure out the meanings of unknown words, to retain these words in the long-term memory, to recall them, and eventually to utilize the learned words in an oral or written form. Note that effective vocabulary learning can be achieved by introducing learners some facilitative instructions of vocabulary strategies (Brown & Perry, 1991; Chu, 2008; Liao, 2004; Mercer, 2005; Nation, 2001; Oxford & Scarcella, 1994; Schmitt, 2000). Moreover, strategies help create a context of learner autonomy (Cohen, 1998; Nation, 1990; Oxford, 1990; Oxford and Nyikos, 1989). Independent learning is a key to academic success. Identifying what and how vocabulary learning strategies are actually employed by successful learners in their learning process and to what extent such strategies are used is of paramount importance since it provides an insightful guidance for developing a strategy training program tailored to vocabulary learning. The present study was thus set to probe the frequency and type of vocabulary strategy use by more proficient EFL learners (i.e., college English majors) in contrast to the strategies used by those from other academic specializations. It was in the hope that the results would aid the development of a more elaborate curriculum concerning vocabulary instruction in a regular college reading course.

Schmitt (1997) developed a vocabulary learning strategy taxonomy encompassing five strategy categories: determination, social, memory, cognitive, and meta-cognitive. Determination strategies (DET) are the kind of strategies learners utilize to sort out the meaning of a new word without consulting other people. Social strategies (SOC) involve interactions with other people to acquire the language. Approaches which relate new information to learners' existing word knowledge are viewed as memory strategies (MEM). Cognitive strategies (COG) entail learners' mental processing ability to manipulate or transform the target language. Finally, meta-cognitive strategies (MET) refer to learners' conscious awareness of planning for, evaluating, and monitoring of their entire vocabulary learning process. This taxonomy was projected in a concrete questionnaire with a comprehensive list of vocabulary learning strategies geared to the five types. Several VLS studies have been conducted with the inventory developed by Schmitt (e.g., Catalán, 2003; Chen & Yeh, 2004; Liao, 2004; Runtmets, 2005; Sung, 2006; Wang, 2004; Wu, 2005). This VLS inventory was also tested in the present study due to its clearly-defined categories with manageable questionnaire items and its suitability to be administered in the EFL context of Taiwan.

Although over the past two decades ample studies have explored general language learning strategies (e.g., Chang, Liu, & Lee, 2007; Ehrman & Oxford, 1989; Griffiths, 2003; O'Malley, Russo, Chamot, Stewner-Manzanares, & Kupper, 1985; Oxford & Nyikos, 1989; Oxford, Nyikos, & Ehrman, 1988; Riazi, 2007), specific tasks on vocabulary acquisition have not received an equivalent amount of attention from the strategy research. Even so, much research of this sort only examined a small number of vocabulary learning strategies with a focus on mnemonic techniques (Schmitt & Schmitt, 1993). One reason for a noticeable lack of attention on vocabulary strategies was due to the shortage of comprehensive research on strategy use in this specific area (Schmitt, 1997). Therefore, the identification of such strategies was very necessary so as to deal with the deficiency of previous studies.

In addition, much attention has been directed toward the identifications of vocabulary strategies in relation to differing English proficiency (Fan, 2003). Namely, most of the recent strategy studies have been conducted to distinguish good and poor language learners of

homogeneous academic backgrounds in their use of vocabulary strategies (e.g. Fan, 2003; Gu, 1994; Li, 2004; Liu, 2006; Wang, 2004; Wei, 2007). Relatively, not that many studies of this sort have explored major and gender differences in EFL contexts (e.g., Catalán, 2003; Lee, 2007; Liao, 2004; Gu, 2002; Wei, 2007), and the results have, unfortunately, been mixed. Liao (2004) and Wei (2007) found English majors significantly outperformed their non-English-major counterparts in the overall VLS use; however, such significant major effect was not detected in Gu (2002). One inconsistent finding concerning the gender effect was that Catalán (2003) found Spanish-speaking females were more strategic than males over a great range of vocabulary strategies; in contrast, Lee (2007) reported no sex differences in Korean university students' use of VLS. Such inconsistencies call for more research, so that firm conclusions could be reached about the effects of major and gender on VLS use.

In sum, this study aimed to probe major and gender differences in vocabulary strategy use among college EFL learners of technological orientation. There were four specific purposes: (a) to detect the general use of VLS; (b) to examine major differences in the strategy use; (c) to determine the gender effect in the strategy use; (d) to explore the relationship between type of strategy use and overall VLS behavior.

Based on the purposes of the study, four predictions were made. First, since the previous literature consistently indicated that EFL learners were moderate strategy users (e.g., Chang, 2005; Chang, et al, 2007; Gooch & Foong, 1997; Ruutemets, 2005; Su, 2005; Wei, 2007), it was thus predicted that the college participants would not use VLS frequently in studying English. When the major effect was concerned, English majors, generally assumed as more proficient language learners, were expected to use greater strategies than their non-English-major counterparts (e.g., Chang, et al, 2007; Huang, 2008; Liao, 2004; Wei, 2007). As for the gender effect, females were hypothesized to be more strategic than their male counterparts as female superiority in strategy use was generally reported in the previous literature (e.g., Chang, et al, 2007; Ehrman & Oxford, 1989; Goh & Kwah, 1997; Liu, 2004; Nyikos, 1990; Oxford & Nyikos, 1989; Politzer, 1983). Lastly, memory and cognitive strategies were anticipated to have the strongest correlations with overall VLS use because vocabulary acquisition involved various information-processing measures to retain the learned information in the long-term memory (e.g., Mohammad, 2008; Johnson & Obi, 1993; Yek, 2006). It was hoped that the findings would acquaint EFL teachers with the approaches students commonly used in learning or retaining vocabulary, so strategic vocabulary instruction in response to student needs could be intervened to maximize their learning experiences.

LITERATURE REVIEW

The domain of language learning strategies has received great research attention since mid-1970s. Strategies are referred to techniques, steps, or behaviors taken by learners to facilitate the process of learning (Oxford & Crookall, 1989; Oxford & Nyikos, 1989). Many empirical studies have confirmed that use of appropriate learning strategies has positive effect on the acquisition of second/foreign languages (e.g., Ehrman & Oxford, 1989; Oxford & Crookall, 1989; Oxford & Nyikos, 1989). Oxford and Nyikos (1989) raised several important variables pertinent to learners' use of strategies. These influential factors were language proficiency, field of specialization, gender, attitude, motivation, learning style, and cultural background. Among these factors, the variable of gender/sex differences has received great attention. Chang, et al. (2007), Ehrman and Oxford (1989), Green and Oxford (1995), Griffiths (2003), Oxford and Nyikos (1989), and Politzer (1983) reported that females use language learning strategies significantly greater than males did. Particularly, females were found more social-strategy-orientated learners than their male counterparts in language learning. This phenomenon could be well-explained by women's longing for good

grades and social approval (Oxford & Nyikos, 1989). Previous literature also has revealed that learners majoring in humanities/social science use strategies more frequently than other-major counterparts (e.g., Chang, et al., 2007; Oxford & Nyikos, 1989). Relatively, fewer studies have been conducted to probe strategy use in the specific area of vocabulary learning; thus, the effects of gender and field of specialization on the selections of vocabulary strategies have remained undetermined. Therefore, the current study intended to explore both variables in relation to learners' vocabulary strategy use in the context of EFL.

A considerable body of research about vocabulary learning strategies has been conducted in different nationality contexts, such as in China (Gu, 1994; Gu, 2002; Gu & Johnson, 1996; Wei, 2007), in Japan (Schmitt, 1997), in Taiwan (Chiang, 2005; Li, 2004; Liao, 2004; Wang, 2004; Wu, 2005), in Hong Kong (Fan, 2003), in Korea (Lee, 2007), and in Spain (Catalán, 2003). The strategy-related factors being studied in these studies include: gender (Catalán, 2003; Gu, 2002; Lee, 2007; Liao, 2004), major (Gu, 2002; Liao, 2004; Wei, 2007), vocabulary proficiency (Gu, 2002; Gu & Johnson, 1996; Lee, 2007; Li, 2004; Wang, 2004), language proficiency (Gu, 1994; Gu, 2002; Gu & Johnson, 1996; Wei, 2007), and age group (Schmitt, 1997; Wu, 2005). In order to understand how words are acquired by learners, the development of a clear set of vocabulary learning strategies is very necessary (Schmitt, 1997). To date, researchers have proposed taxonomies with different strategy categories about vocabulary learning (e.g., Fan, 2003; Gu & Johnson, 1996; Schmitt, 1997). Gu and Johnson (1996) developed an extensive taxonomy of vocabulary learning strategies with two contrastive types: cognitive and meta-cognitive. Fan (2003) raised another VLS taxonomy with 60 strategy items geared to nine categories. They were management, sources, dictionary, guessing, four memory-related strategies (analysis, grouping, repetition, and association), and consolidating the knowledge of known words which dealt with reviewing, practicing, or learning new usage of acquired words. The results of surveying on 1067 Cantonese-speaking English learners showed that consolidating the knowledge of known words, and dictionary strategies were frequently-used, effective strategies. In addition, Schmitt (1997) initiated a comprehensive taxonomy of VLS with 5 types: Determination, Social, Memory, Cognitive, and Meta-cognitive, as mentioned previously. The present study mainly concerned about the inventory developed by Schmitt.

Major differences could be a decisive factor in learners' choice of strategies. Oxford and Nyikos (1989) discovered that Social Science/Humanities/Education majors employed functional practice strategies more frequently than did their technical counterparts. Such result coincided with what was reported in Chang, et al. (2007) that Chinese EFL learners majoring in Humanities and Social Science significantly used more language learning strategies than those majoring in Science and Engineering as well as in Business and Management. One plausible explanation for this result might be that the majority of Humanities and Social Science was females who were commonly regarded as the greater strategy users than males (Chang, et al., 2007). Liao (2004) probed use of vocabulary learning strategies on English-major and non-English-major university freshmen in Taiwan. Determination strategies were found to be used most often, and meta-cognitive and social strategies, least often for all learners. A significant major effect was produced with the strategy superiority found in the English majors. However, Gu (2002) reported no significant major effect on college EFL learners' use of VLS.

To date, very few studies have examined gender differences in the context of vocabulary strategy use (e.g., Catalán, 2003; Gu, 2002; Lee, 2007; Liao, 2004). Unfortunately, with a closer examination into these studies, different results have manifested on the sex effect. Catalán (2003) probed whether sex differences would affect learners' preference of VLS with Spanish-speaking students ranging from 11 to 56 of age. The results indicated that females were more strategic than males over a greater range of strategies. A similar female

superiority of VLS was also detected in Gu (2002) and Liao (2004) with Chinese EFL learners. In contrast, the results of Lee (2007) contradicted those of the aforementioned studies in that no gender differences were exhibited in the use of VLS among Korean university students. Note that most of these studies (i.e., Catalán, 2003; Lee, 2007; Liao, 2004) conducted a survey research with the VLS questionnaire developed by Schmitt (1997), but the results were quite mixed. With the limited literature, the distinctions between males and females on their use of VLS still remained uncertain.

In sum, the number of empirical studies on VLS in relation to gender and major effects has not been sufficient enough to allow a consistent pattern of vocabulary learning behavior to be drawn from. Although a substantial body of research on VLS has been conducted in EFL settings (e.g., Chen & Yeh, 2004; Chiang, 2005; Fan, 2003; Gu & Johnson, 1996; Schmitt, 1997; Wang, 2004; Wei, 2007; Wu, 2005), little of which explored the effects of these two factors on learners' choice of VLS (e.g., Catalán, 2003; Gu, 2002; Lee, 2007; Liao, 2004; Wei, 2007). Moreover, certain studies such as Chiang (2005) and Wu (2005) analyzed learners' strategy differences from individual VLS items in which it was hard to see a general pattern of VLS use. Hence, the current research intended to examine major and sex differences with an emphasis on exploring systematic VLS types. It was in the hope that the study could provide a general but clear picture on EFL learners' use of VLS to add credence to the deficiency of the current literature.

METHODOLOGY

The current study attempted to explore vocabulary learning strategies used by Chinese EFL learners of differing field specializations. Basically, it was a quantitatively-oriented survey research. The questionnaire employed in this study was adapted from that developed by Schmitt (1997), which was a standardized test with a comprehensive categorization and was used in several strategy studies (e.g., Catalán, 2003; Chen & Yeh, 2004; Liao, 2004; Sung, 2006; Wang, 2004; Wu, 2005). A surveying approach is a common practice in the strategy research in discerning the perceived use of strategies among ESL/EFL learners and is distinguished by its non-threatening and cost-effective nature (Oxford, 1990). The five types of vocabulary strategies to be probed were determination, social, memory, cognitive, and meta-cognitive. The effects of academic major and gender would be examined in the context of frequency and type of strategy use. The subjects, instruments, procedures of data collection, and data analyses are described in the following.

Subjects

The participants of the study, 187 college sophomores, were drawn from 4 intact classes of two technological colleges in southern Taiwan, two from the department of Applied Foreign Languages, and one from International Trade, and one from International Business. Specifically, there were 100 English-major students (M= 15, F= 85), and 87 non-English-major students (M= 10, F= 77). The non-English majors were fundamentally business-oriented. Due to the nature of field of specialization, the number of females was relatively greater than that of males in each of the English and business groups.

Instrument

A questionnaire of vocabulary strategies adapted from Schmitt (1997) served as the key instrument to collect data regarding vocabulary learning patterns among the college participants. These strategies were originally derived from various sources and with references to vocabulary textbooks, students' reports on the way they approached vocabulary learning, teachers' observations on students' vocabulary learning, etc. To tailor to the nature and intention of the current study, some items of vocabulary learning strategies were eliminated from the Schmitt's inventory because of different cultural-educational contexts in learning vocabulary. The number of the questionnaire items was reduced from the original

58 to the current 51. A Chinese version of the adapted questionnaire was developed for the survey to avoid any potential comprehension problems due to a deficiency in English knowledge. Both of the English and Chinese versions were then reviewed by one college professor specializing in learning strategy research and several other experienced English teachers to ensure the content validity of the questionnaires. A pilot study of the revised questionnaire on 5 college students of homogeneous backgrounds as the target learners was conducted to ascertain that the test items were clearly-stated and well-conveyed in Chinese. Eventually, the Chinese version yielded an overall test reliability of Cronbach alpha as .93 on the 187 college participants, specifically, with determination (8 items) as .69, social (7 items) as .78, memory (22 items) as .89, cognitive (9 items) as .73, and meta-cognitive (5 items) as .62. Although the reliability coefficient of each type was relatively lower due to a limited number of items involved in each type of strategies, the overall test coefficient .93 suggested the questionnaire was quite reliable.

Basically, the questionnaire was composed of two main parts: demographic data and strategy items about vocabulary learning. The first part was designed to elicit the participants' background information in relation to major, gender, and student ID number. The second part included 51 test items featured by five different types of strategies. A five-point scale ranging from "Always" (5 points), "Frequently" (4 points), "Sometimes" (3 points), "Seldom" (2 points), to "Never" (1 point) was used to measure the frequency of occurrence of each of these strategic behaviors. In addition, some space was left at the end of the questionnaire to elicit suggestions or comments about vocabulary strategy use from the college participants. At the end of the questionnaire, a written request was made on the subjects to double check what they had done to ensure every item was well taken care of. Sample items of the five types of vocabulary learning strategies are provided in Table 1.

Table 1 *Sample Items for the Five Strategy Types*

Strategy Type	Sample Items
Determination	Guess from textual context Analyze part of speech Analyze affixes and roots
Social	Ask teacher for an L1 translation Ask classmates for meaning Discover new meaning through group work activity
Memory	Study word with a pictorial representation of its meaning Connect the word to a personal experience Connect the word to its synonyms and antonyms
Cognitive	Verbal repetition Written repetition Keep a vocabulary notebook
Meta-cognitive	Skip or pass new word Testing oneself with word tests Continue to study word over time

Procedures

Before the questionnaire was administered, a pretest was conducted. Permissions of conducting the research were also gained from the English teachers in charge of the four target classes. On the arranged surveying days, an orientation to the study was given in the first place. The students were informed about the purpose of the study and the accurate way to fill out the questionnaire. Moreover, they were reassured that their responses would not affect their course grades, and that all the collected data would be analyzed corporately. The individual information would be kept confidential. The students were encouraged to complete the questionnaire honestly and intuitively. Afterwards, while they were working on the questionnaire, sufficient assistance was provided. The researcher and her assistant walked around to take questions from the students. Each returned questionnaire was carefully checked on the spot to ensure each item was completed appropriately and no items were left undone. Eventually, each of the participants was rewarded with a small gift for his/her assistance in the study. The entire surveying in each class took about 25 minutes to complete.

Data Analyses

The data collected in the survey research were analyzed statistically. First, to probe the general use of vocabulary learning strategies among college learners, the descriptive statistics of means, standard deviations and frequencies were reported. Second, independent-samples t-tests were conducted to detect if there were significant differences between English and non-English majors in their overall and specific types of strategy use. Third, similar t-tests were executed to determine the gender effect on the strategy use. Lastly, Pearson product-moment correlation analyses were performed to examine the relationships between the individual types of strategy use and overall strategic behavior. The level of statistical significance for all of the t-tests was set at .05, and for the correlation tests, at .01.

RESULTS

General Use of Vocabulary Strategies. The overall mean of strategy use among all college participants was 3.05 (See Table 2). The statistic was close to the 3-point (sometimes) category in the frequency scale. The result indicated that EFL college students used vocabulary learning strategies in a rather moderate degree. In addition, determination (M=3.40) was found to be used most often, and meta-cognitive (M=2.93) and social (M=2.65), least often.

Table 2 *General Vocabulary Strategy Use Among College Students*

abulary Strategy Type	Mean	SD	N.	Rank
Determination	3.40	.57	187	1
Social	2.65	.62	187	5
Memory	3.08	.55	187	3
Cognitive	3.16	.55	187	2
Meta-cognitive	2.93	.60	187	4
Overall	3.05	.44	187	

Major Differences in Vocabulary Strategy Use. The overall VLS use between the English and non-English majors were 3.05 and 3.04 respectively, as indicated in Table 3.

The result of an independent-samples t-test for the major effect on the strategy use showed there was no significant difference on the two means ($t(185) = .14, p = .89 > .05$). That is, the two groups of learners exhibited an equivalent frequency in terms of overall strategy use. In addition, no significant mean difference was detected in each of the five types of strategies between English and non-English majors (See Table 3). Furthermore, the frequency rankings of VLS use for the two studied groups were found identical. Although major differences were not found in overall and specific types of strategy use, they were detected in the use of individual vocabulary strategies. In particular, English majors significantly used more (1) studying the sound of a word, (2) imaging word form, (3) analyzing part of speech, (4) using a bilingual dictionary, and (5) keeping a vocabulary book, than their non-English-major counterparts.

Table 3 *T-tests on VLS Use Between English and Non-English Majors*

Type	Major	English Majors			Non-English Majors			Independent T-test		
		Mean	SD	Rank	Mean	SD	Rank	t	df	Sig.
Determination		3.41	.53	1	3.39	.62	1	.19	185	.85
Social		2.64	.60	5	2.67	.64	5	-.34	185	.73
Memory		3.08	.49	3	3.07	.61	3	.14	185	.89
Cognitive		3.16	.51	2	3.17	.60	2	-.08	185	.94
Meta-cognitive		2.96	.57	4	2.90	.64	4	.64	185	.52
Overall		3.05	.40		3.04	.49		.14	185	.89

Gender Differences in Vocabulary Strategy Use. Although no significant mean difference was found on the overall VLS use between males and females ($t(185) = -.64, p = .52 > .05$), a closer analysis of the five different types of VLS revealed that two of them reached the statistically significant level (See Table 4). More precisely, the females significantly used more determination strategies ($t(185) = -2.09, p = .04 < .05$), and the males significantly employed more meta-cognitive strategies ($t(185) = 2.1, p = .04 < .05$) than their opposite gender. Moreover, the frequency rankings for the two gender groups were somewhat inconsistent in such a way that determination and meta-cognitive were ranked as the top two most frequently used types of VLS for the males, and determination and cognitive for the females. Generally, although male and female learners exhibited no significant difference in their overall strategy use, significant gender effects were detected in the use of specific types of vocabulary strategies as well as in the rankings of strategy use.

Table 4 *T-tests on VLS Use between Male and Female Learners*

Gender & Values Type	Males			Females			Independent T-test		
	Mean	SD	Rank	Mean	SD	Rank	t	df	Sig.
Determination	3.18	.65	1	3.44	.55	1	-2.09*	185	.04
Social	2.55	.72	5	2.67	.60	5	-.88	185	.38
Memory	3.07	.47	3	3.08	.56	3	-.12	185	.91
Cognitive	3.00	.47	4	3.19	.56	2	-1.6	185	.11
Meta-cognitive	3.17	.71	2	2.90	.58	4	2.1*	185	.04
Overall	2.99	.40		3.05	.45		-.64	185	.52

Note. * signifies significant gender effect on the type of VLS use ($p < .05$).

Relationship between Type of Strategy Use and Overall Strategic Behavior. The strongest correlations were detected between memory and overall strategy use ($r = .84$, $p = .00 < .05$), and between cognitive and overall strategy use ($r = .80$, $p = .00 < .05$), as indicated in Tables 5. Memory and cognitive strategies were most related to and best predicted EFL learners' use of vocabulary learning strategies.

Table 5 *Correlations between Each Type and Overall Strategy Use*

Types of VLS	Overall Strategy Use		
	r	Sig. (two-tailed)	N
Memory	.84*	.00	187
Cognitive	.80*	.00	187
Determination	.76*	.00	187
Meta-cognitive	.73*	.00	187
Social	.69*	.00	187

Note. * signifies a significant correlation ($p < .01$).

DISCUSSION

General Use of Vocabulary Strategies. The current study found that EFL college students were moderate users of vocabulary learning strategies. Such finding was consistent with those of Liao (2004) and Wei (2007), revealing a medium use of overall VLS among the same type of learners. It also echoed what was reported in the previous research that language learning strategies were not highly used among EFL learners (Chang, 2005; Chang, et al, 2007; Goh & Foong, 1997; Ruutemets, 2005; Su, 2005; Wei, 2007). Moreover, determination was found to be the most used type of VLS, and social and meta-cognitive strategies as the least used two types, which confirmed what was produced in Liao (2004). Determination strategies involve learners' will and persistence to figure out the meaning of a word. Guessing from the textual context and analyzing affixes and roots are two examples of such type. As a matter of fact, EFL learners view vocabulary learning as individual learning process in which they tend to figure out new words on their own rather than resorting to others (Liao, 2004). There is, therefore, no surprise that determination strategies were found more frequently used than social ones in learning vocabulary.

Furthermore, with regard to a lack of the use of social strategies, it might be greatly associated with the local educational contexts marked by the conventional teacher-centered, grammar-oriented instruction. Kudo (1999) precisely uncovered that social strategies were the least frequently used type of VLS since language learners did not learn vocabulary

collaboratively in class. With the tightly controlled instruction by the teacher, the students are generally not given many chances to learn from peer-interactions in class (Chou, 2007; Qin, 2007). As for the least use of meta-cognitive strategies, it might be due to learners' limited English knowledge for such type of VLS (Gu & Johnson, 1996). Meta-cognitive strategies, which involve planning for, monitoring, and evaluating the progress of language learning, are in fact a higher level of cognitive strategies (Gu & Johnson, 1996; Myers & Paris, 1978). They would surely demand more on foreign language learners. This helps explain why the college learners did not make good use of such type of strategies, particularly when they generally did not use vocabulary strategies often.

Major Differences in Vocabulary Strategy Use. Different from what had been predicted earlier, the present study found no significant mean differences in overall and specific types of VLS use between English and non-English majors. This result suggested there were generally no major differences in college EFL learners' use of vocabulary strategies. Although such result was consistent with what was reported in Gu (2002) with no significant major effect on college EFL learners' use of VLS, this piece of evidence contradicted those of previous learning strategy studies in that English majors used strategies more extensively than their non-English-major counterparts (e.g., Chang, et al, 2007; Huang, 2008). Particularly, empirical research on vocabulary strategies did discover English majors exceeded non-English majors in their reported overall strategy use (e.g., Liao, 2004; Wei, 2007). Since the study identified all the college learners as "sometimes" ($M=3.05$) strategy users in learning vocabulary, a mean in fact lower than what had been found in most strategy studies with an average strategy use around 3.50, this deficiency in the use of VLS might obscure or dilute the interested major effect in this study.

Another plausible explanation for such inconsistent findings with the former research literature involves a different composition of the studied subjects. Liao (2004) surveyed the VLS use of 625 EFL students from a university in Taiwan and Wei (2007) studied the same strategy use of 60 Beijing university students, whereas the present study chiefly drew 187 subjects from two technological colleges in South Taiwan. In addition, the non-English majors in this study were basically business-oriented from the departments of International Trade and International Business. The literature has consistently reported that there is a strong link between English proficiency and the frequency of vocabulary strategy use (e.g., Fan, 2003; Gu, 1994; Li, 2004; Liu, 2006; Wang, 2004; Wei, 2007). It could thus be inferred that the English and business groups approximated in their English proficiency (particularly, in technologically oriented colleges), so the discrepancy in the consequent strategy use was not that evident. And what's more, both major groups were characterized by female dominance. Females have consistently been identified as more strategic than males in the past literature (Nyikos, 1990; Oxford & Nyikos, 1989). This homogeneous sex trait in the subject composition of the studied groups might contribute to the insignificant major effect in this study. With the limited literature examining the effect of major on vocabulary strategy use, it calls for more research to clarify the inconsistent findings, in particular, in the context of EFL vocabulary learning. A sufficient research body of this sort would allow solid conclusions to be arrived at on the major effect.

Gender Differences in Vocabulary Strategy Use. Previous strategy literature has generally shown that females are better strategy users than males in language learning. The present study found that there was no significant gender effect on overall strategy use; male and female college learners used vocabulary strategies in an equivalent frequency. This result of no gender effect on VLS strategic behaviors conflicted those of studies on general language learning strategies (e.g., Chang, et al., 2007; Ehrman & Oxford, 1989; Goh & Foong, 1997; Liu, 2004; Nyikos, 1990; Oxford & Nyikos, 1989; Polizter, 1983), but it corresponded to those of research on more specific strategic reading habits (e.g., Hsu, 2006;

Phakiti, 2003; Poole, 2005; Sheorey & Mokhtari, 2001) with insignificant gender effect. This piece of evidence suggests gender is a more determinant factor on general language learning than on specific reading tasks. However, in the limited VLS studies on gender effect, Catalán (2003) and Liao (2004) reported that females significantly used more VLS than males. Different college orientations and age levels from the studied subjects might account for such inconsistencies between the two studies and this study. This study adopted technological EFL students, while Liao (2004) recruited EFL learners from a regular university, and Catalán (2003) surveyed on Spanish learners with varying age level ranging from 11 to 56.

Although significant gender effect was not found in overall strategy use, it was detected in certain types. This study found that female students significantly used more determination strategies, whereas male students resorted to meta-cognitive strategies more frequently in their vocabulary learning. Particularly, in dealing with new words, the former tended to find out the meaning of an unfamiliar word persistently by themselves rather than seeking for others' expertise; the latter were more likely to pass or skip an unknown word while reading. The sex differences relating to VLS choice were well-disclosed from this perspective. According to Oxford and Nyikos (1989) and Nyikos (1990), female learners are marked by a desire for academic success and a need for social recognition of their achievement, so there is a high chance that they would select strategies involving functional practice in authentic language use more frequently than male learners. Due to the limited documentation of VLS studies probing sex differences, more research is needed to confirm this message. It is also worth noting that, with the significant sex differences detected on these specific types, gender was relatively a more potent factor than major to influence the adult EFL learners' choice of vocabulary strategies.

Relationship between Type of Strategy Use and Overall Strategic Behavior. This study also found that memory and cognitive strategies were highly and most correlated with overall VLS use. This piece of evidence suggested these two types of strategies were the best predictors of adult learners' use of vocabulary strategies. They were also most related to the overall strategic behavior. Oxford (1990) defined memory strategies as "powerful mental tools which make learning easier, faster, more enjoyable, more self-directed, more effective, and more transferable to new situations" (p. 8). Mohammad (2008) further indicated that vocabulary learning was greatly associated with memory; therefore, mnemonics were crucial for foreign language vocabulary learning. Moreover, the effect of memory strategies on long-term vocabulary retention was well-revealed in Johnson and Obi (1993). In addition to memory strategies, cognitive strategies facilitate vocabulary learning in that effective vocabulary learning requires the information manipulating skills of learners to digest the incoming new information. These information-processing strategies, particularly verbal or written repetitions, are popular among EFL learners because they help promote learners' levels of vocabulary proficiency (Schmitt, 1997). In sum, memory and cognitive strategies found to have the strongest relations with adults' overall VLS use came up with no surprise at all.

PEDAGOGICAL IMPLICATIONS

The current study provided an overview picture on the frequency and type of VLS use among college students in Taiwan. To help EFL teachers gain more insightful knowledge about their students' vocabulary learning and related strategy use, pedagogical implications are provided based on the major findings of the study. First and foremost, vocabulary learning plays an indispensable and fundamental role in all language learning and is an everlasting learning process (Oxford & Crookall, 1999; Wei, 2007). Without this lexical competence, other language skills, reading and writing in particular, would fail to develop.

The previous literature has indicated that there is a close link between proficiency and use of vocabulary strategies. Good and poor language learners differ in their use of vocabulary strategies (e.g. Gu, 1994; Liu, 2006; Wei, 2007). Therefore, to increase learners' awareness of such strategies use was remarkably important. Empirical research has also asserted that effective vocabulary learning can be achieved by introducing learners some facilitative instructions of vocabulary strategies (e.g., Brown & Perry, 1991; Chu, 2008; Mercer, 2005; Oxford & Scarcella, 1994). EFL teachers should assist their students in becoming autonomous strategy users (Liao, 2004). In this study, the college participants of technological orientation were found to be very moderate users of vocabulary strategies. It is thus suggested that explicit vocabulary strategy training or instruction be integrated into such EFL classrooms to promote lifelong autonomy of learning for the adult learners.

In addition, it was also found that college students seemed unfamiliar with meta-cognitive and social strategies. As for such concern, encouraging learners to reflect on the vocabulary learning process of their own as a way to activate their meta-cognitive behaviors is highly recommended. Instructions for promoting the conscious use of meta-cognitive vocabulary strategies help narrow the gap between English and non-English majors' strategic behaviors (Liao, 2004). To enhance the use of social strategies, more cooperative vocabulary learning opportunities should be given in class to prompt interpersonal interactions. Although this study discovered major differences might not be a crucial factor to influence choice of VLS, certain individual strategies did mark a great difference between English and non-English majors in their use of vocabulary strategies. Particularly, those of (1) studying the sound of a word, (2) imaging word form, (3) analyzing part of speech, (4) using a bilingual dictionary, and (5) keeping a vocabulary book are advised to be instructed to non-English majors because they were found to be more effective strategies used by the English-major participants in learning new words. Another important result found in the present study was that females significantly employed more determination strategies, and males used more meta-cognitive strategies. EFL teachers should be aware of such gender differences in VLS use. Lastly, emphasis should be placed on the training of memory and cognitive strategies since these strategies are most relevant to the strategic behaviors of vocabulary learning.

SUGGESTIONS FOR FURTHER RESEARCH

When the limitations of this study are concerned, suggestions for further research on VLS are addressed. First of all, the findings of the current study can only be generalized to the college EFL learners of homogeneous academic backgrounds in South Taiwan. Additional subjects from other regions of the country could be recruited for further studies so as to increase the generalizability of any VLS-related findings with a more representative sample. In addition, this study mainly recruited the non-English majors from a business track. Further research may add a technical track with learners drawn from engineering, computer science, etc. This will definitely elicit more intriguing findings with a greater versatility on the adult learners of differing academic backgrounds. Moreover, major and gender factors were detected in this research. Other variables such as learning style and motivation as noted in Oxford and Nyikos (1989) are worth probing under the framework of VLS use. Last but not least, this study centered on a survey approach to collect data. Additional interviews and reflective journals on vocabulary strategy use may be integrated into such survey research to achieve methodological triangulation for firm conclusions to be reached.

REFERENCES

Barcroft, J. (2004). Second language vocabulary acquisition: A lexical input proceeding

- approach. *Foreign Learning Annals*, 37(2), 200-208.
- Boote, C. (2006). Vocabulary: Reasons to teach it, an effective teaching method, and words worth teaching. *The NERA Journal*, 42(2), 24-28.
- Brown, T. S., & Perry, F. L. (1991). A comparison of three learning strategies for ESL vocabulary acquisition. *TESOL Quarterly*, 25, 655-70.
- Catalán, R. (2003). Sex differences in L2 vocabulary learning strategies. *International Journal of Applied Linguistics*, 13(1), 54-77.
- Chang, Y. P. (2005). English learning strategy use: A study of Yu-Da college students. *Yu-Da Academic Journal*, 9, 21-41.
- Chang, C. Y., Liu, S. C., & Lee, Y. N. (2007). A study of language learning strategies used by college EFL learners in Taiwan. *Ming-Dao Journal of General Education*, 2, 235-261. (ISSN 1993-6346)
- Chen, H. C., & Yeh, H. C. (2004). College students' difficulties and strategies in EFL vocabulary learning. *Wu-Feng Junior College of Technology and Commerce Journal*, 12, 107-114.
- Chiang, H. (2005). The use of vocabulary learning strategies of EFL university students in Taiwan. *CEPS*, 26-54.
- Chou, L. (2007). Exploring student talk in a Taiwanese English reading class. *Journal of Educational Research and Development*, 3(2), 137-162.
- Chu, C. W. (2008). *The effects of memory strategy instruction on vocabulary learning: A study of elementary fourth-grade students*. Unpublished master's thesis, National Pingtung Institute of Commerce, Pingtung, Taiwan.
- Cohen, A. D. (1998). *Strategies in learning and using a second language*. U.K.: Longman.
- Ehrman, M., & Oxford, R. (1989). Effects of sex differences, career choice, and psychological type on adult language learning strategies. *Modern Language Journal*, 73, 3-13.
- Fan, M. Y. (2003). Frequency of use, perceived usefulness, and actual usefulness of second language vocabulary strategies: A case study of Hong Kong learners. *The Modern Language Journal*, 87(2), 222-241.
- Ghazal, L. (2007). Learning vocabulary in EFL contexts through vocabulary learning strategies. *Novitas-ROYAL*, 1(2), 84-91.
- Goh, C. (1999). How much do learners know about the factors that influence their listening comprehension? *Hong Kong Journal of Applied Linguistics*, 4(1), 17-40.
- Goh, C. M., & Foong, K. P. (1997). Chinese ESL students' learning strategies: A look at frequency, proficiency, and gender. *Hong Kong Journal of Applied Linguistics*, 2(1), 39-53.
- Goulden, R., Nation, P., & Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics*, 11, 341-358.
- Green, J. M., & Oxford, R. L. (1995). A closer look at learning strategies, L2 proficiency, and gender. *TESOL Quarterly*, 29, 261-297.
- Griffiths, C. (2003). Patterns of language learning strategy use. *System*, 31, 367-383.
- Gu, Y. (1994). Vocabulary learning strategies of good and poor Chinese EFL learners. *Language and Learning*, 3, 376-401.
- Gu, Y. (2002). Gender, academic major, and vocabulary learning strategies of Chinese EFL learners. *RELC Journal*, 31(1), 35-54.
- Gu, Y., & Johnson, R. K. (1996). Vocabulary learning strategies and language learning outcomes. *Language Learning*, 46(4), 643-679.
- Hsu, S. (2006). The reading strategies used by EFL technical students. *Journal of Nan-ya Institute of Technology*, 26, 159-174.
- Huang, H. (2008). *The investigation of English proficiency and listening strategy use in*

- Taiwanese university students*. Unpublished master's thesis, Providence University, Taichung, Taiwan.
- Johnson, D., & Obi, S. C. (1993). Mnemonics: Can you spell it? (ERIC Document Reproduction Service No. ED 394 220)
- Khaldieh, S. A. (2001). The relationship between knowledge of I'raab, lexical knowledge, and reading comprehension of nonnative readers of Arabic. *The Modern Language Journal*, 85(3), 416-431.
- Kudo, Y. (1999). *L2 vocabulary learning strategies*. Unpublished master's thesis. University of Hawaii, Honolulu.
- Lawson, M. J., & Hogben, D. (1998). Learning and recall of foreign-language vocabulary: Effects of a keyword strategy for immediate and delayed recall. *Learning and Instruction*, 8(2), 179-194.
- Lee, S. (2007). Vocabulary learning strategies of Korean university students: Strategy use, vocabulary size, and gender. *English Teaching*, 62(1), 149-169
- Li, D. (2005). How to foster learn autonomy in English teaching and learning. *Sino-US English Teaching*, 2(7), 45-47.
- Li, Y. C. (2004). *A study of technological university freshmen's performances in decoding vocabulary and strategy selection*. Unpublished master's thesis, National Yunlin University of Science & Technology, Yunlin, Taiwan.
- Liao, Y. I. (2004). A survey study of Taiwan EFL freshmen's vocabulary learning strategies. *Journal of National Pingtung Normal University*, 21, 271-288.
- Lin, W. T. (2003). *A study of TVBS college students' vocabulary size and the vocabulary of their English field-specific textbooks*. Unpublished master's thesis, National Yunlin University of Science & Technology, Yunlin, Taiwan.
- Liu, D. (2004). EFL proficiency, gender, and language learning strategy use among a group of Chinese technological institute English majors. *ARECLS E-Journal*, 1, A5. (ISSN 1743-159X)
- Liu, J. (2006). A comparison between good and poor language learners in vocabulary learning. *Sino-US English Teaching*, 3(3), 18-20.
- Mercer, M. (2005). Vocabulary strategy work for advanced learners of English. *English Teaching Forum*, 43(2), 24-35.
- Miao, Y. (2006). An investigation of the factors affecting Chinese students' listening comprehension of English. *Sino-US English Teaching*, 3(3), 27-30.
- Min, H. Z., & Hsu, W. S. (2008). The impact of supplemental reading on vocabulary acquisition and retention with EFL learners in Taiwan. *Journal of National Taiwan Normal University: Humanities & Social Sciences*, 53(1), 83-115.
- Mohammad, M. F. (2008). In search of the best technique for vocabulary acquisition. *Estonian Papers in Applied Linguistics*, 4, 121-138.
- Myers, M., & Paris, S. G. (1978). Children's metacognitive knowledge about reading. *Journal of Educational Psychology*, 70, 680-690.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Boston: Heinle & Heinle Publishers.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nyikos, M. (1990). Sex-related differences in adult language learning: Socialization and memory factors. *Modern Language Journal*, 74(3), 273-287.
- O'Malley, J. M., Russo, R. P., Chamot, A. U., Stewner-Manzanares, G., & Kupper, L. (1985). Learning strategy applications with students of English as a second language. *TESOL Quarterly*, 19(3), 557-584.
- Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*.

- Boston: Heinle & Heinle Publishers.
- Oxford, R. L., & Crookall, D. (1989). Research on language learning strategies: Methods, findings, and instructional issues. *Modern Language Journal*, 73, 404-19.
- Oxford, R. L., & Nyikos, M. (1989). Variables affecting choice of language learning strategies by university students. *The Modern Language Journal*, 73(3), 291-300.
- Oxford, R. L., & Scarcella, R. C. (1994). Second language vocabulary learning among adults: State of the art in vocabulary instruction. *System*, 22(2), 231-243.
- Oxford, R. L., Nyikos, M., & Ehrman, M. (1988). Vive la difference? Reflections on sex differences in use of language learning strategies. *Foreign Language Annals*, 21, 321-329.
- Phakiti, A. (2003). A closer look at gender and strategy use in L2 reading. *Language Learning*, 53, 649-702.
- Politzer, R. L. (1983). An exploratory study of self reported language learning behaviors and their relationship to achievement. *Studies in Second Language Acquisition*, 6(1), 54-68.
- Poole, A. (2005). Gender differences in reading strategy use among ESL college students. *Journal of College Reading and Learning*, 36(1), 7-20.
- Qin, L. (2007). EFL teachers' factors and students' affect. *US-China Education Review*, 4(3), 60-67.
- Riazi, A. (2007). Language learning strategy use: Perceptions of female Arab English majors. *Foreign Language Annals*, 40(3), 433-440.
- Rubin, J. (1994). A review of second language listening comprehension research. *The Modern Language Journal*, 78, 199-221.
- Ruutemets, K. (2005). *Vocabulary learning strategies in studying English as a foreign language*. Unpublished master's thesis, University of Tartu, Estonia.
- Sheorey, R., & Mokhtari, K. (2001). Differences in the metacognitive awareness of reading strategies among native and non-native speakers, *System*, 29(4), 431-449.
- Su, M. M. (2005). A study of EFL technological and vocational college students' language learning strategies and their self-perceived English proficiency. *Electronic Journal of Foreign Language Teaching*, 2(1), 44-56.
- Sung, C. (2006). *An investigation into English vocabulary learning strategies used by college students in Taiwan*. Unpublished Master thesis, Leader University, Tainan, Taiwan.
- Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition, and pedagogy* (pp. 199-226). Cambridge: Cambridge University Press.
- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
- Schmitt, N., & Schmitt, D. R. (1993). Identifying and assessing vocabulary learning strategies. *Thai, TESOL Bulletin*, 5, 27-33. (ERIC Document Reproduction Service No. ED 380 994)
- Smith, C. B. (2003). *Vocabulary's Influence on Successful Writing*. *ERIC Topical Bibliography and Commentary*. ERIC Clearinghouse on Reading, English and Communication. (ERIC Document Reproduction Service No. 480 633)
- Wang, Y. H. (2004). *An investigation into vocabulary learning strategies used by senior high schools in Taiwan*. Unpublished Master thesis, National Chengchi University, Taiwan.
- Wei, M (2007). An examination of vocabulary learning of college-level learners of English in China. *The Asian EFL Journal*, 9(2), 93-114.
- Wu, W. S. (2005). Use and helpfulness rankings of vocabulary learning strategies

- employed by EFL learners in Taiwan. *Journal of Humanities and Social Sciences*, 1(2), 7-13.
- Yek, S. M. (2006). *Memory enhancing vocabulary learning strategy instruction*. Unpublished master's thesis, Ming Chuan University, Taipei, Taiwan.
- Zimmerman, C. B. (2000). Historical trends in second language vocabulary instruction. In J. Coady & T. Huckin (Eds.), *Second Language Vocabulary Acquisition*. Cambridge: Cambridge University Press.
- Zimmerman, K. J. (2004). *The role of vocabulary size in assessing second language proficiency*. Unpublished master's thesis, Brigham Young University, Provo, Utah.

Teacher Questioning Strategies and Wait Time

Mei-chen Chen (陳美貞), Yen-hua Lee (李妍樺)

National Pingtung Institute of Commerce

meichen@npic.edu.tw

onmy_7@hotmail.com

Practices and perceptions of questioning strategies have been found as having direct impact on student learning outcomes. To date, not much research has discussed teachers' questioning techniques, in particular, in the context of EFL classroom. Empirical research on the comparison of practiced and perceived teacher questioning is even less. This study thus aimed to detect the discrepancies between real classroom practices and teachers' self-perceptions of their use of questioning strategies and wait time. Bloom's cognitive taxonomy (1956) served as the crucial criterion to differentiate the six hierarchical levels of teacher questioning. Six experienced EFL teachers and their six 10th-grade classes of a total of 221 students took part in the study. Classroom observations and survey questionnaires targeting these teachers' questioning behaviors were triangulated for the data collection. The present study found: (a) on the average, about 25 questions were given in each period of class; the top two questioning types were knowledge (55%) and comprehension (29%), and the bottom ones, synthesis (0.5%) and evaluation (0.5%); (b) in the classroom observation, the teachers were often found to take the initiative to answer their own questions to save wait time, especially for higher-level questions; (c) no significant relationship was detected between the cognitive level of teacher questioning and observed wait time; (d) all teacher participants thought highly of the role of teacher questioning in teaching and learning English; (e) there was a significant discrepancy between the actual classroom practice and teacher perceptions of questioning strategy use; analysis, synthesis and evaluation questions (accounting for 5% totally) were rarely found in the actual classroom practice, but these questions were reported to be used frequently in the questionnaire; (f) the teachers generally believed that they used the six types of questioning in an equivalent frequency, which contradicted what had been observed in their questioning behaviors; (g) there was a significant discrepancy between the wait time of actual classroom practice ($M=1.70$ seconds) and reported wait time in the questionnaire; EFL teachers tended to give shorter wait time in class than they believed they actually would. Pedagogical and research implications of the findings were presented.

INTRODUCTION

Questioning is a vital instructional behavior in an ESL/EFL classroom (Chang, 1990; Chen, 2005). A good questioning initiates dialogues between teachers and students and determines the quality and progress of language teaching. Wilen (1991) specified the various functions of teacher questioning as "to stimulate students participation, to conduct a review of materials previously read or studied, to stimulate discussion of a topic, issue, or problem, ...and to support student contributions in class" (p. 8). In the past decades, the effects of questioning strategies have been probed in the framework of language instruction (e.g., Chang, 1990; Chen, 2005; Chuska, 1995; Cortazzi & Jin, 1996; Dillon, 1988; Hsu, 2001; Ralph, 1999). Some empirical studies have, however, indicated that most language teachers are still unaware of the importance and effect of teacher questioning on student learning (e.g., Chen, 2005; Dillon, 1988; Wong, 2005). There is an imperative need to improve the questioning skills used by these teachers.

Bloom (1956) proposed a cognitive taxonomy with six hierarchical levels of thinking associated with teacher questioning behaviors. Knowledge questions aimed to examine student recall of facts, concepts, and information. Comprehension questions dealt with the

ability to understand given information and required students to explain, compare, and translate. Application questions explored students' problem-solving skills. In higher-level analysis questions, students were asked to analyze and identify components in an event. Synthesis questions were concerned about the abilities to connect, conclude and create different themes. In the highest-level evaluation questions, students were required to argue, justify and predict according to certain criteria. The following table illustrates different questioning levels with corresponding model sentences based on the Bloom's taxonomy.

Table 1 *Questioning Types and Model Sentences*

Types	Instructions	Model Sentences
1. Knowledge	Fact, concept and information	What is...? Can you recall...?
2. Comprehension	Explanation and comparison	How would you know...? Can you translate...? Can you distinguish between...?
3. Application	Solution, transformation and illustration	How would you use...? What examples can you find to...?
4. Analysis	Identification and classification	What characteristics are there...? What components are there...? What are some of the motives behind...?
5. Synthesis	Summary and connection	What conclusions can you draw...? What inference can you make...?
6. Evaluation	Argument, justification and prediction	Do you agree or disagree with...? Why? What would happen if...? What changes to...would you recommend?

Different types of question basically serve different instructional purposes. English teachers could not only explore information through knowledge questions to assess students' progress of learning, but also initiate more classroom interactions via higher-level evaluation questions. Particularly, higher-cognitive synthesis and evaluation questions can engage students in learning, elicit more thinking, and motivate them to achieve higher learning goals (Wilén, 2001). Unfortunately, a common problem for teachers' use of verbal questioning in the classroom is a lack of knowledge about questioning taxonomies and sequencing (Vogler, 2005). As observed in Wong (2005), over fifty-four percent of questions asked by language teachers were geared to the knowledge level, and a similar percentage (i.e., 60%) was also detected in Gall (1970) for the same type of questions in the reviewing of research articles about questioning practices. With the emphasis on the acquisition of knowledge or facts, higher-order questions are often neglected in the questioning process. Teacher questioning does pose certain problem in the regular classroom practice. Moreover, wait time plays an important role in questioning. It can help reduce student confusion and frustration in the process of questioning (Vogler, 2005). The length of wait time is greatly associated with higher-level thinking behavior found in students' responses (Lake, 1975). Longer wait time induces higher cognitive thinking, which in turn brings more divergent answers from students. Note that language teachers in Taiwan tend to give short wait time, restate questions frequently, and provide negative feedback (Zhang, 1999). Generally, sufficient wait time is very necessary to bring about successful teacher questioning (Lake, 1975); however, this quality is rarely found in the language learning classroom here (Zhang, 1999).

Although a great number of studies have explored teacher questioning behavior, not many of them have examined such quality in EFL settings (e.g., Chang, 1990; Cortazzi & Jin,

1996; Huang, 1998; Wong, 2005). In particular, little research has been conducted in observing the actual classroom practices of vocational high school English teachers in their use of questioning skills. Furthermore, the empirical link between questioning level and wait time has remained surprisingly unexplored, especially in the EFL context of Taiwan. Considering this relative lack of regional research, this study attempted to expand the existing literature by detecting the association between the two variables. In addition to classroom observation on teacher questioning behavior and related wait time, a deeper exploration into teachers' self-perceived questioning via a questionnaire survey would be conducted in this study. Research on English teachers' awareness of their own questioning strategy use is of paramount importance since not much information of this sort has been documented in the local literature (Chen, 2005). A further comparison between what teachers believe they do and what they actually do in class would, for certain, shed additional light on the existing research of teacher questioning.

Generally, this study attempted to probe teacher questioning and associated wait time in the EFL classrooms of vocational high school in southern Taiwan. More precisely, there were three main purposes: (1) to examine the type and frequency of teacher questioning, (2) to explore the relationship between questioning level and wait time, and (3) to compare English teachers' self-reported (or self-perceived) questioning and wait time to their actual questioning behavior in class. With these purposes in mind, several predictions were made. Although the previous literature had claimed teacher questioning as one crucial teaching technique, higher-level questions were less frequently used in classroom instruction (e.g., Chen, 2005; Chuska, 1995; Cortazzi & Jin, 1996; Dillon, 1988; Hsu, 2001; Ralph, 1999; Vogler, 2005; Weinbaum, Allen, & Blythe, 2004; Zhang, 1999). It was, therefore, hypothesized that lower-level questions, especially the knowledge and comprehension types, would be used most often in the EFL setting since they were relatively easier to ask and respond (Chuska, 1995). When the relationship between teacher questioning and wait time was considered, the factor of wait time was usually neglected by the inservice teachers while giving questions (Dillon, 1981; Lake, 1975; Rowe, 1974; Vogler, 2005; Wong, 2005; Zhang, 1999). It was thus anticipated that the level of teacher questioning would not be correlated with the involved wait time length. Lastly, in the previous research, gaps had been noted in teachers' beliefs and their actual teaching behaviors (e.g., Chen, 2005; Dillon, 1988; Wong, 2005; Zhang, 1999). Therefore, discrepancies were expected to be detected in the teachers' perceived use and their actual classroom practices of questioning strategies. It was hoped that the present study would provide a deep insight into the questioning practices of the EFL teachers of vocational high school, so effective training programs on questioning techniques for both preservice and inservice teachers could be developed and intervened.

LITERATURE REVIEW

Teacher questioning has been viewed as one crucial teaching technique in a conventional classroom instruction (Chen, 2005; Chuska, 1995; Cortazzi & Jin, 1996; Dillon, 1988; Hsu, 2001; Ralph, 1999; Vogler, 2005; Weinbaum, et. al., 2004; Zhang, 1999). The most pronounced and longstanding taxonomy of such questioning was first proposed by Bloom (1956). Ever since then, related research has followed along, and educators have started to value this educational hierarchy of differing levels of cognitive behavior. Gallagher and Aschner (1963) further introduced cognitive-memory and three other types of thinking (i.e., convergent, divergent and evaluative thinking) to categorize teacher questioning behavior in their proposed taxonomy. Cognitive-memory questions involved simple skills such as memory recall or recognition. Convergent thinking, a higher level in this taxonomy, required integration and analysis skills. Divergent thinking dealt with the ability to use new perspectives or ideas on the learned knowledge. Evaluative thinking, the highest level, was

concerned about the ability of making judgments, sharing values, and making choices. Krathwohl (2002) reviewed Bloom (1956) and suggested a similar taxonomy with six types of questioning: remember, understand, apply, evaluate, analyze, and create. The researcher renamed the categories into verbs in order to fit the form of teaching objectives. Particularly, the highest level, create, required the ability to put elements together to make an original product. Regardless of the existence of varying questioning taxonomies, it was the Bloom's construct of six levels of cognitive thinking that was of great concern in the current research on teacher questioning.

There are several functions of teacher questioning. Hunkins (1972) categorized the functions of teacher questioning into centering, expansion, distribution and order. Centering was the first function in which teachers helped focus students' attention on certain materials and procedures via asking questions. Expansion function dealt with a deeper exploration into target materials and the comprehension of them. It aimed to activate students' thinking from knowledge memory to creativity. Distribution function was to increase students' participation in the classroom and to encourage discussion about given issues. Then, order function depicted that appropriate questioning from teachers promoted classroom management. Generally, these questioning functions facilitated both teaching and learning.

Gall (1970) examined questioning by reviewing a few articles about classroom questioning practices, effect of questions on students' behavior, and programs to improve questioning behavior. The findings in these studies on questioning practices were quite consistent, and there was no vital change in the types of questions which teachers commonly asked in the classroom. About 60% of classroom questions required students to recall facts, 20% demanding students to think, and the remaining 20%, set up for classroom management. The researcher further noticed a difficulty in differentiating students' responses because what constituted good answers to higher-level questions still remained unclear. It was concluded that to provide teachers with effective training programs of questioning strategies was very necessary. In an empirical study, Chang (1990) explored how questioning influenced student participation in EFL classroom activities with three English teachers and 174 senior high students in Taiwan. Classroom observation was the main approach to collect data in this study. Eliciting and responding questions were found to be asked more frequently in the classrooms. Cortazzi and Jin (1996) further examined teacher questioning types and reasons students did not like to ask questions. Two hundred and seven experienced English teachers were interviewed with 88 recruited from Great Britain, 59 from Malaysia, and 60 from China. It was found that many English teachers asked questions in a sequence from the easiest to more difficult. These teachers usually started with either-or questions because they were easy for their students to respond since the answer was in the question itself. Yes-no questions came next because they gave learners a limited choice. Then, the teachers proceeded to ask more thought-provoking questions, namely, Wh- questions. This study also surveyed on 429 students and found that most of them did not want to ask questions because they were shy and feared that others might laugh at their mistakes. Vogler (2005) specifically indicated that a common problem for teachers' use of verbal questioning in the classroom was a lack of knowledge about questioning taxonomies and sequencing. According to the researcher, teacher questioning should be valued due to its positive impact on student learning, such as monitoring student comprehension and stimulating more cognitive behavior.

Wait time is considered as important as questioning skills in the empirical research. Rowe (1974) noted that sufficient wait time was critical for students to think, especially for doing higher-level thinking, after a question was raised. The researcher noticed that the average teacher wait time was only three seconds. When the teacher participants had received trainings on questioning strategies and use of wait time for classroom practice, the

quantity and quality of their students' responses improved dramatically. Such finding suggested the importance of sufficient wait time for students to do higher-level thinking and to respond more precisely. Lake (1975) conducted an experimental study to determine the effect of wait-time on the cognitive complexity of students' responses. Seventy-five American fifth-grade students were randomly assigned to control and treatment groups which were respectively given one-second and three-second wait time to the same questions over three-week instruction. The results indicated that the students in treatment groups showed a significant progress on thinking behaviors and reactions to the given questions. Longer length of wait time was found highly associated with higher-level thinking. Vogler (2005) further explained that longer wait time could help reduce student confusion and frustration in the process of questioning. Short wait time would, in contrast, induce such undesirable behaviors and students were even unwilling to think in response to teachers' questions.

Although the length of wait time was found to be one of the important elements that could affect student learning, language teachers in Taiwan were more likely to give short wait time (Zhang, 1999). Wong (2005) explored four Elementary School Chinese teachers' questioning skills via qualitative classroom observation. The study found that three out of four teachers tended to provide more knowledge questions and fewer creation questions in the five types of questioning: knowledge, application, creation, evaluation, and management. The Chinese teachers mostly gave three-second wait time to their students for knowledge questions. Although all of the teachers indeed extended the length of wait time based on different levels of questioning, the difference between the wait time of knowledge questions and that of evaluation questions was not statistically significant, suggesting equivalent wait time was given to the two types of questions of very different cognitive levels. Dillon (1981) also indicated that wait time should be varied according to different questioning types and classroom activities. Teachers should give more wait time to their students who could take time to think more for higher-level questions of synthesis and evaluation.

To date, scarce literature has been documented on EFL teachers' awareness of their own questioning strategy use. In Chen (2005), comparisons between one native and one non-native English teachers' questioning behavior and perceived use of such strategy were conducted. Two such teachers and 13 students in a children language school in Taipei City were observed and interviewed during a two-month research period. It was found that the intonation questioning with "OK?" or "Right?" was most frequently used in the classroom. The results from the interview about the teachers' perceptions of their frequently-used questioning types were consistent with what was observed in their actual classroom practices. Both of the native and non-native teachers knew they provided a high proportion of intonation questions and wh-questions while they were teaching English. However, these teachers' self-perceived use of questioning in terms of differing cognitive levels was, unfortunately, left unexplored in the study. The present study attempted to do so to expand the current literature of perceived strategy use of questioning.

In sum, the research mentioned above has unveiled the value of teacher questioning, the relationship between questioning types and wait time, and the perceived use of questioning strategies. However, with the limited literature available, the discrepancies between what EFL teachers believe they do and what they actually do in class in terms of questioning and involved wait time have still remained unclear. Very few studies, to date, have explored such relationship of questioning level and wait time, and the discrepancies in teacher beliefs and teaching practices, in particular, at vocational high schools based on the Bloom's taxonomy. The present study thus directed the research attention to these issues to discern the gaps between practiced and perceived teacher questioning.

METHODOLOGY

The purpose of the present research was to explore teacher questioning type and its relation to wait time for English teachers in vocational high schools. In addition, a comparison of English teachers' self-reported questioning types and wait time to their actual questioning behavior in class was made. Classroom observation and questionnaire survey were conducted to collect data. These data collecting approaches are common research practices of teacher questioning strategies (e.g., Chen, 2005; Hsu, 2002; Wong, 2005; Yu, 2006). A classroom observation elicits real-life information about actual classroom practices (e.g., Iwami, 2001; Richard, 1990). A questionnaire survey is a common practice in strategy research because it assesses strategy use in a neutral, non-threatening context (Oxford & Burry-Stock, 1995; Oxford, 1990). It is a regular measure to explore attitudes and perceptions.

Subjects

Six EFL teachers and six of their 10th-grade classes of a total of 221 students took part in the study. These classes were recruited from three public vocational high schools in Tainan, Kaohsiung and Pingtung, two classes from each school. The six teachers in charge of these classes were experienced in teaching English (see Table 2). The number of the male teachers was equal to that of the female ones. This arrangement was to avoid any potential gender effect on the use of questioning strategies. Each teacher agreed to give one of his/her classes to be observed for this research study. There were two observations in each class, 50 minutes for each observation. When the observations were completed, the teachers were surveyed for the awareness of their own questioning behavior. The entire observing and surveying procedures took about three weeks to complete. In addition, the six classes under these teachers were instructed with the same textbook "English Reader I" published by Lungteng for vocation-oriented, tenth-grade students. The possibility of confounded results caused by the use of different textbooks could thus be eliminated.

Table 2 *Demographic Data for Teacher Participants*

School	School A		School B		School C		Total
	Male	Female	Male	Female	Male	Female	
Teaching Length							
3-5 years	0	1	1	0	0	0	2
6-10 years	1	0	0	1	1	0	3
11 years or over	0	0	0	0	0	1	1
Total	1	1	1	1	1	1	6

Instruments

The instruments in this study included a question analysis sheet, and a questionnaire of questioning strategies and wait time. The analysis sheet helped identify the types of questions raised by the six teacher participants in class. When the classroom observations were completed, all of them were asked to complete the questionnaire. The six questioning types of hierarchical levels were derived from the cognitive taxonomy proposed by Bloom (1956). In addition, a video camera and two digital tape recorders were used for each classroom observation. The researcher filmed the entire teaching process and took field notes about the materials and progress of instruction, classroom activities, and observed questioning behaviors. Note that the wait time in this study was defined as "the length of the silent period which occurs immediately before a teacher utterance" after a question is raised, which was the criterion of wait time used by Lake (1975, p. 241). The wait time length for each teacher question was identified from the careful analyses of the recorded films

and voice tapes.

Question Analysis Sheet. Based on Rawadieh (1998), a question analysis sheet was developed to identify distinct questioning types (see Table 3). It was theoretically based and empirically studied. This assessment sheet was designed according to Bloom's Taxonomy, which served as the major criterion to differentiate the level of teacher questioning. The validity of this Taxonomy had been established through theoretical and empirical practices (e.g., Chen, 2005; Chuska, 1995; Cortazzi & Jin, 1996; Dillon, 1988; Gall, 1970). Its reliability had been demonstrated in the frequent use as a foundation for creating other taxonomies in the cognitive domain (e.g., Gallagher & Aschner, 1963; Krathwohl, 2002). In this study, it was the crucial measure to classify the questions raised by the teacher participants into six categories. Each category of questioning strategies was further divided into two or three subcategories. The specific questioning behaviors listed in the subcategories would assist in making accurate decisions on the assigning of teacher questions into six types (Rawadieh, 1998).

Table 3 *Question Analysis Sheet*

Question Number	Bloom's Taxonomic Category														
	Knowledge			Comprehension		Application			Analysis		Synthesis		Evaluation		
	Fact	Concept	Information	Explanation	Comparison	Solution	Transformation	Illustration	Identification	Classification	Summary	Connection	Argument	Justification	Prediction
1															
2															
3															
4															
5															

Questionnaire of Questioning Strategies and Wait Time. The questionnaire was designed to assess the teachers' awareness of their own questioning strategy use and the associated wait time. There were two parts: value of questioning and actual questioning behaviors. The former was composed of five questions and aimed to explore the teachers' general attitudes toward questioning strategies. The Likert's scale of five possible responses ranging from "strongly agree (five points)" to "strongly disagree (one point)" was utilized to exhibit varying degrees of beliefs with each item (see Table 4).

Table 4 *Part-one Items in the Questionnaire*

Part One: Value of Questioning	Strongly agree	Agree	Uncertain	Disagree	Strongly disagree
1 Questioning is one important teaching technique.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2 Improving questioning strategies can make teaching more efficient.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3 Questioning increases students' learning motivation.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4 Teacher questioning engages students and thus increases student participation in learning.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5 During the class, I usually note whether the questions I give match the target materials or not.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Moreover, part two comprising 10 questions was developed to probe EFL teachers' perceptions of their questioning behaviors and use of wait time in class (see Table 5). Question 1 was concerned about teacher awareness of varying questioning levels. Questions 2-7 were geared to the six hierarchical categories (i.e., knowledge, comprehension, application, analysis, synthesis, and evaluation) based on the criteria set by the Bloom's Taxonomy of the Cognitive Domain (1956). In addition, questions 8-10 involved wait time. The frequency scale of five possible responses ranging from "always (five points)" to "never (one point)" was used to measure a self-perceived questioning behavior in each item. All of the items in this questionnaire were carefully reviewed and revised by three English professors to achieve content validity of the survey. The Cronbach Alpha of the overall test reliability was .76, with the internal-consistency coefficients of the items of value of questioning (part one) and actual questioning behaviors (part two) as .69 and .71, respectively. Generally, the statistics reported above suggested the questionnaire was reliable.

Table 5 Part-two Items in the Questionnaire

Part Two: Actual Questioning Behaviors	Always	Usually	Sometimes	Seldom	Never
1. I ask students questions in different cognitive levels.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. I ask students to define and describe the facts taught in class. (Knowledge)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. I ask students to explain something and tell about similarities and/or differences between certain concepts or things. (Comprehension)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. I ask students to apply what they have learned to solve problems. (Application)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. I ask students to identify the characteristics or features of something and classify them. (Analysis)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. I ask students to connect and integrate different main points and make conclusion accordingly. (Synthesis)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. I ask students to judge and argue about something, and make prediction accordingly. (Evaluation)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. During questioning, I give students at least three seconds to respond to the question being asked.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. I adjust the wait-time span based on the cognitive levels of questions. (e.g., the wait time of evaluation questions is longer than that of knowledge questions.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. Generally, how many seconds are given to a student to respond to the question being asked?	About _____seconds.				

Procedures

Having acquired the research permissions from six English teachers, the researcher studied on two tenth-grade classes from each of three vocational high schools. Each class was observed for 2 separate class sessions of 50 minutes each, and the whole teaching procedure was filmed for each observation. The total of 12 classroom observations took about two weeks to complete. One female assistant had been trained to set up the camcorders and tape recorders to help collect the data. One week after the classroom observations, all six teachers were asked to complete one questionnaire concerning questioning strategies and wait time. In the surveying procedure, the purpose of the research was first told, and specific instruction on how to fill out the questionnaire was then given. In addition, a list of six types of questioning derived from the Taxonomy of Educational Objectives (Bloom, 1956) was shown to these teacher participants. Generally, it took about 15 minutes to complete the whole surveying procedure. When all the classroom observations were finished, the film recordings were analyzed, and the questions raised by the teachers in class were identified and transcribed verbatim. Such transcriptions

were later given to two experienced English teachers to sort all these questions into six types based on the criteria listed in the question analysis sheet mentioned previously. If inconsistencies were found between these two raters in identifying the type of questioning for certain questions, a third rater would get involved to make the final decisions. The data were then analyzed statistically via the SPSS program.

Data Analyses

All of the collected data of this study were analyzed by two ways: qualitative analysis and quantitative analysis. As mentioned previously, the entire teaching process of each classroom observation was filmed, and the questions raised by the teachers were transcribed verbatim. These questions were listed according to the progressing of each observed class. One question analysis sheet with clearly-stated criteria of six questioning types was given to two experienced raters to categorize these transcribed questions based on the Bloom's taxonomy. Product-Moment Pearson correlation procedure was conducted between the identified question levels obtained from the two judges to generate the interrater reliability. The result ($r=.76$, $p=.00<.05$) indicated favorable scoring agreement between the two independent raters. More importantly, when a discrepancy in assigning the level of a question was detected, it would be settled by a third judge. Generally, great caution was taken for the rating stage in analyzing the collected teacher questions to strengthen the validity of the results.

Moreover, descriptive statistics such as frequencies, means, and standard deviations were used to assess the teachers' questioning behaviors and wait time uses, as well as their responses to the questionnaire items. Then, the procedures of one-way analysis of variance (ANOVA) on the six question types were executed to explore if there were significant differences in the uses of these questioning types (from knowledge to evaluation) in the practices and perceptions, respectively. In addition, a Spearman correlation analysis was conducted to examine if there was a significant relationship between the level of questioning and the use of wait time. A paired-samples *t*-test was further performed to determine if there was a significant difference between the reported and practiced wait times. The level of significance for all of the statistical tests was set at .05.

RESULTS

This section mainly focuses on reporting the statistical results of the current study. The actual questioning behaviors of the six teachers in their respective classes are described. Then, the correlation between questioning level and wait time is provided. At last, the discrepancies between actual questioning practices and reported questioning strategy use are presented. Interpretations of the results are also given along.

Teacher Questioning Behaviors in EFL Classrooms. This study observed the questioning behaviors of six EFL teachers for 12 periods of class, 2 separate observations for each teacher. As reported in Table 6, the overall frequency of the questioning behaviors occurring in the entire observations was 302 with the use of six types ranging from the maximum of 167 to the minimum of 1. On the average, about 25 questions were given in each period of class. Out of the 302 questioning strategies, the top two types were knowledge (55%) and comprehension (29%), and the bottom two types, synthesis (0.5%) and evaluation (0.5%). Apparently, the uses of questioning behaviors varied dramatically among these types. A successful F test with one-way ANOVA procedure on the uses of the six questioning types was detected ($F(5, 296)=5.535$, $p=.001<.05$), which suggested these types of questions were used in significantly different frequencies by the English teachers. Additional paired-samples *t*-tests revealed that the average usage of knowledge questions was significantly different from those of application ($t(5)=22.16$, $p=.02<.05$), analysis ($t(5)=25.83$, $p=.00<.05$), synthesis ($t(5)=27.66$, $p=.00<.05$) and evaluation questions ($t(5)=27.66$,

$p=.00<.05$). Generally, there were significant differences in the observed use of the six types of questions among EFL teachers; knowledge and comprehension were most frequently used, and synthesis and evaluation, least frequently used.

Table 6 *Questioning Strategies in Classroom Practices*

Questioning			Each teacher (2 class periods)		Each class period		Rank order of usage
Type	Number	(%)	Mean	SD	Mean	SD	
Knowledge	167	(55%)	27.83	20.70	13.91	10.35	1
Comprehension	87	(29%)	14.50	17.78	7.25	8.89	2
Application	34	(11%)	5.66	5.04	2.83	2.65	3
Analysis	12	(04%)	2.00	2.89	1.00	1.44	4
Synthesis	1	(0.5%)	0.16	0.40	0.08	0.20	5
Evaluation	1	(0.5%)	0.16	0.40	0.08	0.20	5
Overall Strategies	302	(100%)	50.30	7.89	25.15	3.95	

Relationship between Questioning Level and Wait Time. There were 302 questions raised by the six vocational high school teachers, and the average wait time of those questions was 1.70 seconds ($SD=1.87$), as indicated in Table 7. The means of the wait time to the six questioning types varied with the highest wait time on application questions ($M=3.67$), and the lowest, on synthesis ($M=0.83$) and evaluation ($M=1.03$) questions. Furthermore, Spearman correlation analysis was conducted to probe the relationship between the six questioning levels (from 1 to 6) and the involved wait times, and produced an insignificant association ($r=-.006$, $p=.91>.05$), suggesting wait time did not vary proportionally according to the cognitive level of questioning type. Generally, no significant relationship was found between the cognitive level of the teacher questions and the associated wait time span.

Table 7 *Level of Teacher Questioning and Associated Wait Time Span*

Questioning Type	Level of Questioning	Wait Time (by second)	SD	Rank Order
Knowledge	1	1.63	1.26	3
Comprehension	2	1.12	0.78	4
Application	3	3.67	4.17	1
Analysis	4	1.65	1.15	2
Synthesis	5	0.83	0	6
Evaluation	6	1.03	0	5
Average		1.70	1.87	

Discrepancies in the Questioning Practices and Perceptions. With the grand mean of 4.40, the 6 teachers highly agreed with the listed values of questioning from the questionnaire (see Table 8). More specifically, they perceived questioning strategies as (1)

one important teaching technique (M=4.30), (2) promoting teaching efficacy (M=4.30), (3) increasing learning motivation (M=4.50), (4) engaging students in learning (M=4.50), and (5) relating to the target materials (M=4.30). In sum, they all thought highly of the role of teacher questioning in teaching and learning English.

Table 8 *Value of Teacher Questioning*

Teacher questioning	Mean	SD
1. as one important teaching technique	4.30	0.51
2. promotes teaching efficacy	4.30	0.51
3. increases learning motivation	4.50	0.54
4. engages students in learning	4.50	0.54
5. relates to the target materials	4.30	0.81
Overall Mean	4.40	0.35

As for the awareness of the use of the six questioning types, the teachers reported that they gave comprehension questions most often (M=4.33), and evaluation questions least often (M=2.83). The remaining four types of questions were given equivalently (M=3.83) to their students (see Table 9). Note that as detected from the observations, analysis, synthesis and evaluation questions (accounting for 5% totally) were rarely found in the actual classroom practices, but these questions were reported to be more frequently used in the questionnaire with the means of 3.83 (usually), 3.83 (usually), and 2.83 (sometimes), respectively. The distinct gap between practices and perceptions in terms of questioning strategy use was well-disclosed here. Apparently, the teacher participants were not very aware of their own strategy use of questioning in class.

Furthermore, the result of one-way ANOVA revealed that there was no significant mean difference among the six perceived types of questioning ($F(5, 6)=1.85, p=.13 >.05$), as shown in Table 10. This suggested that EFL teachers generally thought they used the six types of questioning in an equivalent frequency. This finding conflicted what was found previously that there was a significant difference in the use of the six types of questioning in the real classroom setting ($F(5, 296)=5.535, p=.001 <.05$). The discrepancy in the practiced and perceived questioning was evident in the statistical analyses.

In addition, the paired-samples t-test ($t(5)=-3.143, p=.014 <.05$) indicated that there was a significant discrepancy between the wait time of actual classroom practice (M=1.70 seconds) and reported wait time in the questionnaire (M=31.6 seconds) (see Table 9). This means EFL teachers tended to give shorter wait time while questioning than they thought they actually would. In the questionnaire, the teachers also reported that they usually gave students at least three seconds to respond to a question being asked (M=4.10, SD=.98), and that they would adjust the wait-time spans based on the differing cognitive levels of questions (M=3.33, SD=1.36), which was inconsistent with what was actually observed in the classroom practices. In sum, significant differences were detected in the practices and perceptions of the concerned questioning strategies and wait time.

Table 9 Self-perceived Use of Questioning Strategies and Wait Time

	Survey Item	Mean	SD
Teacher Questioning Strategies	1. I ask students questions in different cognitive levels.	3.50	0.83
	2. I ask students to define and describe the facts taught in class. (Knowledge)	3.83	1.16
	3. I ask students to explain something and tell about similarities and/or differences between certain concepts or things. (Comprehension)	4.33	0.81
	4. I ask students to apply what they have learned to solve problems. (Application)	3.83	0.75
	5. I ask students to identify the characteristics or features of something and classify them. (Analysis)	3.83	0.40
	6. I ask students to connect and integrate different main points and make conclusion accordingly. (Synthesis)	3.83	0.98
	7. I ask students to judge and argue about something, and make prediction accordingly. (Evaluation)	2.83	0.98
Wait Time	8. During questioning, I give students at least three seconds to respond to the question being asked.	4.10	0.98
	9. I adjust the wait-time span based on the cognitive levels of questions. (e.g., the wait time of evaluation questions is longer than that of knowledge questions.)	3.33	1.36
	10. Generally, how many seconds are given to a student to respond to a question being asked? _____ seconds.	31.60	45.20

Note. N=6

Table 10. A Comparison between Practiced and Perceived Questioning

Type	<u>Actual Classroom Practice</u>		<u>Self-perception</u>	
	Rank	Number of Questions	Rank	Frequency Scale (Mean)
Knowledge	1	167 (55%)	2	3.83
Comprehension	2	87 (29%)	1	4.33
Application	3	34 (11%)	2	3.83
Analysis	4	12 (04%)	2	3.83
Synthesis	5	1 (0.5%)	2	3.83
Evaluation	5	1 (0.5%)	3	2.83
ANOVA Test	* $F(5,296)=5.535, p=.001<.05$		$F(5,6)=1.85, p=.13>.05$	

DISCUSSION

Questioning is considered as an important aspect of teaching and learning because of its effect on foreign language instruction. It gives students an opportunity to articulate their thoughts, to interact among themselves, and to benefit from divergent thinking of others

(Vogler, 2005; Wilen, 1991). The purpose of the study was to analyze the cognitive levels of the questions asked in EFL classes at vocational high schools. Based on the Bloom's Cognitive Taxonomy (1956), the collected teacher questions were categorized into six hierarchical levels. The study also examined the relationship between questioning level and wait time use. The discrepancies between teachers' actual classroom behaviors and their perceptions were further probed under the framework of questioning strategies.

Teacher Questioning Behaviors in EFL Classrooms. Consistent with what had been predicted earlier, the study found that there were significant differences in the use of the six types of questions among EFL teachers; knowledge and comprehension were most frequently used, and synthesis and evaluation, least frequently used. It was interesting to find that as the cognitive level of questioning increased, the frequency of the observed questioning behavior decreased dramatically. The results confirmed what had been found in the previous literature that teachers tended to provide lower cognitive questions with a high percentage but seldom asked higher-level questions (Chen, 2005; Chuska, 1995; Cortazzi & Jin, 1996; Dillon, 1988; Gall, 1970; Hsu, 2001; Wong, 2005). This finding consistency might result from the fact that knowledge and comprehension questions are less time-consuming and, more importantly, easier for teachers to ask and for their students to respond (Chuska, 1995; Zhang, 1999), and the follow-up correction or feedback is relatively more precise (Gall, 1970; Hsu, 2001; Wong, 2005). Some researchers have specifically pointed out that classroom teachers' unawareness of questioning levels or insufficient preparation for teaching materials might result in a high tendency of asking lower-level questions (Vogler, 2005; Wong, 2005). In contrast, a higher cognitive question is more divergent and effective in such a way that it usually prompts students to use their own knowledge, experiences, backgrounds, and beliefs to come up with a response from a broader perspective instead of rendering a single correct answer (Chuska, 1995). This type of question is thus seldom found in the regular classroom practices due to its demanding nature on teacher questions and student responses.

Generally speaking, a balanced mix of lower and higher cognitive questions induces an optimal learning effect (Chuska, 1995; Dillon, 1988; Rawadieh, 1998). Lower level questions will remain important because they enhance the acquisition of factual knowledge and lay foundation for attaining higher thinking skills. However, questions which stress knowledge and comprehension alone will not prepare students for more advanced intellectual skills. In fact, higher level questions are more likely to elicit learning experience necessary to such skills as critical thinking, problem solving, decision making, and creative thinking (Bruner, 1996; Shank, 1994). Providing students with an additional number of questions of this sort will not decrease their knowledge, but assist them in employing that knowledge in a good, divergent way. On the other hand, a failure to provide sufficient higher cognitive questions as a part of the instructional process may lead to a negative effect on the development of various thinking levels (Elizabeth, 2000; Painter, 2005).

This study also detected that, on an average, 25 questions were asked in each period of class, which was similar to what was observed by Wong (2005) with 28 questions per class in the elementary school Chinese courses. It seems that the questioning frequency of twenty something is a common teaching practice in a language classroom.

Relationship between Questioning Level and Wait Time. The present study found that the average wait time identified from the 302 questioning tokens was 1.70 seconds, which was shorter than the suggested wait time (i.e., 3 to 5 seconds in average) in Wilen (1991) and Zhang (1999). Regardless of the types of questions the teacher participants asked in class, insufficient wait time was given to their students. As indicated in Zhang (1999), sufficient wait time is very necessary for students to comprehend the questions and to think over how to respond. Although in many cases, teachers might shorten the wait time

length with certain considerations, the fact that their students need more time to develop and organize their answers can not be neglected. Insufficient wait time may force the students to choose not to think about the questions or to provide incorrect answers before thinking twice. The efficacy of questioning could thus be diminished, especially for higher-level questions if sufficient wait time is not present (Lake, 1975; Lin, 2004).

From the classroom observations, the study detected an insignificant relationship between types of teacher questioning and involved wait time spans, which suggested the wait time lengths would not increase or decrease proportionally in accordance with the cognitive levels of the given questions. Namely, the use of wait time had nothing to do with the cognitive level of questioning type, which coincided with what had been predicted earlier. This finding supported those reported in Wong (2005) with an insignificant wait time difference detected between knowledge and high-order evaluation questions. Although language teachers might be aware of questioning levels, they always cut down the wait time span unknowingly due to the tight teaching schedule (Wong, 2005). In the present study, it was further observed that the EFL teachers tended to shorten wait time and take the initiative to answer their own questions, especially for higher-level questions as those of analysis, synthesis, and evaluation. This might, in turn, encourage students to hold their responses to these divergent questions and wait for their teachers to answer for them. EFL teachers should be cautious about this wait time problem and the undesirable consequence it could bring to their students in the process of questioning.

In addition, a sufficient wait time is essential to student thinking, especially at the higher cognitive level. Previous research has suggested that a three- to five-second wait time following a question results in more comprehensive, higher-quality answers (Rowe, 1974). Effective discussion usually relies on good questioning skills, of which wait time may be the most influential (Wilén, 1991). The importance of wait time to questioning is well-revealed from this perspective. Generally, sufficient wait time given according to the cognitive level of question helps assess students' learning process, activate student thinking, and more importantly, induce better answers and more successful classroom discussions.

Discrepancies in the Questioning Practices and Perceptions. In this study, a survey was conducted to probe the teacher participants' perceptions of their questioning strategy use. The results of the questionnaire indicated that the teachers generally thought very highly of the role of teacher questioning in promoting teaching efficacy, strengthening learning motivation, engaging students in learning, and relating to the teaching materials. Regardless of those merits of questioning, the detected insignificant difference in the reported use of the questioning types suggested the teachers believed that they gave the six types of questions in an equal proportion in their classroom questioning. This belief contradicted what was observed in class with a significant discrepancy in the use of high- and low-order questions. The fact that higher-level analysis, synthesis, and evaluation only accounted for 5% of the overall strategy use marked the gap between the reported use and the actual practice in questioning. In addition, there was a significant discrepancy between wait time of actual classroom practice and reported wait time in the questionnaire; EFL teachers tended to give shorter wait time while questioning their students than they thought they actually would. The gaps between what teachers perceive and what they actually do in class are greatly associated with a lack of knowledge about questioning taxonomies and sequencing (Vogler, 2005), or a lack of sufficient training on questioning techniques (Gall, 1970). Such discrepancies between perceptions and actual teaching behaviors may also reflect a high expectation from the EFL teachers for their own questioning behavior, which is, though, not happening in their classroom instruction. The inconsistent results associated with reported, or perceived, questioning behaviors from the questionnaire imply a conflict or struggle between ideal instruction and real-life teaching for these teachers. Further research on

teacher questioning may incorporate additional interviews on individual teachers to discern the gaps in perceptions and practices.

PEDAGOGICAL IMPLICATIONS

There are some instructional implications that can be drawn from the major findings of the study. Questioning is an integral component of the teaching-learning process and an important means through which differing levels of thinking skills are enhanced. This study observed that the EFL teachers of vocational high school seldom used higher-level questions in class. It is thus suggested the Bloom's Cognitive Taxonomy (1956) or other theoretical constructs be used as a criterion for the development of questioning in the lesson plan. These questions should be geared to the goals of instruction and target materials. Note that a balanced mix of lower and higher cognitive questions is of paramount importance here. An elaborate lesson plan of this sort reinforces the awareness of the cognitive levels of the planned questions, and thus induces more effective behaviors in classroom questioning.

In addition, the study found that insufficient wait time was given in the practice of questioning. It is essential for teachers to realize that the function of wait time is more than a period of time waiting for a student response to a teacher question, but a precious chance to observe the reactions from the students toward the questions. A good teacher can often catch the messages from the facial expressions of the students, and he/she can always make good use of wait time. EFL teachers are also advised not to repeat the questions too often or answer their own questions during the wait time span because too many occurrences of these behaviors will interrupt their students' thinking and thus affect the willingness to respond to the questions.

The study also noted that there were significant discrepancies in the teachers' practices and perceptions of questioning and wait time. To shorten such gaps, questioning strategy use should be integrated into preservice and inservice teachers' training programs (Gall, 1970; Rowe, 1974). Questioning is one important teaching technique; however, most EFL teachers do not use it very well in their everyday practice. This may be due to either an unfamiliarity of questioning taxonomies or a lack of realization that their teaching behaviors can be improved through effective questioning. Basically, five domains of questioning deserve our attention: (1) differentiating question levels conceptually, (2) understanding the functions of questioning strategy, (3) improving the use of wait time, (4) integrating questions of different cognitive levels into the lecture, and (5) learning tips on the time management during questioning. These issues are suggested to be discussed in such training programs. In sum, both value and practice of teacher questioning should be stressed in the training to improve the teaching skills of preservice and inservice teachers. A follow-up peer evaluation on the questioning levels and concerned wait time spans is highly recommended to advance the training to the actual classroom practices (Boud & Feletti, 1991; Finn, 1972; Rice, 1977; Riley, 1980). This evaluation should be implemented on a regular basis to achieve optimal teaching and learning through effective questioning.

SUGGESTIONS FOR FURTHER RESEARCH

There are limitations in the study, and suggestions for further research are addressed accordingly. First, an assumption was made that the participants represented a sample of EFL teachers of vocational high school in southern Taiwan. Therefore, the findings and implications of the study should be generalized to the extent that the groups of teachers are similar to the teacher participants in the study. Additional studies on questioning may incorporate a larger sample of teachers; a more representative sample size strengthens the generalizability of the findings and brings more convincing results about the questioning practices. Second, a comparison on the practices of teacher questioning from different

grade levels (e.g., senior high school vs. university) or from different regions (e.g., urban vs. suburban areas) is worth conducting through classroom observations. Third, in addition to grade and regional differences, other variables concerning teachers' demographic data and learning materials deserve more of our attention. These factors could be sex differences (i.e., male vs. female teachers), length of teaching (i.e., experienced vs. novice teachers), and complexity of learning tasks (i.e., simple vs. complex materials). More intriguing findings on teacher questioning could be generated with the examination on some of these suggested variables. Fourth, a different cognitive taxonomy such as the one used in Gallagher and Aschner (1963), or a different form in the presentation of questions as the one studied in Cortazzi and Jin (1996) is worth probing under the framework of questioning and associated wait time. A deeper exploration into the effect of the level of teacher questions on the cognitive complexity of student responses will definitely bring more significant results for the teacher questioning practices. Lastly, additional face-to-face interviews on individual teacher participants should be conducted after the classroom observations and questionnaire survey to triangulate the data collection measures to reconfirm the discrepancies detected in teacher questioning practices and perceptions.

REFERENCES

- Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals*. New York: David McKay.
- Bobrowski, P. (2004). Bloom's taxonomy-expanding its meaning. Retrieved March 11, 2006, from <http://pcrest.com/Bloom.pdf>
- Boud, D., & Feletti, G. (1991). Cognition and improvisation: Differences in mathematics instruction by expert and novice teachers. *American Education Research Journal*, 26(4), 473-498.
- Bruner, J. (1996). *The culture of education*. Cambridge, MA: Harvard University Press.
- Chang, C. P. (1990). A case study of questioning and responding in English conversation classes. *Proceedings of the International Conference on English Teaching and Learning in the Republic of China*, Taipei, Taiwan, 437-456.
- Chen, S. Y. (2005). *Questioning strategy use and perceptions of native and non-native teachers and students' responses in an EFL class at elementary school level*. Unpublished master's thesis, National Taipei Teachers College, Taipei, R.O.C.
- Chuska, K. R. (1995). *Improving classroom questions*. Indiana: Bloomington.
- Cortazzi, M., & Jin, L. (1996). Changing patterns of questioning. *Proceedings of the Fifth International Symposium on English Teaching*, Taipei, Taiwan, 269-282.
- Dillon, J. T. (1981). Duration of response to teacher questions and statements. *Contemporary Educational Psychology*, 6, 1-11.
- Dillon, J. T. (1988). *Questioning and teaching: A manual of practice*. London: Croom Helm.
- Elizabeth, S. (2000). Questioning techniques in the gifted classroom. *Gifted Child Today*, 23(5), 14-21.
- Ellis, R. (2003). *The study of second language acquisition*. Oxford: Oxford University Press.
- Finn, J. D. (1972). Expectation and education environment. *Review of Educational Research*, 42, 387-412.
- Gall, M. D. (1970). The use of questions in teaching. *Review of Educational Research*, 40(5), 707-721.
- Gallagher, J. J., & Aschner, M.J. (1963). A preliminary report: Analysis of classroom interaction. *Merrill-Palmer Quarterly*, 2, 183-194.
- Goodman, L., & Berntson, G. (2000). The art of asking questions: Using directed inquiry in the classroom. *The American Biology Teacher*, 62(7), 473-476.
- Hsu, L. C. (2002). *Using creative questions to facilitate L2 acquisition*. Unpublished master's

- thesis, National Chengchi University, Taipei, R.O.C.
- Hsu, W. H. (2001). *How classroom questioning influences second language acquisition*. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign, Illinois, United States. UMI No.3023080.
- Huang, H. Y. (1998). *A case study of the functions of teacher talk in an EFL classroom*. Unpublished master's thesis, National Kaohsiung Normal University, Kaohsiung, R.O.C.
- Hunkins, F. P. (1972). *Questioning strategies and techniques*. Boston: Allyn and Bacon.
- Iwami, I. (2001). Observation of high school students' real-life communication during a study tour abroad. *JABAET Journal*, 5, 13-37.
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice*, 41(4), 212-218.
- Lake, J. H. (1975). The influence of wait-time on the verbal dimensions of student inquiry behavior. *Dissertation Abstracts International*, 36, 6343A.
- Lin, Y. A. (林育安) (2004). Research on how questioning techniques affect seventh-grade students' learning in math (發問技巧對國中七年級學生數學學習影響之研究). Unpublished master's thesis, National Kaohsiung Normal University, Kaohsiung, R.O.C.
- Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. Boston: Heinle & Heinle.
- Oxford, R. L., & Burry-Stock J. A. (1995). Assessing the use of language learning strategies worldwide with the ESL/EFL version of the Strategy Inventory for Language Learning (SILL). *System*, 23(1), 1-23.
- Painter, J. (2005). Questioning techniques for gifted students. Retrieved October 17, 2005, from <http://www.nexus.edu.au/teachstud/gat/painter.htm>
- Ralph, E. G. (1999). Oral-questioning skills of novice teachers: Any questions? *Journal of Instructional Psychology*, 26(4), 286-296.
- Rawadieh, S. M. (1998). *An analysis of the cognitive levels of questions in Jordanian secondary social studies textbooks according to Bloom's taxonomy*. Unpublished doctoral dissertation, Ohio University, Ohio, United States. UMI No.9829185.
- Rice, D. R. (1977). The effect of question-asking instruction on preservice elementary science teachers. *Journal of Research in Science Teaching*, 14(4), 353-359.
- Richard, R. D. (1990). Teacher observation in second language teacher education. In J.C. Richards & D. Nunan (Eds.) *Second language teacher education* (pp. 43-61). New York: Cambridge University Press.
- Riley, J. P. (1980). A comparison of three methods of improving preservice science teachers' questioning knowledge and attitudes toward questioning. *Journal of Research in Science Teaching*, 17(5), 419-424.
- Rowe, M. B. (1974). Wait-time and rewards as instructional variables, their influence on language, logic, and fate control. *Journal of Research in Science Teaching*, 11(2), 81-94.
- Shank, G. (1994). Shaping qualitative research in educational psychology. *Contemporary Educational Psychology*, 19, 340-350.
- Vogler, K. E. (2005). Improve your verbal questioning. *ProQuest Education Journals*, 79(2), 98-103.
- Weinbaum, A., Allen, D., & Blythe, T. (2004). *Teaching as inquiry: Asking hard questions to improve practice and student achievement*. New York: Teachers College Press.
- Wilensky, W. W. (1991). *What research says to the teacher: Questioning skills for teachers*. Washington, D.C.: National Education Association of the United States.
- Wilensky, W. W. (2001). Exploring myths about teacher questioning in the social studies

- classroom. *The Social Studies*, 92(1), 26–32.
- Wong, J. F.(王靖芳) (2005). *Research on teacher questioning behaviors in elementary schools* (國小教師發問行為之研究). Unpublished master's thesis, National Hsinchu University of Education, Hsinchu, R.O.C.
- Yu, T. T., & Liang, J. S. (2006). *To ask or not to ask: An observational study of an EFL college teacher's display questions in her CLT classroom*. Unpublished master's thesis, National PingTung University of Education, Pingtung, R.O.C.
- Zhang, Y. C. (張玉成) (1999). *Teacher questioning techniques* (教師發問技巧). Taipei, Taiwan: Shin Li.

Effects of Major and Gender on Language Learning Strategy Use

Mei-chen Chen (陳美貞), Chen-ni Wu (吳貞霓)

National Pingtung Institute of Commerce

meichen@npic.edu.tw

s96307005@student-mail.npic.edu.tw

Previous research has indicated that use of strategies facilitates the acquisition of a foreign language, and proficient language learners employ more strategies during learning process. Although tons of empirical studies have probed strategy use in the context of ESL/EFL, not many of them have discussed the factors of academic specialization and sex differences in relation to the choice of language learning strategies. Research of this sort was even less particularly with adult learners in the regional context of EFL. The results have, unfortunately, been mixed. This study thus aimed to determine the effects of major and gender on adult EFL learners' choice of such strategies. One hundred and twenty-five college students of differing academic specializations of language, business, and management were surveyed with a standardized language learning strategy questionnaire (SILL: Oxford, 1990). They were categorized as English and non-English majors. Important findings were: (a) college students used language learning strategies in a moderate degree; (b) metacognitive and compensation strategies were used most frequently, and social and affective strategies, least frequently; (c) significant differences were detected between English-major and non-English-major learners in their overall and specific types of strategy use; (d) no significant differences were found between male and female learners in their strategy use; (e) relatively, major was a more potent factor than gender in determining adult learners' choice of strategies; (f) English-major females significantly employed more strategies than non-English-major females; however, such significant major effect was not detected on male learners. Pedagogical implications and suggestions for further studies were provided.

INTRODUCTION

Language strategies strengthen learning (Okada, Oxford, & Abo, 1996). They play an important role in learners' acquisition of a foreign language (Chamot & Küpper, 1989; Dreyer & Oxford, 1996; Levine, Reves, & Leaver, 1996; Oxford, 1990; Oxford, Lavine, Felkins, Hollaway, & Saleh, 1996). Grainger (1997) defined a language learning strategy as a technique or a specific action used consciously by a learner to assist certain aspect(s) of language learning. It is also viewed as a tool for active and self-directed involvement during learning. Namely, strategies enhance greater self-direction and learner autonomy which are crucial to the active development of ability in a new language (Oxford, 1990). Despite the merits of strategy use in acquiring a foreign language, the extent and associated types of strategy use are still unclear among EFL college learners, particularly those of different academic specializations (Chang, Liu, & Lee, 2007; Ong, 2005; Wu & Chang, 2005; Zhang, 2005). With very limited literature in this regional research, more research is very necessary to discern the effect of major differences on the strategy preferences.

Oxford (1990) initiated a strategy taxonomy in a questionnaire form (SILL: Strategy Inventory for Language Learning, version 7.0 ESL/EFL) to assess foreign language learners' preferences on strategy use. There are six types of strategies in this standardized inventory. Memory, cognitive and compensation are noted as direct strategies because of their direct

links to the mental processing of language learning. Metacognitive, affective and social are classified as indirect strategies which do not involve language learning directly, but help support and manage the learning process. Elaborately, memory strategies help store and retrieve information. Cognitive strategies involve behaviors to think; summarizing, analyzing, and reasoning the target information are examples of such acts. Compensation strategies are concerned about guessing intelligently or overcoming limitations or difficulties in learning. Metacognitive strategies enable learners to plan, monitor, and evaluate their learning progress. Affective strategies deal with emotions, motivations, attitudes, etc. in learning. Lastly, social strategies induce interaction or cooperation with others. This inventory has been highly praised for its validity and reliability in measuring ESL/EFL learners' preferred types of strategy and has been treated as the chief instrument of data collection in a sufficient number of strategy research (e.g., Grainger, 1997; Hiromori, et al., 2003; Khalil, 2005; Nisbet, Tindall, & Arroyo, 2005; Okada, et al., 1996; Oxford & Burry-Stock, 1995; Park, 1997; Wharton, 2000; Zhang, 2005). The present study intended to incorporate the SILL in this strategy study because of its exhaustive and comprehensive classifications in strategy types and its suitability to be implemented in an EFL college classroom in Taiwan.

Although use of strategies facilitates the process of achieving a target language, not all language learners succeed in language learning. Proficient learners seem to have better access to language learning and learn languages more readily and effectively while less proficient learners don't have such advantage (Chen, 2002; Green & Oxford, 1995; Hsieh, 2006; Okada, et al., 1996; O'Malley, Chamot, Stewner-Manzanares, Russo, & Küpper, 1985; Riazi, 2007; Yang, 1993; Yu, 2006). Riazi (2007) further defined sophisticated language learners as those who made use of a wide range of effective strategies in a systematic and organized way. Research and theory in second language learning also strongly suggest that proficient language learners use a variety of strategies to assist them in gaining command over new language skills (O'Malley, et al., 1985). The value of strategies in ESL/EFL acquisition in relation to language proficiency is well-revealed in the former literature.

To date, a great majority of strategy studies have directed their interests toward examining the effects of English proficiency (e.g., Chen, 2002; Hsieh, 2006; Lo, 2007; Nisbet, et al, 2005; Park, 1997; Yang, 1993), cultural backgrounds (e.g., Bedell & Oxford, 1996; Grainger, 1997; Levine, et al., 1996), learning motivation (e.g., Chang, 2005; Chen, 2007; Chuang, 2007; Huang, 2007; Lai, 2007; Liao, 2000; Peng, 2002; Yu, 2006), language learning beliefs (Lin, 2006; Liu, 2004; Shen, 2006; Shih, 2003), learning styles (Lo, 2007; Tsai, 2006), and grade level (Huang, 2005; Riazi, 2007). Although tons of empirical studies have probed strategy use in the context of ESL/EFL, not many of them have discussed the effects of major and gender on adult EFL learners' use of language strategies. Oxford and Nyikos (1989) specifically indicated that field of specialization and sex were two potential factors that could affect the choice of language learning strategies by university students. Unfortunately, empirical studies on these two factors have been limited and the results have been mixed. With the scant strategy research on the major effect (e.g., Chang, et. al, 2007; Ong, 2005; Wu & Chang, 2005; Zhang, 2005), the findings have been contradictory although significant major effect has generally been detected. Chang, et al. (2007) reported that humanities and social science college students used all types of strategies more than their business, engineering and management counterparts. However, Ong (2005) specifically found that metacognitive strategies were most frequently used by business college students, and social strategies were most frequently used by engineering college students. Zhang (2005) further noted that students of engineering school used more metacognitive strategies than those of arts. In addition to major, inconsistent gender findings have been documented in the strategy literature. Several studies have found that females significantly used more

strategies than males (e.g., Chang, et al., 2007; Green & Oxford, 1995; Huang, 2007; Kaylani, 1996; Yu, 2006). In contrast, male superiority in strategy use was reported in Tran (1988) and Zhang (2005). Nisbet, et al. (2005) and Huang (2005) even produced insignificant gender effects in Chinese college students' strategic behaviors. The male and female adult learners were equivalent strategy users in learning English. Generally, mixed findings have been detected in the previous literature probing the effects of major and gender on strategy use. Such inconsistencies call for further research on the concerned issue.

The purpose of this study was to probe the frequency and type of language learning strategy use between English majors and non-English majors in the context of EFL. Sex differences were further examined under the framework of strategy use. More specifically, this study intended to explore: (1) general strategy use among college learners, (2) the effect of major (English majors vs. non-English majors) on strategy use, (3) the effect of gender (males vs. females) on strategy use, and (4) the effect of major by gender on strategy use.

Research predictions were made based on the purposes of the present study. The previous literature indicated that EFL learners did not use language strategies in a high frequency (e.g., Bedell & Oxford, 1996; Chang, et al., 2007; Khalil, 2005; Ong, 2005; Oxford, 2003; Su, 2005; Yu, 2006); therefore, the college participants in the study were expected to be moderate strategy users. As the academic major effect was concerned, English majors had been found as better strategy users than non-English-major college learners (e.g., Chang, et al., 2007; Wu & Chang, 2005). It was thus hypothesized that the English majors would significantly employ more strategies than their non-English-major counterparts. In addition, sex differences had been examined in SILL-based studies, and females had usually had greater strategy use than males (Green & Oxford, 1995). Significant gender effects had also been detected in most of the strategy research of this sort in favor of female learners (e.g., Chang, et al., 2007; Green & Oxford, 1995; Huang, 2007; Kaylani, 1996; Yu, 2006). Therefore, the college females in this study were expected to be more strategic than their male counterparts. It was hoped that the findings of this study would discern the inconsistencies found in the strategy research literature, and the awareness of strategy use among college EFL learners of differing specializations would be raised, so effective strategy training could be intervened to maximize adult learners' English learning experiences.

LITERATURE REVIEW

Language learning strategies are specific actions that students use to improve their progress in L2 learning (Green & Oxford, 1995). To date, different strategy taxonomies toward ESL/EFL have been launched (See Table 1). In Naiman, Fröhlich, Stern and Todesco (1978), several strategy categories were raised: an active task approach, realization of a language as a system, realization of a language as a means of communication and interaction, management of affective demands, and monitoring of second language performance. Bialystok (1979) developed a model to account for the input and output of four types of information processing: functional practicing, formal practicing, monitoring, and inferencing. Rubin (1981) categorized learning strategies into direct and indirect ones to exhibit two distinct ways of language strategy use. Brown and Palinscar (1982) further classified learning strategies into three categories: cognitive, metacognitive, and affective-social strategies. O'Malley, et al. (1985) raised three similar types as those used in Brown and Palinscar (1982). The two studies shared the same idea on how learning strategies should be grouped. More importantly, Oxford (1990) proposed a rather different taxonomy comprising six types of strategy use. The six categories were: memory, cognitive, compensation, metacognitive, affective, and social. They were specially geared to learning second/foreign languages. The first three types belonged to direct strategies because they directly involved the target language in the learning process; in contrast, the last three types

were viewed as indirect strategies because they served to support and manage language learning without such direct involvement in many instances. The strategy inventory (SILL) developed by Oxford (1990) helped both language teachers and students to be aware of the presence of various types of language learning strategies. The present study was simply interested in using this six-factor inventory due to its well-rounded categorization of strategy types and its feasibility in the EFL context of Taiwan.

Table 1 *Different Taxonomies of Language Learning Strategies*

Advocators	Types of language learning strategies
Naiman, et al. (1978)	<ol style="list-style-type: none"> 1. An active task approach. 2. Realization of a language as a system. 3. Realization of a language as a means of communication and interaction. 4. Management of affective demands. 5. Monitoring of second language performance.
Bialystok (1979)	<ol style="list-style-type: none"> 1. Functional practicing. 2. Formal practicing. 3. Monitoring. 4. Inferencing.
Rubin (1981)	<ol style="list-style-type: none"> 1. Direct strategies (clarification/verification, monitoring, memorization, guessing/inductive reasoning, deductive reasoning, and practice). 2. Indirect strategies (creating practice opportunities and using production tricks such as communication strategies).
Brown & Palinscar (1982)	<ol style="list-style-type: none"> 1. Cognitive strategies. 2. Metacognitive strategies. 3. Affective-social strategies.
O'Malley, et al. (1985)	<ol style="list-style-type: none"> 1. Metacognitive strategies. 2. Cognitive strategies. 3. Socio-affective strategies.
Oxford (1990)	<ol style="list-style-type: none"> 1. Direct strategies (memory strategies, cognitive strategies and compensation strategies). 2. Indirect strategies (metacognitive strategies, affective strategies and social strategies).

A great number of strategy studies have conducted the research using the SILL inventory, and consistent effects of ethnicity, proficiency, and grade level have been detected with adult language learners. Levine, et al. (1996) explored the ethnic effect on language strategy use with 117 college students registering in pre-academic EFL courses in America and found that behaviors of language learning were significantly related to cultural factors. Namely, language learners of differing Israeli and Russian cultural backgrounds used different

strategies in learning English. Significant effects of ethnicity were also detected in Bedell and Oxford (1996) with Chinese learners and learners from some other countries, and in Okada, et al. (1996) with Japanese and Spanish learners. Language proficiency is another factor frequently examined with EFL college students in the strategy literature. Park (1997) noted a strong relationship between English proficiency and frequency of strategy use among 332 Korean university students; high achievers significantly used more strategies than their mid- or low- achieving counterparts. Consistent with these findings, Chen (2002) also detected a significant ability effect between high and low achievers on each of the six strategy types among college learners of technological orientation in Taiwan. On the other hand, Riazi (2007) probed the relationship between different grade levels and strategy use with 120 female Arabic-speaking college students. The freshmen significantly used more strategies than did their sophomore and junior counterparts in terms of overall strategy use. Metacognitive and cognitive strategies were most frequently used, and affective strategies, least frequently used for all learners. In a similar research study, Huang (2005) discovered that technological college freshmen in Taiwan used more strategies than those of the other three college grades in term of overall and specific types (i.e., memory, cognitive, and compensation) of strategy use. Generally, significant ethnic, proficiency, and grade effects were consistently found in the past strategy studies with the SILL inventory.

Note that some SILL-based studies have probed the factors of major (e.g., Chang, et al, 2007; Ong, 2005), gender (e.g., Chang, 1990; Green & Oxford, 1995; Huang, 2005; Nisbet, et al., 2005), or both (e.g., Chang, et al., 2007; Wu, 2007) with adult EFL learners in the past two decades. These studies have, unfortunately, produced mixed results in this limited research literature.

Chang, et al. (2007) found significant major effects on the strategy use of 1758 Taiwanese college students of differing academic specializations. The humanities and social science college students significantly employed more strategies than their business, management and engineering counterparts in terms of overall and each of the six types of strategy use. For all college learners, compensation strategies were used most, and affective strategies were used least. However, inconsistent strategy findings were detected in Ong (2005) with 342 Taiwanese college learners. Particularly, social strategies were more frequently used by the engineering college students than their business and management counterparts. Metacognitive strategies were more frequently used by business college students than engineering, humanities and management college students. Among all types of strategies, compensation strategies were most frequently used; whereas, cognitive strategies were least frequently used. Zhang (2005) further noted that students of engineering school used more metacognitive strategies than those of arts. Although significant major effects were also produced in Wu (2007), inconsistencies were generally detected in the type of strategy use among EFL college learners of differing schools in the past literature.

In addition to major, gender differences have been examined in SILL-based studies. Several of such studies found that females significantly used more strategies than males (e.g., Chang, et al., 2007; Green & Oxford, 1995; Wu, 2007; Yu, 2006); in contrast, male superiority in strategy use was reported in Tran (1988) and Zhang (2005). Moreover, Nisbet, et al. (2005), Chang (1990), and Huang (2005) produced insignificant gender differences in Chinese learners' strategic behaviors. Males and females were equivalent strategy users in learning English. In sum, mixed findings were detected in the previous literature in the effects of major and gender on strategy use with adult language learners. Such inconsistencies called for further research on the concerned issue.

METHODOLOGY

The main purpose of this study was to examine the frequency and type of language learning strategy use among EFL college students. Moreover, the relations of major and gender to strategy preferences were explored. Basically, the nature of this present study was a survey research with quantitatively-oriented data analyses. A great majority of empirical studies on learning strategies have been done with the administration of questionnaires (e.g., Bedell & Oxford, 1996; Chang, 1990; Chang, et al., 2007; Chen, 2002; Grainger, 1997; Green & Oxford, 1995; Huang, 2005; Levine, et al., 1996; Nisbet, et al., 2005; Okada, et al., 1996; Ong, 2005; Park, 1997; Riazi, 2007; Wu, 2003; Wu, 2007). A survey research involving the self-reported strategy use from language learners has been a common practice in the strategy literature. In this section, demographic data of the subjects, descriptions of the instrument, data collection procedures, and quantitative approaches employed to analyze the data are presented in sequence in the following.

Subjects

Originally, 154 college students from four intact classes of two colleges in southern Taiwan took part in this study. They were all freshmen from three academic fields: Applied Foreign Languages, International Trade, and Business Administration. These participants were asked to complete a questionnaire of language learning strategies. Although the returning rate of the survey was 100%, twenty-nine questionnaires were discarded due to unfinished test items and/or inappropriate checking on the items. Eventually, 125 questionnaires were valid for further data analyses. These college participants included 58 English majors (M= 8, F= 50) and 67 non-English majors (M= 7, F= 60) from the business and management schools. The two types of learners were compared for their strategy use.

Instrument

In this study, the participants were required to fill out the Strategy Inventory for Language Learning (SILL) from Oxford (1990) and provide some relevant information concerning their background profile of major and gender. This inventory was a standardized test and the original English version 7.0 (ESL/EFL) was specially designed for speakers of other languages who were learning English. Oxford and Burry-Stock (1995) and Oxford (1990) reported high indexes of Cronbach alpha reliability (mostly .91 to .94) for this ESL/EFL version of the SILL across many cultural groups. In Oxford and Nyikos (1989), the SILL yielded a Cronbach alpha of .96 based on a sample of 1,200 university students, and a content validity of .95 using classificatory agreement between two independent judges. Those statistics generally suggested that this inventory was quite valid and reliable.

A Chinese version of SILL adapted from the one used in Liao (2000) with some modifications on the Chinese wording was administered to the student participants to eliminate any sources of misunderstanding due to limited English knowledge. According to Liao (2000), the Chinese version yielded a Cronbach alpha of .96 on the internal test consistency. The present study also produced a similar reliability coefficient as .94 (more specifically, memory strategies as .64, cognitive strategies as .85, compensation strategies as .60, metacognitive strategies as .91, affective strategies as .71, and social strategies as .74). With the Chinese SILL, the overall and specific types of strategy use toward English learning were probed. Comparisons were also made in the frequency and type of strategy use between the English and non-English majors, and between the male and female college learners.

There were fifty questions in this survey with the use of a five-point Likert scale, ranging from 5 points "Always or almost true of me" to 1 point "Never or almost never true of me", which measured the varying degrees of the frequency of strategy use. The test items could be sorted into six categories: memory strategies (questions 1 to 9), cognitive strategies (questions 10 to 23), compensation strategies (questions 24 to 29), metacognitive strategies

(questions 30 to 38), affective strategies (questions 39 to 44), and social strategies (questions 45 to 50). Sample items of the six types of strategies in the SILL are presented in Table 2.

Table 2 *Sample Items of Six Types of Learning Strategies*

Types	Sample Items
Memory	<ol style="list-style-type: none"> 1. I use new English words in a sentence so I can remember them. 2. I connect the sound of a new English word and an image or picture of the word to help me remember the word. 3. I remember new English words or phrases by remembering their location on the page, on the board, or on the street sign.
Cognitive	<ol style="list-style-type: none"> 1. I first skim an English passage (read over the passage quickly) then go back and read carefully. 2. I find the meaning of an English word by dividing it into parts that I understand. 3. I try not to translate word-for-word.
Compensation	<ol style="list-style-type: none"> 1. To understand unfamiliar English words, I make guesses. 2. When I can't think of a word during a conversation in English, I use gestures. 3. If I can't think of an English word, I use a word or phrase that means the same thing.
Metacognitive	<ol style="list-style-type: none"> 1. I notice my English mistakes and use that information to help me do better. 2. I have clear goals for improving my English skills. 3. I think about my progress in learning English.
Affective	<ol style="list-style-type: none"> 1. I encourage myself to speak English even when I am afraid of making a mistake. 2. I give myself a reward or treat when I do well in English. 3. I notice if I am tense or nervous when I am studying or using English.
Social	<ol style="list-style-type: none"> 1. If I do not understand something in English, I ask the other person to slow down or say it again. 2. I practice English with other students. 3. I ask for help from English speakers.

Procedures

As mentioned previously, the questionnaire used in the present study was adapted from Oxford (1990). A Chinese version of SILL was developed to avoid undesired text confusions resulting from a lack of sufficient knowledge in English. The Chinese version was reviewed by one college professor and several experienced English teachers, and revisions were made according to their suggestions. Afterwards, a homogeneous group of three college students were pretested on this questionnaire to detect any unwanted wording ambiguities among the 50 strategic statements. The pretest turned out quite satisfactorily and no further revision on the items was necessary.

With the permission and assistance from four college professors whose classes were to be surveyed, the data collection procedures started. In each class, the participants were informed about the purpose of this survey research in the first place. They were also encouraged to raise any questions about the questionnaire. During the survey, the students worked on their own pace. About fifteen minutes were given to fill out the questionnaire, and the time could be extended if necessary. The questionnaires were then collected for further data analyses. Table 3 summarizes the detailed procedures of data collection.

Table 3 *Procedures of Data Collection*

Stages	Procedures	Time
1	Select four classes from two colleges.	
2	Contact the concerned authorities.	
3	Acquire research permissions from four instructors in charge of these classes.	
4	Conduct survey research on each class.	
	Provide orientation of the study.	5 minutes for each class
	Administer the SILL inventory.	15 minutes for each class
5	Collect the questionnaires on the spot	

Data Analyses

In the present study, there were two independent judges to grade all written responses from the subjects. As mentioned formerly, the five-point Likert scale was used as an indicator of the subjects' frequency of strategy use toward English learning. The descriptive statistics of means, standard deviations, and frequencies were used to signify how often certain strategic behaviors were exhibited by the 125 college participants. In addition, the quantitative analyses on the collected data involved SPSS (Statistics Package for Social Science, version 12.0). Independent-Samples T-test procedures were conducted to examine the effect of major (English vs. non-English majors) on the overall strategies and each of the six types of strategies. The same procedures were also executed to probe the effect of gender (males vs. females) on the overall and individual types of strategy use. All of the statistical tests were set at .05 level of significance.

RESULTS

This section presents the results of the study in the following sequence: (1) general strategy use, (2) effects of major on strategy use, (3) effects of gender on strategy use, and (4) effects of major by gender on strategy use.

General Strategy Use. The frequencies of the six types of strategy use ranged from 2.90 to 3.24 with an average of 3.05 for overall strategy use, as indicated in Table 4. These statistics were all located somewhere around "somewhat true of me" (3 points), which suggested that college students used language learning strategies in a moderate degree. In addition, metacognitive strategies (M=3.24) were used most frequently, followed by compensation strategies (M= 3.17). Relatively, social strategies (M=2.91) and affective strategies (M= 2.90) were used least frequently among college students.

Table 4 *General Language Learning Strategy Use among College Students*

Type	Mean	SD	Rank
Memory strategies	3.01	.45	4
Cognitive strategies	3.07	.57	3
Compensation strategies	3.17	.57	2
Metacognitive strategies	3.24	.70	1
Affective strategies	2.90	.60	6
Social strategies	2.91	.63	5
Overall strategy use	3.05	.47	

Effects of Major on Strategy Use. Tables 5 and 6 present the descriptive statistics of and t-tests on strategy use by English and non-English majors. The mean of overall strategy use by English majors was 3.23 (SD = .44), and non-English majors, 2.89 (SD = .43). An independent-samples t-test procedure was conducted and a significant difference between the two means was detected ($t(123) = 4.38, p = .00 < .05$). English majors significantly used more strategies than their non-English major counterparts. With a closer check, English majors obtained higher means on all of the six strategy types than non-English majors. Via additional t-tests, significant mean differences were found in almost all categories of strategies with memory ($t(123) = 2.12, p = .04 < .05$), cognitive ($t(123) = 5.65, p = .00 < .05$), metacognitive ($t(123) = 3.83, p = .00 < .05$), affective ($t(123) = 2.25, p = .03 < .05$), social ($t(123) = 5.03, p = .00 < .05$), except for compensation ($t(123) = 1.51, p = .14 > .05$). In addition to overall strategy use and specific types of strategy use, certain discrepancies were detected in the ranking of strategy use between the two major groups. The top type used by English majors was metacognitive strategies, and by non-English majors, compensation. Affective strategies were least frequently used by English majors, and social strategies, by non-English majors. In sum, significant differences were detected between the two contrasting groups of EFL learners in terms of overall, specific types, and rank order of strategy use. Major did affect college students' choices of strategies.

Table 5 *Strategy Use by Major*

Type	English majors (N=58)			Non-English majors (N=67)		
	M	SD	Rank	M	SD	Rank
Memory strategies	3.10	.46	5	2.93	.43	3
Cognitive strategies	3.34	.54	2	2.83	.47	4
Compensation strategies	3.25	.49	3	3.10	.63	1
Metacognitive strategies	3.48	.63	1	3.02	.69	2
Affective strategies	3.02	.59	6			
Social strategies	3.19	.61	4	2.79	.58	5
				2.67	.54	6
Overall strategy use	3.23	.44		2.89	.43	

Table 6 *Independent-Samples T-Tests on Strategy Use by Major*

Type	t	df	Sig.
Memory strategies	2.12*	123	.036
Cognitive strategies	5.65*	123	.000
Compensation strategies	1.51	123	.135
Metacognitive strategies	3.83*	123	.000
Affective strategies			
Social strategies	2.25*	123	.026
	5.03*	123	.000
Overall strategy use	4.38*	123	.000

Note. * $p < .05$, two-tailed.

Effects of Gender on Strategy Use. The type most frequently used by both males and females was metacognitive strategies; the type least frequently used by males was memory strategies, and by females, social strategies (see Table 7). The overall frequency of language

learning strategies used by males was 3.12, and by females, 3.04; no significant mean difference was detected between these two groups in their overall strategy use ($t(123) = .66$, $p = .51 > .05$) (see Table 8). In addition, an independent t-test was conducted on each strategy type between the two groups of learners, and no specific significant differences were found for the six individual types (see Table 8). Generally, from the probing of strategy use in overall frequency, specific types and rank order, gender did not affect strategy use.

Table 7 *Strategy Use by Gender*

Type	Males (N=15)			Females (N=110)		
	M	SD	Rank	M	SD	Rank
Memory strategies	2.96	.43	6	3.01	.46	4
Cognitive strategies	3.15	.48	3	3.06	.58	3
Compensation strategies	3.00	.64	5	3.19	.56	2
Metacognitive strategies	3.38	.66	1	3.22	.70	1
Affective strategies	3.04	.73	4			
Social strategies	3.21	.71	2	2.88	.58	5
				2.87	.61	6
Overall strategy use	3.12	.49		3.04	.46	

Table 8 *Independent-Samples T-Tests on Strategy Use by Gender*

Type	t	df	Sig.
Memory strategies	-.47	123	.639
Cognitive strategies	.59	123	.558
Compensation strategies	-1.24	123	.218
Metacognitive strategies	.83	123	.406
Affective strategies			
Social strategies	1.03	123	.307
	1.97	123	.051
Overall strategy use	.66	123	.513

Effects of Major by Gender on Strategy Use. There were four dimensions to examine the interaction effects of major nested within gender and gender nested within major on strategy use: (1) English-major (males vs. females), (2) non-English-major (males vs. females), (3) males (English-major vs. non-English-major), and (4) females (English-major vs. non-English-major), as indicated in Table 9. Four independent-samples t-tests were conducted and no significance mean differences were detected on overall strategy use, except for the group of females (English-major vs. non-English-major): $t(108) = 4.87$, $p = .00 < .05$. Significant mean differences were further found in this female group in almost all individual types with memory ($t(108) = 2.42$, $p = .02 < .05$), cognitive ($t(108) = 6.18$, $p = .00 < .05$), metacognitive ($t(108) = 4.58$, $p = .00 < .05$), affective ($t(108) = 2.44$, $p = .02 < .05$), social ($t(108) = 5.93$, $p = .00 < .05$), except for compensation strategies ($t(108) = 1.19$, $p = .24 > .05$), as shown in Table 10.

Table 9 *Descriptive Statistics of Strategy Use (Major X Gender)*

Major \ Gender	Male (N=15)	Female (N=110)
English-major (N=58)	M=3.15 SD=.57	M=3.25 SD=.44
Non-English-major (N=67)	M=3.10 SD=.45	M=2.86 SD=.41

Table 10 *Effects of Major X Gender on Strategy Use*

Strategy Group	Overall Strategy Use	Significant Type
English-major (males vs. females)	$t(56) = -.93, p = .35 > .05$	None
Non-English-major (males vs. females)	$t(65) = 1.72, p = .09 > .05$	Cognitive strategies ($t(65) = 2.16, *p = .04 < .05$) Metacognitive strategies ($t(65) = 2.36, *p = .02 < .05$) Social strategies ($t(65) = 3.53, *p = .00 < .05$).
Males (English-major vs. non-English-major)	$t(13) = -.21, p = .84 > .05$	None
Females (English-major vs. non-English-major)	$t(108) = 4.87, *p = .00 < .05$	Memory strategies ($t(108) = 2.42, *p = .02 < .05$) Cognitive strategies ($t(108) = 6.18, *p = .00 < .05$) Metacognitive strategies ($t(108) = 4.58, *p = .00 < .05$) Affective strategies ($t(108) = 2.44, *p = .02 < .05$) Social strategies ($t(108) = 5.93, *p = .00 < .05$)

Note. * $p < .05$, two-tailed.

DISCUSSION

General Strategy Use. As what had been predicted earlier, the results indicated that college students used language learning strategies with a moderate frequency, which confirmed those reported in Chang, et al. (2007), Riazi (2007) and Su (2005). In particular, the most frequently used type was found to be metacognitive strategies. This piece of

evidence also echoed Zhang (2005) that Taiwanese college students tended to be metacognitive strategy users. As a matter of fact, metalinguistic awareness is crucial for foreign language learning because it elicits the acts of self-monitoring and self-correction to facilitate the learning process between L1 and L2 (Yu, 2006).

Moreover, social and affective strategies were found to be unpopular among EFL college students, which was congruous with what was reported in Bedell and Oxford (1996), and Chang (1990). Chen (2006) stated “socialization is clearly linked to cultural transmission” (p.169). Oxford (1990) also indicated that cultural or ethnic backgrounds influenced learners’ choice of strategies. Although Asian students are motivated to learn foreign languages, they seldom ask for help from others during the learning process (Bedell & Oxford, 1996; Chang, 1990; Zhang, 2005). Also, the factor of emotion during learning is often neglected by these learners (Bedell & Oxford, 1996; Chang, 1990; Chang, et al., 2007). The finding of social and affective strategies as the least used two types is usually well-expected in the EFL strategy research.

Effects of Major on Strategy Use. Consistent with what was produced in the previous studies (i.e., Chang, et al., 2007; Ong, 2005; Wu, 2007), there was no surprise that students majoring in English significantly employed more strategies than their non-English-major counterparts in terms of overall strategy use. Significant major effects were also detected in all individual types, except for compensation strategies. These results suggested that English majors were, as expected, more proficient strategy users than their non-English-major counterparts. It echoed what was found in Chang, et al. (2007), and Wu and Chang (2005) that humanities and social science college students used more strategies than their business and management counterparts. It is worth noting that compensation strategies were used most often by non-English majors, and metacognitive strategies, by English majors. The result that non-English majors tended to employ more compensation strategies signified they were lower-level strategy users during English learning process. Oxford (1990) stated that compensation strategies viewed as guessing or inferring strategies “involve using a wide variety of clues- linguistic and nonlinguistic- to guess the meaning when the learner does not know all the words” (p.47). In contrast, metacognitive strategies were higher order of cognitive strategies and were found to be frequently used by more proficient language learners (Chen, 2002; Yu, 2006).

Effects of Gender on Strategy Use. Different from what had been predicted earlier, the study found no significant gender differences in strategy use; male and female learners used an equivalent number of strategies in learning a foreign language. Apparently, academic major was a more potent factor than gender on Chinese adults’ choice of language learning strategies. This result of insignificant gender effect contradicted the findings of previous EFL strategy literature which generally indicated that females used more strategies than males: with elementary school students (e.g., Hsu, 2007; Tseng, 2005), junior high students (e.g., Huang, 2004; Lai, 2007; Shen, 2006; Yu, 2006) or senior high students (e.g., Kaylani, 1996; Lo, 2007). However, several studies reported an insignificant gender effect with EFL college students as did the present study (e.g., Huang, 2005; Nisbet, et al., 2005; Wu & Chang, 2005). These findings suggested that the strength of sex differences on strategy use could be moderated or dilated with the growing of age, so significant gender effect would no longer exist among adult strategy users. It appears that age factor has played a critical role on male and female EFL learners’ use of strategies. This inference echoed Oxford (1993), who indicated that factors that affect “L2 students’ choice of learning strategies include: motivation, career/academic specialization, cultural background, nature of task, age, and stage of language learning” (p.19). Generally, gender factor might only influence the strategy use of younger learners rather than that of adult learners. Further strategy research is required to confirm this message.

Effects of Major by Gender on Strategy Use. As for the effects of major by gender, there was a significant difference in the strategy use between English-major and non-English-major females, while such significant major effect was not detected between English-major and non-English-major males. That is, English-major females were more strategic than their non-English-major female counterparts. The effect of major by gender only existed among female college students of differing academic specializations. Since this study found that English majors significantly employed more learning strategies than their non-English-major counterparts, it could be inferred that females were more responsible for the significant major differences found in this study than males. In addition, the literature has consistently found that there is a significant relationship between English proficiency and the frequency of use of language learning strategies (e.g., Chen, 2002; Park, 1997). From this perspective, the proficiency gap between the English-major and non-English-major females could have been much greater than that in the corresponding males, so significant major effect was only detected in the college females in their strategy use. Note that the number of male adults in the present study was relatively low (M=15, F=110), which might limit the generalizability of this piece of finding and its relevant inference as well. It calls for more research on this major by gender issue, so consistent strategic behaviors in college male and female learners of differing majors can be identified.

PEDAGOGICAL IMPLICATIONS

Based on the results of the present study, four pedagogical suggestions for EFL teachers are addressed. Firstly, strategy training should be integrated into the EFL courses, specifically those geared to the acquisition of the four language skills since college students do not use language learning strategies frequently. In particular, social and affective strategies should be introduced in such training to all college EFL students to promote interpersonal interaction and self-efficacy during language learning as they were the least used two types. Secondly, the study found that metacognitive strategies were most frequently used by more effective learners identified as those majoring in English for the significant major effect. This type of strategies should be specially taught to non-English majors to strengthen the behaviors of planning, monitoring and evaluating so as to raise the awareness of their strategy use. Thirdly, although significant gender effect on strategy use was not found in adult language learners, EFL teachers should be aware of its potential effect on younger learners. Lastly, EFL teachers should be cautious about the strategy use among non-English-major females since this specific group was found to use language learning strategies the least.

SUGGESTIONS FOR FURTHER RESEARCH

This study has some limitations that should be taken into deep consideration. Several suggestions are thus made for further research. Since the generalization of the findings was only limited to the college EFL learners in southern Taiwan, future strategy studies may expand the coverage of surveying areas with a greater number of college students to strengthen the generalizability of any potential results. In addition, this study recruited subjects from language, business, and management colleges, and the number of males was not large enough to get a clear picture on their strategy use. EFL learners with male dominance from colleges of science, engineering, etc. could be added to balance the number of males and females for the gender effect and meanwhile, to explore deeper into the effect of academic specializations with adult learners of greater versatility. More intriguing results are more likely to produce. Furthermore, this study centered on examining major and gender effects on strategy use. Other factors such as motivation, personality type, and learning style as noted in Oxford and Nyikos (1989) and Oxford (1993) are worth probing in

the context of EFL. Finally, additional data collection measures of individual interviews and reflective journals may be incorporated into such research to triangulate the methodology to increase the credence of any potential findings.

REFERENCES

- Bedell, D., & Oxford, R. L. (1996). Cross-cultural comparisons of language learning strategies in the People's Republic of China and other countries. In R. L. Oxford (Eds.), *Language learning strategies around the world: Cross-cultural perspectives* (pp. 47-60). University of Hawaii Press.
- Bialystok, E. (1979). The role of conscious strategies in second language proficiency. *Canadian Modern Language Review*, 35, 372-394.
- Brown, A. L., & Palinscar, A. S. (1982). Inducing strategic learning from contexts by means of informed self-control training. *Technical Report*, 262, 1-49.
- Chamot, A. U., & Küpper, L. (1989). Learning strategies in foreign language instruction. *Foreign Language Annals*, 22, 13-24.
- Chang, C. Y., Liu, S. C., & Lee, Y. N. (2007). A study of language learning strategies used by college EFL learners in Taiwan. *Center Bulletin*, 2, 235-262.
- Chang, H. H. (2005). *The relationship between extrinsic/intrinsic motivation and language learning strategies among college students of English in Taiwan*. Unpublished master's thesis, Ming Chuan University, Taipei, Taiwan.
- Chang, S. J. (1990). *A study of language learning behaviors of Chinese at the University of Georgia and the relation of those behaviors to oral proficiency and other factors*. Unpublished doctoral dissertation, University of Georgia, Athens.
- Chen, I. J. (2002). *Language learning strategies used by high and low English proficiency students in a technology college*. Unpublished master's thesis, National Changhua University of Education, Changhua, Taiwan.
- Chen, M. L. (2007). *Vocational high school students' English learning motivation and their English learning behaviors*. Unpublished master's thesis, Tunghai University, Taichung, Taiwan.
- Chen, W. C. (2006). Some literature review on difference of British and Chinese cultural and educational expectations. *Whampoa - An Interdisciplinary Journal*, 51, 169-176.
- Chuang, H. Y. (2007). *The study of foreign language anxiety, English learning motivation and strategies in the elementary school*. Unpublished master's thesis, National Chiayi University, Chiayi, Taiwan.
- Dreyer, C., & Oxford, R. L. (1996). Learning strategies and other predictors of ESL proficiency among African speakers in south Africa. In R. L. Oxford (Eds.), *Language learning strategies around the world: Cross-cultural perspectives* (pp. 61-74). University of Hawaii Press.
- Grainger, P. R. (1997). Language-learning strategies for learners of Japanese: Investigating Ethnicity. *Foreign Language Annals*, 30(3), 378-385.
- Green, J. M., & Oxford, R. L. (1995). A closer look at learning strategies, L2 proficiency, and gender. *TESOL Quarterly*, 29(2), 261-297.
- Hsieh, P. C. (2006). *Junior high school students' English learning strategy use and its relationships with their English learning achievements*. Unpublished master's thesis, Tunghai University, Taichung, Taiwan.
- Hsu, M. Y. (2007). *EFL leaning in the language learning strategy use by elementary school Students in central Taiwan*. Unpublished master's thesis, National Taipei University of Education, Taipei, Taiwan.
- Huang, C. (2004). *A study on the relationships among family educational capital, English learning strategies and English learning achievement of middle school students*.

- Unpublished master's thesis, National Changhua University of Education, Changhua, Taiwan.
- Huang, R. F. (2007). *Comparison of English learning motivation and strategies between elderly students of different genders in southern Taiwan*. Unpublished master's thesis, National Kaohsiung Normal University, Kaohsiung, Taiwan.
- Huang, S. C. (2005). *A study of factors affecting applied English majors' use of language learning strategies*. Unpublished master's thesis, Southern Taiwan University, Tainan, Taiwan.
- Kaylani, C. (1996). The influence of gender and motivation on EFL learning strategy use in Jordan. In R. L. Oxford (Eds.), *Language learning strategies around the world: Cross-cultural perspectives* (pp. 75-88). University of Hawaii Press.
- Khalil, A. (2005). Assessing of language learning strategies used by Palestinian EFL learners. *Foreign Language Annals*, 38(1), 108-119.
- Lai, K. T. (2007). *A study of junior high school students' English acquisitive environment, English acquisitive motivation, English acquisitive strategy, and English acquisitive achievement in Miaoli*. Unpublished master's thesis, National Changhua University of Education, Changhua, Taiwan.
- Levine, A., Reves, T., & Leaver, B. L. (1996). Relationship between language learning strategies and Israeli versus Russian cultural-educational factors. In R. L. Oxford (Eds.), *Language learning strategies around the world: Cross-cultural perspectives* (pp. 35-45). University of Hawaii Press.
- Liao, Y. F. (2000). *A study of Taiwanese junior high school students' EFL learning motivation and learning strategies*. Unpublished master's thesis, National Changhua University of Education, Changhua, Taiwan.
- Lin, T. C. (2006). *EFL technological university students' use of language learning strategies and beliefs about language learning*. Unpublished master's thesis, Southern Taiwan University, Tainan, Taiwan.
- Liu, M. L. (2004). *A study of beliefs about language learning and learning strategies of senior high school students*. Unpublished master's thesis, National Kaohsiung Normal University, Kaohsiung, Taiwan.
- Lo, P. C. (2007). *A study of language learning strategies used by comprehensive high school students*. Unpublished master's thesis, Southern Taiwan University, Tainan, Taiwan.
- Naiman, N., Fröhlich, M., Stern, H. H., & Todesco, A. (1978). The good language learner. *Research in Education Series*, 7. The Ontario institute for Studies in Education, Toronto.
- Nisbet, D. L., Tindall, E. R., & Arroyo, A. A. (2005). Language learning strategies and English proficiency of Chinese university students. *Foreign Language Annals*, 38(1), 100-107.
- Okada, M., Oxford, R. L., & Abo, S. (1996). Not all alike: Motivation and learning strategies among students of Japanese and Spanish in an exploratory study. In R. L. Oxford (Eds.), *Language learning motivation pathways to the new century* (2nd ed., pp.105-119). University of Hawaii Press.
- O'Malley, J. M., Chamot, A. U., Stewner-Manzanares, G., Russo, R. P., & Küpper, L. (1985). Learning strategy applications with students of English as a second language. *TESOL Quarterly*, 19(3), 557-584.
- Ong, S. Y. (2005). *English language learning strategies used by the sophomore students of a technology university*. Unpublished master's thesis, Southern Taiwan University, Tainan, Taiwan.
- Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. Boston: Heinle & Heinle.
- Oxford, R. L. (1993). Language learning strategies in a nutshell: Update and ESL suggestions.

- TESOL Quarterly*, 3, 18-22.
- Oxford, R. L. (2003). Language learning strategy profiles of elementary school students in Taiwan. *International Review of Applied Linguistics in Language Teaching*, 41(4), 339-379.
- Oxford, R. L., & Burry-Stock, J. A. (1995). Assessing the use of language learning strategies worldwide with the ESL/EFL version of the Strategy Inventory for Language Learning (SILL). *System*, 23(1), 1-23.
- Oxford, R. L., Lavine, R. Z., Felkins, G., Hollaway, M. E., & Saleh, A. (1996). Telling their stories: Language students use diaries and recollection. In R. L. Oxford (Eds.), *Learning strategies around the world: Cross-cultural perspectives* (pp. 19-34). University of Hawaii Press.
- Oxford, R. L., & Nyikos, M. (1989). Variables affecting choice of language learning strategies by university students. *Modern Language Journal*, 73(2), 291-300.
- Park, G. P. (1997). Language learning strategies and English proficiency in Korean university students. *Foreign Language Annals*, 30(2), 211-221.
- Peng, I. N. (2002). *EFL motivation and strategy use among Taiwanese senior high school learners*. Unpublished master's thesis, National Taiwan Normal University, Taipei, Taiwan.
- Riazi, A. (2007). Language learning strategy use: Perceptions of female Arab English majors. *Foreign Language Annals*, 40(3), 433-440.
- Rubin, J. (1981). Study of cognitive processes in second language learning. *Applied Linguistics*, 2, 117-131.
- Shen, M. C. (2006). *The relationship between beliefs about language learning and learning strategy use of junior high school students in remote districts*. Unpublished master's thesis, Leader University, Tainan, Taiwan.
- Shih, H. M. (2003). *The study of learning belief and strategy use on EFL junior high school students in Taiwan*. Unpublished master's thesis, National Changhua University of Education, Changhua, Taiwan.
- Su, M. H. (2005). A study of EFL technological and vocational college students' language learning strategies and their self-perceived English proficiency. *Electronic Journal of Foreign Language Teaching*, 2(1), 44-56.
- Tran, T. V. (1988). Sex differences in English language acculturation and learning strategies among Vietnamese adults aged 40 and over in the United States. *Sex Role*, 19, 747-758.
- Tsai, L. C. (2006). *An investigation of learning style preferences in relation to language learning strategies*. Unpublished master's thesis, Ming Chuan University, Taipei, Taiwan.
- Tseng, S. W. (2005). *The relationship between foreign language anxiety and English learning strategy of elementary school students*. Unpublished master's thesis, National Taichung University, Taichung, Taiwan.
- Wharton, G. (2000). Language learning strategy use of bilingual foreign language learners in Singapore. *Language Learning*, 50(2), 203-244.
- Wu, C. C. (2003). *A study of EFL learning strategies used by vocational high school students in Taiwan*. Unpublished master's thesis, National Yunlin University of Science and Technology, Yunlin, Taiwan.
- Wu, C. D. (2007). *English learning strategy uses by the students of southern Taiwan University of Technology*. Unpublished master's thesis, Southern Taiwan University, Tainan, Taiwan.
- Wu, C. J., & Chang, S. (2005). Learning strategies: A study of Yu Da college students. *Yu Da Academic Journal*, 9, 21-41.
- Yang, P. L. (1993). *Learning strategies and techniques among high-achieving students of English as a foreign language in ROC's high schools*. Unpublished master's thesis, National Kaohsiung Normal University, Kaohsiung, Taiwan.
- Yu, H.C. (2006). *Motivation and learning strategy use among junior high school students*

with different levels of academic achievement. Unpublished master's thesis, National Pingtung Institute of Commerce, Pingtung, Taiwan.

Zhang, C. (2005). The study of language learning strategies of non-English majors. *Sino-US English Teaching*, 2(5), 36-41.

English Collocation Competence of College Students and Learner Factors

Mei-hua Chen (陳政樺)
National Tsing Hua University
chen.meihua@gmail.com

As collocation competence has attracted EFL scholars' attention in Taiwan, the study addressed the collocation competence of advanced EFL learners and the problems they encountered. Most previous studies concentrated on V + N combinations of lexical collocations mainly; this study examined both lexical collocations and grammatical collocations. Moreover, different from previous studies, this study explored the impact of learners' prior language learning experiences on collocation competence.

440 first-year non-English major college students from three universities participated in this study. By means of a 50-item multiple-choice collocation test and a questionnaire, the study explored two issues: a significant positive correlation between the participants' achievement in the collocation test and their English subject scores on the College Entrance Examination, suggesting that collocation competence is an important aspect of language proficiency; learning experiences had significant effects on collocation competence. For example, early start of English learning, the experience and the frequency of extracurricular English learning, English-speaking environment, appear to be crucial factors in benefiting collocation competence or even in developing language proficiency. These results were consistent with the participants' opinions in the questionnaire and their scores on the collocation test.

The study suggested that collocation dictionaries are highly recommended, and memorizing chunks or phrases may accumulate and consolidate EFL learners' collocation knowledge. Further, EFL teachers are suggested to introduce their students to the concept of collocation and learners are encouraged to concentrate their attention on proper word combinations.

INTRODUCTION

Most EFL learners agree that they have experienced difficulty in selecting appropriate word combinations when producing writing or utterance even though they have memorized a large vocabulary. It appears that lacking awareness of the importance of collocation causes EFL learners' failure to express their thoughts properly (Hill, 2000). For EFL learners, unfortunately, words do not co-occur freely; instead, there are various co-occurrence restrictions on word combinations which are highly predictable for native speakers (Allerton, 1984). For instance, we say "*set the table*" but "*make the bed*". Consequently, EFL learners are encouraged to pay more attention to collocational usages (Zughoul & Hussein, 2001).

This study attempts to identify the relationship between collocation competence and overall English proficiency and reveal how learning experience, including learning background and learning styles, influences EFL learners' collocation competence.

RESULTS AND DISCUSSION

A 50-item multiple-choice collocation test, consisting of verb, adjective and preposition subtests, and a questionnaire were administered to a total of 440 first-year non-English major students recruited from three universities (NU, NUST, and PU) to investigate the participants' receptive knowledge of collocation. The results of statistic analysis concerning the relationship of collocation competence among overall English proficiency and learning experience are discussed as follows.

The relationship between collocation competence and overall English proficiency

This issue explores the relationship between the college students' collocation competence and their English subject scores on the College Entrance Examination (CEE), and evaluates the overall performance of the participants across these three universities.

Correlations between Collocation Competence and English Subject Scores on CEE

The current college entrance examinations consist of three types: Subject Competence Test (SCT), Assigned-Subject Test (AST) for enrolling in universities and Joint-Entrance Examination (JEE) for entering Technical universities or colleges.

Table 1 shows that there exists a positive correlation between the participants' collocation competence and their achievement in the English subject on CEE that measures the general English proficiency. It can be inferred that that collocation competence is a critical aspect of language proficiency. The findings are consistent with those of Bonk (2000), Chen (2002), Lin (2001), Al-Sibai (n.d.).

Table 1. *Correlation between Participants' Collocation Competence and Their English Subject Scores on the College Entrance Examination*

	Verb subtest	Adjective subtest	Preposition subtest	Overall test
SCT	.711**	.573**	.714**	.785**
AST	.736**	.599**	.737**	.818**
JEE	.132	.198	.257	.318*

Note. **. Correlation is significant at the 0.01 level (2-tailed).

Participants' Overall Performance in the Collocation Test

This subsection not only evaluates the overall performance of the participants across universities in the collocation test but also explores the challenging lexical categories for the participants.

Figure 1 illustrates that the participants from NU achieved almost 60 percent of the scores, whereas the participants from NUST and PU did not perform as well as the participants from NU. However, the results indicate that participants did not have sufficient collocation knowledge. Even if the level of vocabulary in the collocation test was quite basic, the participants did not achieve the mean score at a rate of 50 percent.

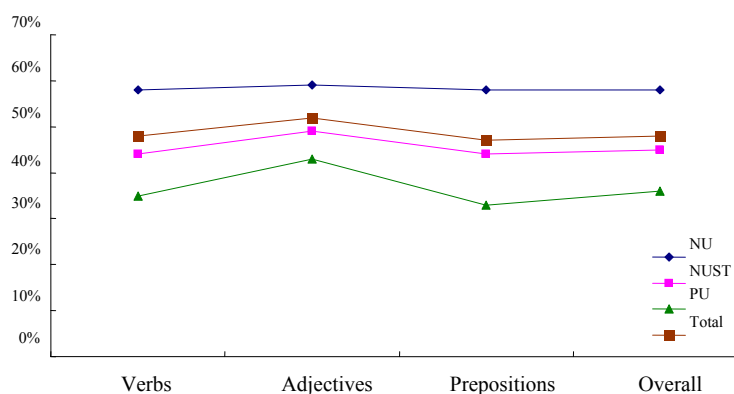


Figure 1. Line graph of overall performance of participants across universities in the collocation test.

Concerning lexical categories, most participants felt that prepositions were the most difficult lexical category and verbs were also quite difficult. As shown in Figure 2, the participants' opinions seem to be consistent with their scores on the collocation test.

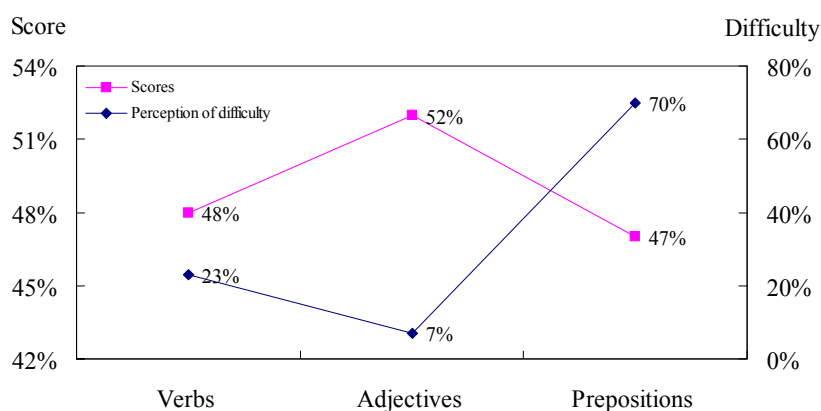


Figure 2. The comparison between participants' collocation test scores and their perception of difficulty on three lexical categories.

The relationship of the participants' scores on the collocation test with their learning experience

The second issue explores the relationship of the participants' scores on the collocation test with their learning background and learning styles respectively.

Correlations between Collocation Competence and Learning Background

Table 2. *Effects of Learning Background on Collocation Test*

Question No.	Items	Significantly related
Q1	The starting age of English learning	✓(negative)
Q2	The experience & the frequency of extracurricular English learning	✓
Q3	The experience of having stayed in English-speaking countries	✓

Note. "✓" indicates the correlation is significant, whereas "×" indicates the contrary results.

As shown in Table 2, there existed significant relationships between learning background and collocation competence. These questions are discussed in more detail as follows.

First, more than half (56%) of the participants started English learning after grade five whereas the other half began English learning before grade four. A negative relationship between the starting age of English learning and collocation competence indicates that the earlier one starts learning English, the better his collocation proficiency can be expected. Some empirical studies (e.g., Lin, 2001; Kuo, 2002; Chen, 2004; Ho, 2004) have also proved that the age effect exerts a positive impact on English attainment. Next, the participants with experience of extracurricular English learning ($M = 24.48$) outperformed those without such experience ($M = 20.13$). The frequency of extracurricular English learning was also found to make a contribution to collocation competence. In other words, extracurricular English learning did produce positive effects on the performance of participants' collocation test. Moreover, having stayed in English-speaking countries was demonstrated to facilitate students' collocation competence. It seems reasonable that in an English-speaking

environment, EFL learners may have a better chance of communicating with native speakers. According to Bonk (2000), the interaction with native speakers did facilitate the acquisition of collocations.

Correlations between Collocation Competence and Learning Styles

Table 3. *Effects of Learning Styles on Collocation Test*

Question No.	Items	Significantly related
Q4	The preference for dictionaries	✓
Q5	Awareness of word combinations when looking up words in dictionaries	✓
Q6	Memorizing phrases or chunks when learning new English words	✓
Q7	L1 reliance in L2 communication	✓(negative)

Note. “✓” indicates the correlation is significant, whereas “✗” indicates the contrary results.

Learning styles has a strong correlation with collocation competence (Table 3). It is apparent that EFL learners gained significant benefit from the dictionaries with English explanations and examples because they provide real world usages of words. In addition, the habit of being aware of collocations and memorizing phrases or chunks can practically facilitate the productive knowledge of words. Tseng (2002), Tang (2004), Mahmoud (2005) and Wang (2005) declared the importance of dictionaries. Lin (2001) also pointed out that high achievers were more aware of the collocations. Further, participants who had heavier reliance on L1 got lower scores on the collocation test. The results are consistent with earlier research, such as the studies of Tseng (2002), Nesselhauf (2003), Tang (2004), Mahmoud (2005) and Wang (2005). As a result, EFL learners are encouraged to pay more attention to collocation instead of relying on L1 translation when learning English.

CONCLUSIONS

The findings of the study reveal that collocation competence is part of overall language proficiency. Additionally, the empirical evidence in the study suggests that learning experience has a significant impact on collocation competence.

Based on the findings of this study, some pedagogical implications are suggested for English teachers and learners. First, dictionaries play a crucial role in language learning. English-English dictionaries with numerous examples and collocation dictionaries, such as *The BBI Dictionary of English Word Combination*, *Collins Cobuild English Collocations*, *Oxford Collocations Dictionary for Students of English*, and *A Dictionary of English Collocations* (Wang, 1997) may facilitate collocation competence. Further, computerized tools such as corpora and concordancers are highly recommended. For example, *British National Corpus*, *Cobuild Concordance and Collocations Sampler* and *Tango* (Jian, 2004), a collocational concordancer developed by Institute of Information Systems and Applications, National Tsing Hua University, provide analytical and effective methods to benefit EFL learners' English proficiency. Then, memorizing chunks or phrases may help EFL learners think directly in the target language rather than rely on L1 transfer, and then reduce collocation errors due to L1 interference (Nattinger, 1980; Sinclair, 1987; Lewis, 1993; Brown, 2001). Moreover, specific instructions may facilitate EFL learners' collocation knowledge. Many researchers such as Tseng (2002), Martyńska (2004), Mahmond (2005), Wang (2005) Keshavarz and Salimi (2007) have stated the importance of collocation teaching.

Moreover, Lewis (1993), Ellis (1997) and Nation (2001) pointed out that teachers should introduce their learners the concept of collocation and concentrate learners' attention on chunks instead of analyzing the words in the target language because words stored in chunks are retrieved more quickly and accurately. Further, Liu (1999) advised that the instruction on differentiating the usages of synonyms should be emphasized. On the other hand, learners should be encouraged to develop good habits of checking collocational usages by consulting collocation dictionaries, and take notes systematically, with the result that they may increase the quantity of collocations (Hsueh, 2002).

On the other hand, some limitations of the study should be pinpointed while suggestions are offered for future research. Future research should be conducted on a larger scale. More participants should be recruited from universities in different areas in Taiwan to further explore the collocation competence of university students in Taiwan. In addition, larger sample size may explore various sources of collocation errors. Moreover, some other collocation categories may be investigated in future research. On the other hand, the multiple-choice collocation test in the study only investigated the participants' receptive knowledge of collocation. In addition to EFL learners' receptive knowledge of collocation, the productive knowledge of collocation need testing in the future research. Furthermore, additional instrument such as interviews can be adopted in future research to lead to a better understanding of the participants' perceptions of collocation.

REFERENCES

- Allerton, D. J. (1984). Three (or four) levels of word co-occurrence restriction. *Lingua*, 63, 17-40.
- Al-Sibai, D. (n.d.). *The development of English collocation competence*. Unpublished manuscript. Retrieved April 23, 2008, from <http://faculty.ksu.edu.sa/dinaalsibai/Research%20Papers/04.%20Collocation%20Review.pdf>
- Bonk, W. J. (2000). *Testing ESL learners' knowledge of collocations*. (ERIC Document Reproduction Service No. ED 442309)
- Chen, P.C. (2002). *A corpus-based study of the collocational errors in the writings of the EFL learners in Taiwan*. Unpublished master's thesis, National Taiwan Normal University, Taiwan, R.O.C.
- Chen, Y. P. (2004). *Investigation of 7th graders' EFL proficiency discrepancy in relation to their family background*. Unpublished master's thesis, National Kaohsiung Normal University, Taiwan, R.O.C.
- Ellis, R. (1997). *SLA research and language teaching*. New York: Oxford University Press.
- Hill, J. (2000). Revising priorities: From grammatical failure to collocational success. In Lewis, M. (Ed.). *Teaching collocation: Further developments in the lexical approach* (pp. 47-69). London: Language Teaching Publications.
- Ho, H. L. (2004). *Effects of pre-junior high English learning experience on students' English achievement: From the perspective of instructions in the private children language institute*. Unpublished master's thesis, National Tsing Hua University, Taiwan, R.O.C.
- Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics*, 19, 45-72.
- Hsueh, S. C. (2002). *An analysis of lexical collocational errors in the English compositions of senior high school EFL students*. Unpublished master's thesis, National Kaohsiung Normal University, Taiwan, R.O.C.
- Jian, J. Y., Chang, Y. C., & Chang, J. S. (2004) TANGO: Bilingual collocational concordancer. Post & demo in ACL 2004, *the 42th Annual Meeting of Association for Computational Linguistics*, Barcelona.

- Keshavarz, M. H. & Salimi, H. (2007). Collocational competence and cloze test performance: a study of Iranian EFL learners. *International Journal of Applied Linguistics*, 17(1), 81-92.
- Kuo, Y. S. (2002). *The effects of age on Taiwanese EFL learners' long-term English proficiency. The Proceedings of the Eleventh International Symposium on English Teaching*, 313-332.
- Lewis, M. (1993). *The lexical approach: The state of ELT and a way forward*. Hove, England: Language Teaching Publications.
- Lin, Y. P. (2001). *The Effects of collocation instruction on English vocabulary development of senior high students in Taiwan*. Unpublished master's thesis, National Kaohsiung Normal University, Taiwan, R.O.C.
- Liu, C. P. (1999). An analysis of collocational errors in EFL writing. *Proceedings of the Eighth English International Symposium on English Teaching*, 483-494.
- Mahmoud, A. (2005). Collocation errors mad by Arab learners of English. *The Professional Teaching Articles Journal*, 5. Retrieved April 6, 2008, from http://www.asian-efl-journal.com/pta_August_05_ma.php
- Martyńska, M. (2004). Do English language learners know collocations? *Investigationes Linguisticae*, XI.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nattinger, J. R. (1980). A lexical phrase grammar for ESL. *TESOL Quarterly*, 14(3), 337-344.
- Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 24, 223-242.
- Sinclair, J. (1987). Collocation: a progress report. In R. Steele & T. Threadgold (Eds.), *Language topics: Essays in honour of Michael Halliday* (pp.319-331). Amsterdam : John Benjamins.
- Tang, Y. T. (2004). *A Study of the collocation errors in the oral and written production of the college students in Taiwan*. Unpublished master's thesis, National Taiwan Normal University, Taiwan, R.O.C.
- Tseng, F. P. (2002). *A Study of the effects of collocation instruction on the collocational competence of senior high school students in Taiwan*. Unpublished master's thesis, National Taiwan Normal University, Taiwan, R.O.C.
- Wang, H. C. (2005). *The relationship between EFL learners' depth of vocabulary knowledge and oral collocational errors*. Unpublished master's thesis, National Kaohsiung Normal University, Taiwan, R.O.C.
- Wang, W. C. (1997). *A dictionary of English collocations*. Taipei: Zhong Yang.
- Zughoul R. & Hussein, A. (2001). *Collocational competence of Arabic speaking learners of English: A study in lexical semantics*. (ERIC Document Reproduction Service No. ED479650)

A Comparison of Two Assessment Instruments of English Oral Performance⁹

Yuh-mei Chen (陳玉美)
National Chung Cheng University
folcym@ccu.edu.tw

This study aims to compare two instruments of English oral performance in the EFL university classrooms by examining the correlation and difference between student and teacher assessments. Two cohorts of English majors in a national university participated in this study. They were asked to assess their own and peers' English oral performances with two instruments in different sequences. One instrument (category-based) is designed on 4 subscales (including content, language, delivery, and manner); the other (item-based) has 20 items subsumed under 4 criteria. Data included student and teacher ratings, written feedback on the assessment sheets as well as students' perceptions of the two assessments on a 5-point scale questionnaire.

Results showed that student and teacher ratings generated by the two different instruments were positively correlated. The item-based instrument generated lower mean scores and scores spread out over a wider range, while the category-based instrument elicited more feedback comments. Significant difference was found between the yielded ratio of students' commentaries and the approved ratio for rating in both assessment instruments. A significant majority of students preferred using the item-based instrument because of its specificity, but some noted that assessing oral performances via discrete items distracted their attention from observing performances. This study demonstrated that both category-based and item-based instruments, if designed appropriately, are valid and reliable tools for student assessment, but the category-based helped students gave more feedback.

INTRODUCTION

When it comes to assessing classroom performances, there are by and large two major types of methods: holistic rating and analytical rating. Holistic rating focuses on a performance as a whole on a single scale in which one score is awarded while analytical rating breaks the performance into separate categories (subskills) and each category of the performance is assessed independently on a sub-scale. Holistic rating may be relatively fast and easy in practice; however, holistic judgment is regarded as less reliable than analytical evaluation (Davies et al., 1999; Weigle, 2002). In a study comparing the two rating approaches to writing assessment, Nakamura (2002) recommends that analytical assessment with several evaluation items should be used to avoid risky idiosyncratic ratings. She also suggests that the best choice is to have multiple raters and multiple rating items; the second best is to have one overall evaluation item and multiple raters; the worst situation is to have one rater and an impressionistic scale.

In the classroom setting, setting clear criteria for assessment is the first and foremost to ensure fairness and accuracy of scoring performance tasks. Airasian (1997) argues that clear criteria should be articulated to students in advance. As students are aware of criteria for success, the assessment is a valid assessment. Stiggins (2001) also urges teachers to identify and discuss performance criteria with students. When they clarify their expectations and explicitly state the desired performance outcomes to students, students' standards of

⁹ Project funded by National Science Council (NSC 97-2410-H-194-072-). Initial findings are presented in this paper.

achievement are likely to be raised. Given that assessment is moving towards partnership in the classroom context, criteria are often developed by the teacher and students collaboratively. A typical example is Falichov's (1986) model of self, peer, and tutor collaborative assessment, in which the teacher first generates a set of criteria of excellent essays, students then construct their criteria as individuals and in groups, and after comparisons and discussions about teacher/student generated criteria, a list of agreed criteria is decided for assessment. Involving students to construct their own marking criteria, Orsmond, Merry, and Reiling (2000) have observed more spontaneous and active student discussions.

Classroom performance assessment is different from general language proficiency testing. What to be looked at in a performance depends on the learning goal and the task nature. According to Davies et al. (1999), "The levels or bands are commonly characterized in terms of what subjects can do with the language (tasks and functions which can be performed) and their mastery of linguistic features (such as vocabulary, syntax, fluency and cohesion)." (p. 154) Luoma (2004) alerts us to the danger that scores become meaningless when criteria are developed in isolation with the tasks. That is, rating scales should be task-specific. In an individual description task, criteria such as comprehensibility, discourse organization, accuracy, and fluency might be taken into consideration. But a discussion task might be scored for interaction skills, appropriateness and vocabulary range. In peer assessment, Luoma especially notes that "linguistic criteria may not be suitable...whereas task-related criteria may prove more effective." (p. 189)

Speaking of oral performance/presentation evaluation, criteria used in previous studies include content, organization, language, delivery, manner, and interaction (AlFallay, 2004; Chen, 2005, 2006, 2007; Cheng and Warren, 2005; Patri, 2002). For example, AlFallay (2004) and Patri (2002) asked their students to evaluate their own and peers' performances with six features 14 items on a 5-point scale. Cheng and Warren (2005) used a 5-point scale of 20 items to ask their students to evaluate peers' oral presentations. Chen (2005, 2006) however asked her students to look at only four categories during peer and self assessment. These studies, using either item-based or category-based instruments, yielded conflicting results. Cheng and Warren (1999) reported student awarded marks varied in reliability coefficients and did not agree with teacher awarded marks. But Patri (2002) noted that given specific assessment criteria and peer feedback, students assessed their peers in a manner identical to their teachers. Also, AlFallay (2004) found high correlations between peer, self-, and teacher assessment. Although more items increases instrument reliability, Chen's findings revealed positive effect of the category-based instrument; student assessment resembled teacher assessment.

To date little has been explored regarding the use of different instruments for EFL students' assessment of their oral performances. Speaking of specificity of the assessment tool, several studies have been conducted in other classrooms. For instance, in teacher education, Schechtman (1992) and Schechtman and Godfried (1993) claimed that holistic peer assessment results in closer agreement between peers than category-based assessment. In the course of Business Systems Analysis, Lejk and Wyvill (2001) also observed that a holistic approach is more appropriate than a category-based approach in summative peer assessment though the latter is useful for formative feedback. In engineering education, van Duzer and McMartin (2000) found better scoring distribution on peer and self assessment by removing individual items with restricted ranges and high end scores. However, comparing two assessment instruments in the course of Physical Therapy—one with global components of performance; the other with discrete components, Miller (2003) found that the instrument with more specific criteria produces better discrimination of scoring results at the expense of students' qualitative feedback. Reports on instrument specificity and sensitivity have shown mixed results and the topic merits more research effort. The dearth of knowledge in

assessment specificity and EFL students' learning thus builds the ground for the present study.

METHODOLOGY

Purpose

This study aimed to compare two instruments (category-based or item-based) of English oral performance in the EFL university classroom by investigating the correlation and difference between student and teacher assessments. Both instruments require analytical rating. But one looks at 4 categories of oral performance; the other includes 20 discrete items derived from the 4 global categories.

Participants

The participants were two cohorts of English majors taking the course of *English Oral Training* from a national university in southern Taiwan. Cohort 1 was freshmen; Cohort 2 was sophomores and had more experience in peer and self-assessment. When asked to assess their own English skills and learning attitude on a 5-point scale, both cohorts rated themselves below 3, the hypothetical mean score in English proficiency, including speaking and listening, but above 3 in their attitude toward English learning.

Instruments

The instruments used in this study were two different evaluation forms: category-based (Appendix A) and item-based (Appendix B). The former includes four categories—content (30%), language (30%), delivery (30%), and manner (10%); the scoring standards contain five levels-- *excellent*, 90% and above, *good*, 80%-89%, *fair*, 70%-79%, *ok*, 60%-69%, and *poor*, below 60%. The latter consists of 20 discrete items subsumed under the four categories, each item on a 5-point scale. Using the same evaluation form, all the assessors (including students themselves, peers, and the teacher) evaluated performances by giving analytical marks and written comments over two cycles of assessment in the study.

Another type of instruments used in the study was the pre-study and post-study questionnaires. The pre-study questionnaire elicited students' background information as well as their self-ratings of English proficiency. The post-study questionnaire also asked students to self-assess their English proficiency and inquired their perceptions of the two different assessment instruments on a 5-point scale.

Procedure

The assessment task was required and completed in a course of *Freshman English Oral Training*, meeting two hours each week. The assessment procedure was based on Chen's (2007), including the following steps: training, observation, evaluation, discussion, feedback, and response.

In the first two weeks of the class, students received a training of assessment. At the first class meeting, students were divided into groups by their preference and informed of the assessment tasks required for the course. Then students and the teacher discussed jointly about what to be noted in an oral presentation. The class concluded a set of criteria for good oral performances and scoring standards. In the second meeting, students practiced assessing two videotaped performance against the marking criteria. They discussed their observations with their group members and share evaluation results in class. The teacher then commented on the videos to demonstrate her evaluation of the performances and comment on students' evaluations. The training session was to orient students with the assessment criteria and procedures.

Oral performance and assessment proceeded after the training. In a cycle of 4-5 weeks (depending on the number of groups), individual students in each group had a turn speaking about a self-chosen topic for about 5 minutes. Assessment was conducted immediately after the performance. Using the same evaluation form, the teacher, peers and students themselves

were given 3-5 minutes to give analytical marks and write evaluations for each performance. Since the course also intended to profile students' learning from some other activities such as teacher-student interview, pair talk, and group projects, the assessment cycle ran twice in a semester.

After the scheduled performances, oral feedback was given by peers and the teacher. Students first exchanged comments or suggestions, reflections and observations within groups. Then group speakers and the teacher provided feedback to the performers one by one. Student evaluation results were turned in after the feedback session.

Data Collection and Analysis

Data of the study were collected from two assessment instruments which demand student and teacher scorings and written comments as well as pre-study and post-study questionnaires. Students' in-class oral performance and feedback were also videotaped and used to triangulate the results of written feedback analysis.

Scoring distributions were presented with descriptive statistics. Correlations between student and teacher assessments in two cycles of assessment were analyzed to examine which instrument enhances scoring reliability. Written feedback was classified and coded into types (positive, negative, and neutral) and topics (content, language, delivery, manner, and others) by two raters with the constant comparative method. The classification of feedback types is based on Falchikov's (1996) taxonomy: positive feedback refers to the statements identifying strengths, negative feedback, the statements of weaknesses, and neutral feedback, prompts to reflection and suggestions for improvement. Categories of feedback were tabulated and presented with the number of statement per assessment, frequencies and percentages. T-test was used to detect difference between student and teacher assessments in the amount of feedback; chi square test was conducted to examine difference in feedback type and topic. Students' questionnaire responses toward the assessment instruments were categorized into likes and dislikes.

RESULTS AND DISCUSSION

Comparison of Category-based and Item-based Assessment

Table 1 summarizes the comparisons of teacher's and students' assessments with two different evaluation forms. When the item-based assessment instrument was used, in two cohorts generally the scores were spread out wider and averaged lower. The standard deviation values were greater in item-based assessment than in category-based assessment, indicating that scores obtained from the item-based assessment form varied more over a wider range. And the mean scores given by different assessors using the item-based form were lower too. Since the item-based assessment form was designed with 20 subscales and the category-based assessment, 4 subscales, more points could have been deducted in the assessment procedures.

Table 1 Comparisons of teacher's and students' category-based and item-based assessments

Cohort 1 (N=21)		Category-based					Item-based				
	Mean	SD	SE	Max	Min		Mean	SD	SE	Max	Min
Teacher	85.33	5.85	1.30	95	74	Teacher	83.19	9.85	2.15	97	62
Students	83.46	3.74	0.84	90.32	77.46	Students	78.19	6.49	1.42	89.95	69.2
Peers	81.19	5.64	0.85	93	71	Peers	72.86	10.21	2.22	90	69
Self	83.76	3.80	1.29	90	77	Self	78.456	6.75	1.47	91	45
Cohort 2 (N=14)		Item-based					Category-based				
	Mean	SD	SE	Max	Min		Mean	SD	SE	Max	Min
Teacher	88.28	6.64	1.78	96	76	Teacher	87.43	6.66	1.78	96	76
Students	83.30	5.91	1.58	91.5	74.29	Students	85.75	4.36	1.17	91,92	78.21
Peers	83.80	6.01	1.61	91	74	Peers	86.05	4.55	1.22	93	78
Self	77	9.17	2.45	93	66	Self	82.36	6.62	1.77	92	70

*p<.05

As shown in Table 2, correlations of students' ratings with the teacher's ratings in both cohorts with both category-based and item-based evaluation forms were all significant. But r values were higher in the item-based assessment. The z-test results further confirmed that in Cohort 1 the strength to predict holistic teacher ratings with student ratings was greater in the item-based assessment, especially in the aspect of language. The result explained that student ratings were more likely to move in the same direction with teacher ratings during the learning process though in the item-based assessment they became less lenient in rating. But in Cohort 2, no significant difference between category-based and item-based assessments was found although student and teacher ratings were highly correlated. A very possible reason was that about 80% of students in Cohort 2 had prior experience in assessing peers' and own English speaking performances, so their ratings were very much like the teacher's. Using either assessment instrument made no difference in their judgment. But Cohort 1 students, not having much experience in peer and self-assessment, might have benefited more from the assessment task itself and using the item-based assessment instrument. They seemed to become more mature and experienced in assessment by following the item sequence and gave scores identical to the teacher's. That is, the interplay of instrument specificity and assessment experience could provide support to inexperienced student raters in the assessment task and make their ratings more identical to the teacher's, especially in the aspect of language as shown in Cohort 1.

Table 2 *Correlations of student and teacher ratings in Cohort 1 (N=21) and Cohort 2 (N=14)*

Cohort 1	1 st Student vs. Teacher (Category-based)	2 nd Student vs. Teacher (Item-based)	z-test	p (1-tailed)
Holistic	0.64*	0.87*	-1.75	0.040*
Analytical				
Content	0.63*	0.86*	-1.62	0.053
Language	0.68*	0.93*	-2.44	0.007*
Delivery	0.72*	0.76*	-0.32	0.375
Manner	0.67*	0.67*	-0.03	0.488
Cohort 2	2 nd Student vs. Teacher (Category-based)	1 st Student vs. Teacher (Item-based)	z-test	p (1-tailed)
Holistic	0.81*	0.84*	-0.22	0.41
Analytical				
Content	0.75*	0.62*	0.58	0.28
Language	0.83*	0.78*	0.33	0.37
Delivery	0.75*	0.83*	-0.5	0.31
Manner	0.80*	0.85*	-0.37	0.36

*P<.05

In addition to compare how much students' ratings were identical to the teacher's in the two assessment forms, it is necessary to look at feedback quantity and quality as well. Table 3 presents t-test results of averaged statements of student and teacher feedback in category-based and item-based assessments from both cohorts. Significantly the teacher gave each student more feedback statements than student raters (3.56 vs. 2.78, 6.21 vs. 3.54, 3.61 vs. 2.24, 4.21 vs. 2.61), and the category-based assessment form elicited more feedback statements from all assessors than the item-based assessment form. In Cohort 1, t-test results of students' feedback to each perform showed that students gave more feedback comments in category-based assessment than in item-based assessment ($t=8.752$, $p=0$). In Cohort 2, it had the similar result ($t= 6.562$, $p= 0$). But using the item-based assessment, students gave as many feedback statements on language and manner as the teacher (no significant difference was found). Although the similar quantity could not explain the feedback content, it showed that the item-based assessment form were likely to help student assessors look at each item and use its subscale (5-point scale) closely and give identical amount of interest to the designated areas like the teacher.

Table 3 *Students' and teacher's feedback to each performer in category-based and item-based assessments*

Cohort 1 (N=21)		Category-based				
	Students (to each performer)	SD	Teacher (to each performer)	SD	t	p (2-tailed)
Content	0.76	0.15	0.91	0.11	-3.74	0.0006*
Language	0.34	0.02	0.92	0.17	15.79	0*
Delivery	0.93	0.18	0.81	0.25	1.83	0.0761
Manner	0.45	0.05	0.66	0.30	-3.16	0.0048*
Others	0.30	0.18	0.27	0.18	0.576	0.568
Total	2.78	0.15	3.56	0.27	-11.547	0*
Cohort 2 (N=14)		Category-based				
	Students (to each performer)	SD	Teacher (to each performer)	SD	t	p (2-tailed)
Content	1.17	0.28	1.86	0.339	-5.897	0*
Language	0.60	0.49	1.86	0.339	-7.912	0*
Delivery	1.04	0.26	1.71	0.327	-5.983	0*
Manner	0.51	0.5	0.79	0.407	-1.625	0.1165
Others	0.23	0.42	0	0	2.044	0.969127
Total	3.54	0.44	6.21	0.497	-15.126	0*
Cohort 1 (N=21)		Item-based				
	Students (to each performer)	SD	Teacher (to each performer)	SD	t	p (2-tailed)
Content	0.72	0.17	1.27	0.52	-4.607	0.0001*
Language	0.75	0.15	0.91	0.52	-1.381	0.1819
Delivery	0.33	0.08	0.92	0.37	-7.552	0*
Manner	0.24	0.43	0.37	0.48	-0.924	0.3574
Others	0.20	0.17	0.15	0.17	1.096	0.2794
Total	2.24	0.24	3.61	0.97	-6.297	0*
Cohort 2 (N=14)		Item-based				
	Students (to each performer)	SD	Teacher (to each performer)	SD	t	p (2-tailed)
Content	0.89	0.31	1.36	0.296	-4.082	0.0004*
Language	0.46	0.41	0.79	0.498	-1.92	0.0664
Delivery	0.8	0.4	1.57	0.316	-5.652	0*
Manner	0.36	0.48	0.5	0.5	-0.756	0.4562
Others	0.11	0.31	0	0	1.315	0.2112
Total	2.61	0.39	4.21	0.458	-9.95	0*

*p<.05

Table 4 presents frequencies and percentages of student comments in the category-based and item-based assessments. Chi square tests detected differences in the assessment criteria between the yielded ratio and the approved ratio (content 30%, language 30%, delivery 30%, and manner 10%) in both category-based ($\chi^2= 261.31$, $p<.0001$) and item-based assessments ($\chi^2= 54.3$, $p<.0001$). It indicated that students' understanding of the criteria might not be related to the specificity of the assessment forms. But using the category-based assessment form, student assessors tended to give more comments (2223 vs. 1508) and a handful were not related to the approved criteria (9.2% vs. 6.8%). As to the nature of their comments, students were inclined to look at strengths more than weaknesses in both category-based and item-based assessments. Neutral feedback refers to reflections and suggestions for improvement. In item-based assessment, the percentage of neutral comments was higher than that in category-based assessment.

Table 4 *Frequencies and percentages of student feedback comments in category-based and item-based assessments of two cohorts*

Topic	Category-based	χ^2	p	Item-based	χ^2	p
Content	654 (29.4%)	261.31	<.0001	503 (33.4)	54.3	<.0001
Language	313 (14.1%)			416 (27.6%)		
Delivery	708 (31.8%)			310 (20.6%)		
Manner	343 (15.4%)			176 (11.7%)		
Others	205 (9.2%)			103 (6.8%)		
Total	2223 (100%)			1508 (100%)		
Nature	Category-based	χ^2	p	Item-based	χ^2	p
Positive	1321 (59.4%)	680.98	<.0001	784 (52.0%)	255.13	<.0001
Negative	453 (20.4%)			293 (19.4%)		
Neutral	449 (20.2%)			431 (28.6%)		
Total	2223 (100%)			1508 (100%)		

A close look at students' feedback found that their comments given in the two assessment forms were of similar quality, including positive, negative, and neutral statements. Most of the comments addressed the approved criteria (content, language, delivery, and manner), and an emerging type of comments highlighted preparation and effort of the oral performance. Patterns or relationship between the instrument specificity and students' comment type and nature had yet been detected. The followings are some excerpts of students' feedback comments given on the two assessment forms.

Category-based

- Content is very fascinating and interesting. (Content, Positive, C1S03Category)
- She can shorten the introduction of music, and add more details about why she likes it, or something like that. (Content, Neutral, C1S11Item)
- There are some errors in her grammar. (Language, Negative, C1S05Category)
- Used a lot of vocabulary!! (Language, Positive, C2S06Item)
- She used visual effect well. (Delivery, Positive, C2S01Category)
- Maybe she can show us the tempo of music with sticks. (Delivery, Neutral, C1S15Item)
- She is a little nervous. She needs eye contact with the audience. (Manner, Negative, C2S13Category)
- Eva can speak louder to show her confidence. (Manner, Neutral, C1S17Category)
- Anyway, compare with her last performance, she improves. (Other, Positive, C2S05Category)
- Next time I could try the same thing. (Other, Neutral, C2S09Category)

Item-based

- Her conclusion is very complete and clear. (Content, Positive, C2S08Item)
- I love her description about the peaceful night and her deep friendship with her friends. (Content, Positive, C1S21Item)
- The contents were organized, fascinating for me because I'm quite interested with football but don't know the rules of it. (Content, Positive, C1S21Category)
- The content is somewhat not very rich. (Content, Negative, C1S18Item)
- Fluency and intonation should be improved. (Language, Negative, C2S13Item)
- She could have more interaction with the audience. (Delivery, Neutral, C2S01Item)
- She lacks of eye contact (Manner, Negative, C1S01Item)
- She did better than last time. (Other, Positive, C1S020Item)
- I've practiced many times but still not good enough. I'll try harder next time. (Other, C2S02Item)

Students' Preference of the Assessment Instrument

A great majority of students preferred using the item-based instrument; only a few favored the category-based. Among 35 participating students, about 83% (N= 29) preferred the item-based assessment instrument. They chose it form mainly because of their concern about grading objectivity. They considered the form is more specific and helps them assess a performance more objectively.

- It helps our evaluation become more specific and hence to give fewer assessment. (C1S02)
- It is clearer to present my opinions about the presenter and the paper is more beautiful. (C1S08)

- It provides me more objective criteria to judge. Besides, I can compare each performer by the sheet and adjust their overall scores, so everyone is evaluated in the same standard, and it'll be fairer. (C1S11)
- We could base on more specific detail to evaluate the performance. Through detailed descriptions of how to do the evaluation, it is easy to follow the performance. (C2S03)
- It evaluates all aspects of a performance. It's easier to find the speaker's advantages and disadvantages. (C1S06)

Some students also stated the disadvantages of the item-based assessment form. One said, "There are too many blanks to score and the final score is lower." (C2S13) Despite the grading "effort" required more in the item-based assessment, they thought using the form can help improve themselves. The other student said, "It is more detailed. Although it is much crueller, we can also improve ourselves." (C1S05)

Students favored the category-based assessment because they found limitations of the item-based assessment form. They preferred looking at fewer aspects when rating a performance, enjoyed observing without too much pressure from grading, and liked more space in the form for writing comments. Much fewer students chose the category-based form and no students mentioned the inconveniences of the form.

- It is more general. I won't have to pay too much attention to the grading part and miss some of the speech. (C2S01)
- I am not good at grading. The form is pretty suitable to me. I give fewer scores to one presentation. (C1S07)
- With fewer categories, it is easier to give scores. (C2S05)
- The comment blank is bigger so that I can write more about the presentation. (C1S01)

Students were also asked to express their opinions about peer and self-assessment in a questionnaire. Using either assessment instrument, more than 90% of them held positive perceptions of participation in assessment and ensured its learning benefits. They felt that the practice helped them develop English speaking skills, think critically, become aware of their own merits and shortcomings in speaking, and build up confidence in evaluation and commentary, and the assessment task did not have a negative impact on their relationships with peers. They considered that peer and self-assessment was beneficial to their learning, and they especially valued the feedback or comments provided after the oral performances.

CONCLUSION

This study aimed to compare the impact of two assessment instruments on students' ratings and learning. Major findings are summarized below.

Students' ratings were significantly correlated with the teacher's in both category-based and item-based assessments. But when the item-based assessment instrument was used, students' ratings were distributed over a wider range and the mean scores were lower. The interplay of instrument specificity and assessment experience could explain the degree of agreement between students' and the teacher's rating results.

As to the quantity and nature of students' feedback comments, either assessment instrument had its function. The category-based instrument elicited more feedback statements while the item-based instrument helped students generate similar amounts of feedback in language and manner to the teacher's. Both instruments guided student raters to follow the approved criteria, but students' attention to these criteria seemed to be different from the approved ratios.

A great majority of students preferred using the item-based assessment form because of its being specific and boosting assessment objectivity. The limitations of the item-based form were identified as requiring more effort in rating and thus detracting students from observing and learning from peers' performances. Overall, most of the students stated that they enjoyed participating in the assessment task and considered the task beneficial to their learning.

Based on the positive results of using the category-based and item-based assessment

form in the university EFL classroom, this study also has a couple of suggestions for classroom practice and future research.

In the EFL classroom, a more specific assessment instrument may be used first for little experienced student raters. This study demonstrated that both category-based and item-based instruments, if designed appropriately, are valid and reliable tools for student assessment. But for inexperienced student raters, an item-based one seems to be more sensitive and helps students learn to assess their English oral performance more precisely. Classroom teachers may use the item-based assessment form to provide specific guidance for students in the initial assessment stage. After students become more experienced in assessing peers' and own performances, teachers may adopt a tool of global components so as to elicit more of students' feedback and generate more of their own observation and appreciation of oral performances on preparation, effort, or specialty.

The interplay of instrument specificity and assessment experience was speculated as the reason why there was significant difference between the category-based and item-based assessments in Cohort 1, not in Cohort 2. As previous studies of peer and self-assessment suggested (Chen, 2007, 2008), students' learning development in the assessment task resulted from experience. Students constructed and reconstructed their knowledge of good oral performances through constant practices of observation and assessment. It merits further examination of which factor (assessment experience or instrument specificity) affects the reliability of students' rating more. Future research can replicate the present study with a group of students comparable to Cohort 1.

REFERENCES

- Airasian, P. W. (1997) *Classroom assessment*. 3rd Edition. New York: McGraw-Hill Companies, Inc.
- AlFallay, I. (2004). The role of some selected psychological and personality traits of the rater in the accuracy of self- and peer-assessment. *System*, 32(3), 407-425.
- Chen, Y.-M. (2005). Peer and self assessment in the EFL university classroom. *Selected Papers from 14th International Symposium on English Teaching* (pp. 320-329). Taipei: Crane Publishing Co., LTD.
- Chen, Y.-M. (2007). A study of self-assessment in the university oral training classroom (Da xe kou yu xun lian ke shi zi wo ping liang de yan jiu), *Curriculum and Instruction*, 10(4), 85-100.
- Chen, Y.-M. (2008). Learning to self-assess oral performance in English: A longitudinal case study. *Language Teaching Research*, 12(2), 235-262.
- Cheng, W., & Warren, M. (1999). Peer and teacher assessment of the oral and written tasks of a group project. *Assessment & Evaluation in Higher Education*, 24(3), 301-314.
- Cheng, W., & Warren, M. (2005). Peer assessment of language proficiency. *Language Testing*, 22(1), 93-121.
- Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T., & McNamara, T. (Eds.). (1999). *Dictionary of language testing*. Cambridge: Cambridge University Press.
- Falchikov, N. (1986). Product comparisons and process benefits of collaborative peer group and self assessment. *Assessment & Evaluation in Higher Education*, 11, 146-166.
- Lejk, M. & Wyvill, M. (2001). Peer assessment of contributions to a group project: Student attitudes to holistic and category-based approaches. *Assessment & Evaluation in Higher Education*, 27(6), 569-577.
- Luoma, S. (2004). *Assessing speaking*. Cambridge: Cambridge University Press.
- Miller, P. J. (2003). The effect of scoring criteria specificity on peer and self-assessment. *Assessment and Evaluation in Higher Education*, 28, 383-394.

- Nakamura, U. (2002). A comparison of holistic and analytic scoring methods in the assessment of writing. *Proceedings of the 3rd annual JALT Pan-SIG conference* (pp. 45-52). Retrieved Dec. 22, 2007, from <http://jalt.org/pansig/2004/HTML/Nakamura.htm>
- Orsmond, P, Merry, S., & Reiling, K. (2000). The use of student derived marking criteria in peer and self-assessment. *Assessment and Evaluation in Higher Education*, 25(1), 23-38.
- Patri, M. (2002). The influence of peer feedback on self and peer-assessment of oral skills. *Language Testing*, 19(2), 109-131.
- Schechtman, Z. & Godfried, L. (1993). Assessing the performance and personal traits of teacher education students by a group assessment procedure: A study of concurrent and construct validity, *Journal of Teacher Education*, 44(2), 130-138.
- Schechtman, Z. (1992). Interrater reliability of a single group assessment procedure administered in several educational settings. *Journal of Personnel Evaluation in Education*, 6, 31-39.
- Stiggins, R. J. (2001). *Student-involved classroom assessment* (3rd ed.). New Jersey: Merrill.
- van Duzer, E., & McMartin, F. (2000). Methods to improve the validity and sensitivity of a self/peer assessment instrument. *IEEE Transactions on Education*, 43, 153-158.
- Weigle, S. C. (2002). *Assessing writing*. Cambridge: Cambridge University Press.

Appendix A

Category-based Assessment Form

Assessor: _____

Group: _____

Date: _____

Scoring Criterion/Level				Excellent 90% & above	Good 80%-89%	Fair 70%-79%	Ok 60%-69%	Poor below 60%
Content (30%) Language (30%) Delivery (30%)				27-30	24-26	21-23	18-20	17-
Manner (10%)				9-10	8	7	6	5-
Content	wonderful, surprising, interesting, fascinating, flat, boring, not organized, not to the point			Delivery	fluent, good-guidance, with expressive voice, with various gestures, with visual aids, with eye contact with the audience, average, not lively at all, dull			
Language	precise, error-free, clear pronunciation, appropriate, good use of transition words, poor grammar, unacceptable use of words			Manner	calm, polite, graceful, dramatic, appropriate, no indication of nervousness or irrelevant little movements			
Performer	Criteria	%	Score	Comments				
	Content	30%						
	Language	30%						
	Delivery	30%						
	Manner	10%						
	Total	100%						
Performer	Criteria	%	Score	Comments				
	Content	30%						
	Language	30%						
	Delivery	30%						
	Manner	10%						
	Total	100%						
Performer	Criteria	%	Score	Comments				
	Content	30%						
	Language	30%						
	Delivery	30%						
	Manner	10%						
	Total	100%						
Performer	Criteria	%	Score	Comments				
	Content	30%						
	Language	30%						
	Delivery	30%						
	Manner	10%						
	Total	100%						
Performer	Criteria	%	Score	Comments				
	Content	30%						
	Language	30%						
	Delivery	30%						
	Manner	10%						
	Total	100%						

Appendix B

Item-based Assessment Form

Assessor: _____ Group: _____ Date: _____

Please rate the performance on the following scale: Excellent—5, Good—4, Fair—3, OK—2, Poor--1

Item:	Performer 1	Performer 2	Performer 3	Performer 4	Performer 5
C1. Introduction—topic sentence appropriate?					
C2. Introduction—Interesting, attractive, surprising?					
C3. Details supporting the topic —sufficient?					
C4. Details supporting the topic —relevant, interesting, informative, fascinating?					
C5. Ideas well-organized?					
C6. Conclusion—appropriate?					
L7. Grammar—accurate?					
L8. Fluency (流利)					
L9. Pronunciation—clear?					
L10. Intonation (語調)—appropriate?					
L11. Use of words correct?					
L12. Use of words appropriate?					
D13. Good-guidance					
D14. Facial expressions, gestures—appropriate?					
D15. Interaction with the audience?					
D16. Voice-- expressive?					
D17. Use of visual aids (視覺媒體)—effective?					
D18. Creativity					
M19. Confident, calm, polite?					
M20. Eye contact					
Total					
Comments					

Chinese EFL Learners' Context Spelling Knowledge: Position Constraints and Sequential Dependencies

Chun-en Cheng (鄭淳恩), Chieh-fang Hu (胡潔芳)

Taipei Municipal University of Education

chengchunen@ntu.edu.tw

cfhu@tmue.edu.tw

The study examined the development of two types of *context-sensitive* spelling knowledge, position constraints and sequential dependencies, in Chinese EFL learners (fourth graders, sixth graders, and college English majors). Knowledge of position constraints was assessed by the doublet production/choice tasks, which assessed whether the participants preferred words spelt with final doublets at the end of a word rather than at the beginning (e.g., *baff* vs. *bbaf*). Knowledge of sequential dependencies using vowel information to spell consonant was assessed by the onset production/choice tasks as well as the coda production/choice tasks, which examined whether the participants preferred to spell onset /k/ when it preceded /i, ε/ than when it preceded /a, æ, ʌ/ and whether coda was extended (e.g., *-tch*) when it followed a short vowel, but not when it followed a long vowel. The results revealed that the two types of context-sensitive spelling knowledge emerged no later than the fourth grade, the youngest group in the study. Both types of spelling knowledge developed saliently from elementary school ages to college ages. The findings suggested that context-sensitive spelling knowledge, though not directly taught in school, develops implicitly and progressively with accumulation of literacy experience in Chinese EFL learners.

INTRODUCTION

In Taiwan, phonics instruction has been implemented in elementary schools in order to help children establish letter-sound correspondence rules. Phonics chants like *c-c-c*, /k/-/k/-/k/ are used by many teachers when they teach phonics. However, there are a lot of inconsistencies in mapping between letters and sounds. The proportion of inconsistency between sounds and letters is as high as 72% (Ziegler, Stone, & Jacobs, 1997). Spelling in English involves more than just parsing a word into individual sounds and representing the individual sound by a single letter or letters. Children need to learn not only which sound is represented by which letter but also the position or the sequential constraints by which the various letters are combined. For example, the doublet *-ff* occurs at the end of a word but not at the beginning of a word (position constraints). It occurs more often when the preceding vowel is a long one than when the preceding vowel is a short one (sequential dependency). This context-sensitive spelling information is not often provided by the teacher but has to be induced implicitly by the learners as they become familiar with the conventions of English orthography (Thompson, 1998).

A considerable amount of L1 research has indicated that in addition to translating individual sounds into letters, L1 children at early grades have developed a range of context-sensitive spelling knowledge, which is not directly taught in school. For example, knowledge of position constraints in English spelling has been demonstrated in L1 first graders, who chose pseudowords with final doublets (e.g., *baff*) as more word-like than pseudowords with initial doublets (e.g., *bbaf*) (Cassar & Treiman, 1997). Knowledge of sequential dependencies emerges in the second grade in that children begin to take into

consideration the preceding consonant (i.e., the onset) or the following consonant (i.e., the coda) of a vowel in spelling the vowel (Treiman & Kessler, 2006; Treiman, Kessler, Zevin, Bick, & Davis, 2006). For example, children tended to spell /gled/ as *glead* and /glep/ as *glep* (Treiman & Kessler, 2006). This reflects children's knowledge of conventional spelling that in English the sound /ɛ/ tends to be spelt as *ea* before *d* (*head*, *dead*) and as *e* when it is followed by other consonants such as /b/, /k/, and /p/ (e.g., *ebb*, *step*). While children use the consonantal context to spell vowels, they also use vowel information to disambiguate neighboring consonants (Hayes, Treiman, & Kessler, 2006; Kessler & Treiman, 2004). For example, second graders know that *vaff* instead of *vaf* is a better spelling choice for /væf/. They could also decide that *vaif* rather than *vaiff* is the legal answer for /vef/, indicating that they have the knowledge that a monosyllabic word with a short vowel tends to end with extended coda spellings while long vowels tend to be followed by single consonants (Hayes et al., 2006).

How does the context-sensitive spelling knowledge develop in children? Many researchers suggest that this particular knowledge does not derive from school instruction. An instruction like this will be overwhelming to both the teachers and the students. The context-sensitive knowledge is acquired, largely through implicit learning, by induction and abstraction from the orthographic components that are common to the words children have encountered in print (Cassar & Treiman, 1997; Treiman & Bourassa, 2000; Treiman & Kessler, 2006). As children's print vocabulary accumulates, they are able to generalize about some underlying patterns of English orthography which are not explicitly taught in school. The generalization of orthographic patterns gradually forms another kind of spelling knowledge different from simple letter-sound correspondences. It is a spelling knowledge which is sensitive to the phonological/orthographic environment of the target sound to be spelt (Hayes et al., 2006; Treiman et al., 2006).

Do EFL learners in Taiwan develop the context-sensitive spelling knowledge as L1 children do? For most children in Taiwan, the most accessible way to learn the spelling conventions of English is through phonics instruction administered at school. Context-sensitive rules such as sequential dependencies and position constraints are not included in formal instruction at school. In addition, children in Taiwan mostly do not have as extensive and intensive reading experience of English as L1 children. It is estimated that English L1 learners may acquire 3000 words annually between grades 3 through 12 (Nagy, Anderson, & Herman, 1987). The number of words children learn directly from vocabulary instruction is smaller, approximately 200 to 300 words per year (Jenkins & Dixon, 1983). No matter which estimate, the number of words acquired by an L1 learner largely exceeds the number of the words learned by a Chinese EFL learner. In Taiwan, elementary school children are required to learn about 300 English words within the six years of primary education (Ministry of Education, Republic of China, 2003). As the EFL learning experience, both in and out of school, is impoverished in support of the development of context-sensitive spelling knowledge, whether Chinese EFL children have the context-sensitive spelling knowledge remains to be explored.

The present study presents an initial attempt to investigate whether children in Taiwan have acquired context-sensitive spelling knowledge in English. A group of fourth graders with four years of English learning experience at school was examined. With the concern that fourth graders might not have established complete context sensitivity due to the fact that they just began their learning process of a foreign language, a group of sixth graders was also included in the study to see if the knowledge of context sensitive rules was evident in more experienced English learners. A third group of participants included college English majors, who were at the more terminal point of English learning.

Two types of context-sensitive spelling knowledge were investigated in the present

study: knowledge of position constraints of doublets and knowledge of sequential dependencies between vowels and consonants. Knowledge of position constraints of doublets is found in L1 children as early as in the first half of first grade, and knowledge of sequential dependencies is evident in second-grade L1 children (Cassar & Treiman, 1997; Hayes et al., 2006; Treiman & Bourassa, 2000). As these two types of spelling knowledge emerge quite early in L1 children (Cassar & Treiman, 1997; Hayes et al., 2006; Treiman & Bourassa, 2000), they are good candidates of spelling knowledge to be explored in EFL children. Three questions are addressed in the study: 1) Do Chinese EFL learners have knowledge of position constraints of double letters in spelling? 2) Do Chinese EFL learners have knowledge of sequential dependencies in using vowel to spell consonants? 3) Is there a developmental trend in Chinese EFL learners' knowledge of context-sensitive rules of English spelling?

METHOD

Participants

Three groups of students participated in the present study, 42 fourth graders (mean age 9 years 9 months), 43 sixth graders (mean age 11 years 8 months), and 25 freshman English majors (mean age 20 years 10 months), recruited from an elementary school and a university in Taipei. The elementary school students started to learn English formally in the first grade. Thus the fourth graders had received English instruction from school for four years while the sixth graders had studied English at school for six years. All participants were native mandarin speakers who lived in Taipei city, and none of them had been identified with language disorders.

Materials and Procedure

There were two sets of measures in the current study, two background measures and six experimental measures. The background measures consisted of a letter-sound knowledge test and a sound discrimination test. The letter-sound knowledge was not given to college students, as they were expected to have mastered the basic letter-sound correspondences in English. The six experimental measures, doublet production, doublet choice, onset production, onset choice, coda production, and coda choice, assessed the participants' context-sensitive spelling knowledge. The measures were administered in three sessions. In the first session, the sound discrimination test and the letter-sound knowledge test were given. In the second session, the three production tasks of spelling knowledge were given. The three choice tasks of spelling knowledge were given in the last session.

Experimental Measures

The doublet production and choice tasks examined Chinese EFL learners' knowledge of position constraints. The onset production and choice tasks as well as the coda production and choice tasks examined the participants' knowledge of sequential dependencies in using vowel context to spell consonants (i.e., onsets and codas). Sample stimuli of the six experimental measures are presented in Appendix.

Doublet choice. The stimuli were taken from Cassar and Treiman's (1997) study. There were 20 pairs of pseudowords. The pseudowords in each pair were pronounced identically but spelt differently, one spelt with a final doublet (e.g., *luss*) and the other with an initial doublet (e.g., *llus*). All the doublets in the 20 pairs of pseudowords were common doublets in English (*ll*, *ss*, *pp*, *tt*, *nn*, *ff*, *mm*). The test giver read the pronunciation in each pair twice. The participants chose the more acceptable spelt form on their answer sheets.

Doublet production. The stimuli were 20 pseudowords derived from Cassar and Treiman's (1997) study and had the same pronunciations as those used in the doublet choice task. The test giver read each pseudoword twice and the participants were asked to spell the pseudowords they heard on the answer sheets.

Onset choice. The stimuli were developed based on the 10 pronounceable pseudowords

in Hayes et al. (2006). The 10 pronunciations all had /k/ as their onsets, preceding one of the five vowels /ε, ɪ, α, æ, ʌ/. Based on these 10 pronunciations, 10 pairs of spelling choices were developed. The two choices in each pair differed only in the first letter, one with *k* and the other with *c* (e.g., *kift-cift*, *kosk-cosk*). One spelling choice in each pair was more acceptable than the other, according to the tendency in English that the onset /k/ is normally spelt as *k* before /ɪ/ and /ε/ and as *c* before /α/, /æ/, and /ʌ/. The test giver read each of the pronunciations twice. The participants chose the more acceptable spelt form of the pronunciation on their answer sheets.

Onset production. The 10 pronunciations used in the onset choice task were used. The test giver read each pronunciation twice and instructed the participants to spell out the word they heard on their answer sheets.

Coda choice. The stimuli were the same as the ones used by Hayes et al. (2006). There were 40 pairs of monosyllabic pseudowords, pronounced identically but spelt differently. For each pair, one pseudoword was spelt without coda extension and the other with coda extension (e.g., *mek-meck*, *yul-yull*, *geech-geetch*, *soof-sooff*). Half of the 40 pairs contained a short vowel (/e/, /u/, /o/, /i/) and the other half contained a long vowel (/æ/, /ε/, /ʌ/), which allowed for an investigation of whether the participants selected different codas according to the preceding vowels. The test giver read the pronunciation twice for each pair. The participants chose the more acceptable spelt form on their answer sheets.

Coda production. The 40 pronounces used in the coda choice task were used. The test giver read each pronunciation twice and instructed the participants to spell out the word they heard on their answer sheets.

Scoring Procedures

The participants' performances in the experimental measures were scored following similar procedures. First, the number of responses with a particular spelling was tabulated. Then the proportions of the spelling in each context (e.g., word initially or word finally; after a short vowel or after a long vowel) were calculated respectively. If the participants were sensitive to the context that constrains a particular way of spelling, then they should prefer different types of spellings in different contexts and the proportions of the particular spelling should be different in different contexts. For *doublet production*, the percentages of initial and final doublets that appeared in each participant's spelling performances were calculated separately, yielding two scores for each participant: percentage of initial doublet spelling and percentage of final doublet spelling. For *doublet choice*, similar two scores were yielded for each participant: percentage of choices with initial doublets and percentage of choices with final doublets. For *onset production*, the items in which the onset /k/ was spelt with *k* were tallied for each participant. Then the percentage of *k* spelling with vowels /ɪ, ε/ and the percentage of *k* spelling with vowels /α, æ, ʌ/ were calculated for each participant. The same procedure was used for *onset choice*. For *coda production*, the percentage of coda extensions for short vowels and the percentage of coda extensions for long vowels were calculated. The same procedure was used for *coda choice*.

Background Measures

Letter-sound knowledge. The beginning and ending consonants appearing in the experimental measures of spelling knowledge were sampled for the letter-sound knowledge test to ensure that the participants had the basic knowledge about letter-sound correspondences in English. Vowels were not included in this test because the letter-sound correspondence rules of vowels were not crucial for completing the tasks in the experimental measures. There were two sections in the letter-sound correspondence test, the first section tested 18 beginning consonants (*k, v, s, w, m, g, sh, th, n, c, z, j, y, fr, l, p, t, f*) and the second section tested 12 ending consonants (*p, s, t, f, n, l, m, ck, ch, ll, tch, ff*). In order to make the consonants audible and salient, each consonant was combined the vowel *a*. The test giver

read the stimulus sound (e.g., *sa* or *ap*) while the participants circled the corresponding letter form from the three choices on their answer sheets (e.g., *ca*, *va*, *sa* or *ap*, *af*, *an*).

Sound discrimination. The task assessed the ability to discriminate long and short vowels. Five pairs of vowel sounds which were confusing to Chinese EFL learners were sampled in this test, including /u/-/ʊ/, /o/-/ɔ/, /e/-/ɛ/, /i/-/ɪ/, and /e/-/æ/. The vowels were used to construct 10 minimally contrastive pairs of pseudowords (e.g., /væf/-/vef/) and 6 identical pairs of pseudowords (e.g., /vef/-/vef/). The test giver read each pair of stimuli twice and the participants decided whether the two pseudowords they heard were the same or not.

RESULTS

The participants' performances in each measure were transformed into proportional scores for comparability. The means and standard deviations of the proportional scores of the measures were presented in Table 1. A close examination of Table 1 revealed several interesting findings. First, the scores of final double letters were lower in the doublet production task than in the doublet choice task, revealing that the participants had certain knowledge of positional constraints which enabled them to choose spelling forms with doublet letters but they did not actively use the knowledge when they needed to spell out the same words. Similarly, the participants did not actively apply their knowledge of sequential dependencies to spell codas though they demonstrated this knowledge in the choice task. Second, the scores of the letter-sound knowledge test almost reached the maximum score (over 90 % of accuracy), which showed that the fourth graders and the sixth graders had basic knowledge of letter-sound correspondences. In addition, the accuracy of sound discrimination was over 90% for the three groups of participants, indicating that differentiating long vowels from short ones in monosyllable words did not pose difficulties to most of the participants in the present study.

Table 1. Means and Standard Deviations of the Proportional Scores for the Background and Experimental Measures

Measures	Fourth-grade		Sixth-grade		College	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
<i>Background measures</i>						
Letter-sound-(beginning)	91.79	12.03	95.73	10.62	—	—
Letter-sound-(ending)	91.67	10.08	94.76	7.93	—	—
Sound discrimination	94.64	10.91	92.70	23.27	99.25	2.07
<i>Experimental measures</i>						
Doublet production						
final	6.88	8.54	9.56	12.76	11.14	10.32
initial	0.19	1.28	0.00	0.00	0.00	0.00
Doublet choice						
final	60.47	20.14	72.20	19.55	95.80	9.31
initial	39.52	20.14	27.79	19.55	4.20	9.31
Onset production						
onset <i>k</i> + / <i>ɪ</i> , <i>ɛ</i> /	57.42	31.78	57.44	37.10	70.40	18.36
onset <i>k</i> + / <i>ɑ</i> , <i>æ</i> , <i>ʌ</i> /	29.84	30.34	25.96	34.32	21.60	21.54
Onset choice						
onset <i>k</i> + / <i>ɪ</i> , <i>ɛ</i> /	69.52	24.68	67.90	27.30	96.00	8.16
onset <i>k</i> + / <i>ɑ</i> , <i>æ</i> , <i>ʌ</i> /	36.19	25.46	31.62	22.35	28.00	22.36
Coda production						
short V + coda extension	8.03	14.70	18.83	20.74	41.88	20.46
long V + coda extension	7.71	13.45	9.80	13.05	21.47	18.29
Coda choice						
short V + coda extension	44.64	17.43	45.23	16.68	74.40	21.66
long V + coda extension	36.42	14.02	38.02	16.15	21.00	16.77

Context-Sensitive Spelling Knowledge

To investigate whether the participants had the various context-sensitive spelling knowledge, six 3(Grade) x 2(Context) ANOVAs were conducted separately on the spelling scores in doublet production, doublet choice, onset production, onset choice, coda production, and coda choice. The between-subjects variable was grade (fourth, sixth, college). The within-subjects variable was the context of spelling in the various tasks: doublet position (initial and final) for doublet production and doublet choice; vowel set (/ *ɪ*, *ɛ* / vs. / *ɑ*, *æ*, *ʌ* /) for onset production and onset choice, and vowel type (long vs. short) for coda production and coda choice. Whenever there was an interaction effect, follow-up paired *t*-tests, corrected for Type I error using a Bonferroni correction procedure ($p < .016$), were conducted to compare the spelling scores in the different contexts for each group. If one group of the participants, for example, had knowledge of position constraints in doublets, they should perform differently in different spelling contexts. The results of the ANOVAs and the follow-up comparisons are summarized in Table 2. The effects between Grade and Context in the six spelling measures were displayed in Figure 1 for doublet production and choice, Figure 2 for onset production and choice, and Figure 3 for onset production and choice.

Table 2. Results of the Main Effects and Follow-up Analyses for the Various Context-Sensitive Spelling Knowledge

Tasks	Grade <i>F</i> (2,107)	Overall		Follow-up comparisons		
		Context <i>F</i> (1, 107)	Interaction <i>F</i> (2, 107)	4 th <i>t</i> (41)	6 th <i>t</i> (42)	college <i>t</i> (24)
Doublet production	1.22***	73.31***	1.49***			
Doublet choice	30.16***	217.61***	30.16***	3.37 ⁺	7.45 ⁺	24.58 ⁺
Onset production	.20***	154.10***	4.30***	5.31 ⁺	7.07 ⁺	11.69 ⁺
Onset choice	4.52***	192.76***	9.50***	5.62 ⁺	7.08 ⁺	14.72 ⁺
Coda production	20.87***	40.99***	12.82***	.23 ⁺	3.39 ⁺	4.94 ⁺
Coda choice	2.23***	157.79***	58.52***	3.42 ⁺	2.84 ⁺	10.56 ⁺

* $p < .05$ ** $p < .01$ *** $p < .001$ ⁺ $p < .016$.

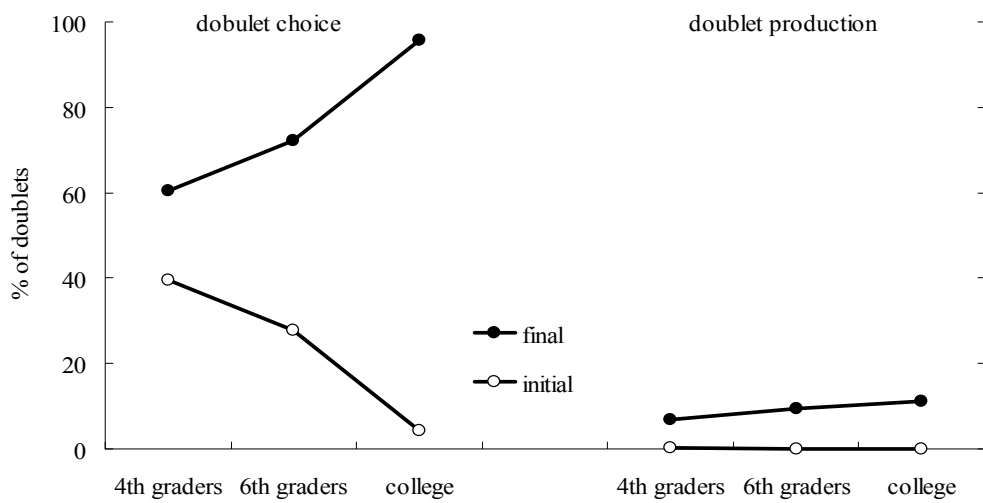


Figure 1. The various effects in doublet choice and doublet production

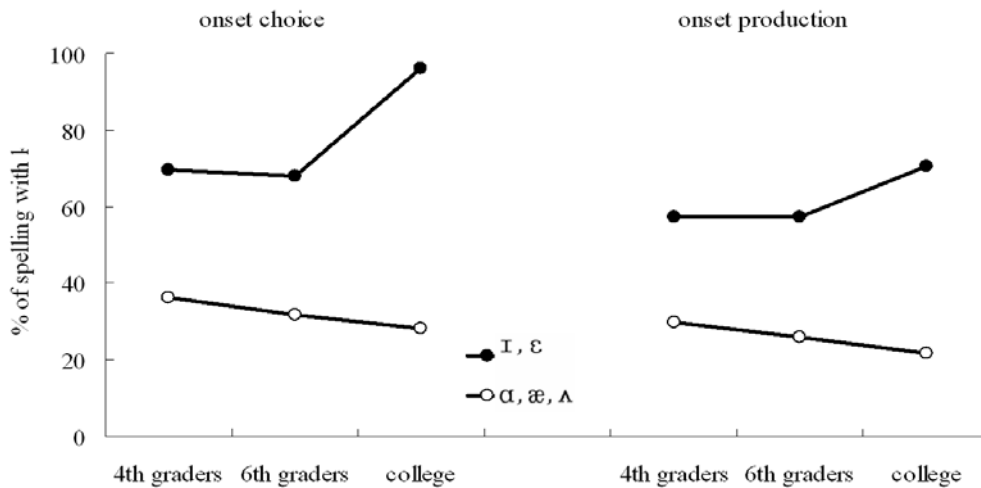


Figure 2. The various effects in onset choice and onset production

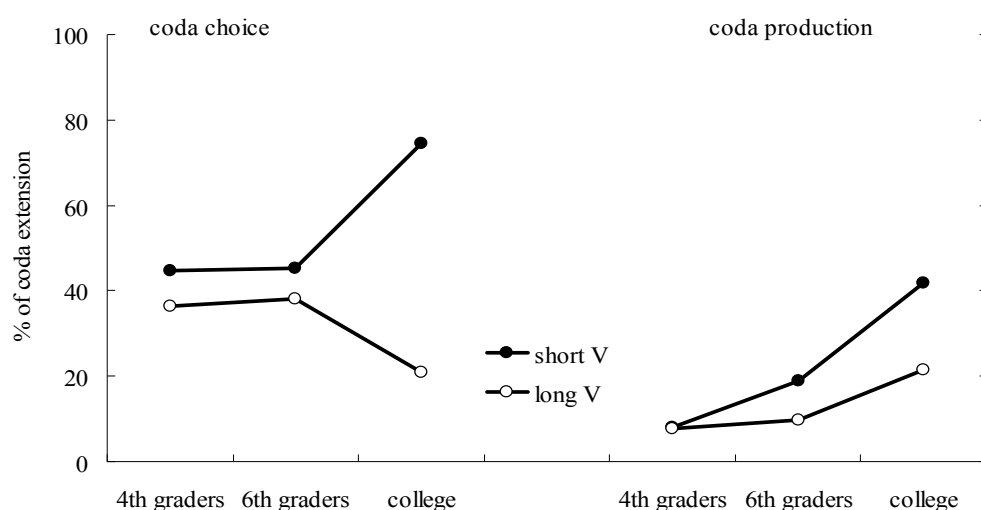


Figure 3. The various effects in coda choice and coda production

As shown by Table 2 and the Figures, the effects of Context in the various spelling measures were significant, indicating the participants in the present study, in general, had developed certain knowledge that doublets were more likely to occur at the end of a word, that onset /k/ was more likely to be spelt with *k* when it preceded vowels / ϵ , ɪ /, and that English coda was more likely to be spelt with multiple letters when it followed short vowels. The results of the follow-up analyses indicated that the context-sensitive spelling knowledge was generally evident in the youngest group of the participants, the fourth graders. The only exception was that the fourth graders did not actively produce coda extension in the coda production task.

The Developmental Trend

To examine whether there was an incremental trend in the development of the various spelling knowledge, mean differences in spelling performances between one context (i.e., initial doublets; *k* before / ϵ , ɪ /; consonant extension after short vowels) and the other corresponding context (i.e., final doublets; *k* before / æ , ʌ , ɑ /; consonant extension after long vowels) were computed for each group. Mean differences in spelling performances between one context and the other were taken as indexes of the strength of the context-sensitive spelling knowledge. Six one-way ANOVAs (using the difference scores as the dependent variable and Grade as the independent variable) were conducted separately for doublets, onset *k* spelling, and coda extension in the choice and the production tasks. When the Grade effect was significant, follow-up *t*-tests were conducted comparing mean differences between adjacent age groups, corrected for Type I error using a Bonferroni correction procedure with a probability level of 0.025. The results of the ANOVAs and the follow-up comparisons are summarized in Table 3.

Table 3. Results of the Follow-up Analyses for the Developmental Trend in Knowledge of Position Constraints and Sequential Dependencies

Effect	Overall	Follow-up comparisons	
	<i>F</i> (2, 107)	4 th and 6 th <i>t</i> (83)	6 th and college <i>t</i> (66)
Doublet production	1.49	—	—
Doublet choice	30.10*	.63	10.96 ⁺
Onset production	4.32*	.72	5.65 ⁺
Onset choice	12.82*	.74	2.82
Coda production	9.50*	7.51 ⁺	2.70
Coda choice	58.52*	.15	3.71

* $p < .05$. ⁺ $p < .025$.

DISCUSSION

The present study investigated two types of context-sensitive spelling knowledge, position constraints and sequential dependencies, in three groups of Chinese EFL learners. The results revealed that EFL learners as young as fourth grade took advantage of contextual clues in spelling. The only exception was that the fourth graders did not use vowel information in active spelling. They did, however, show some knowledge that codas were spelt differently after long or short vowels in coda choice. In addition, there was evidence that the context-sensitive spelling knowledge became consolidated as the learners became more familiar with the conventions of the English orthography, though the increment in the knowledge appeared to vary with the task demands and from one developmental point to another.

Knowledge of Position Constraints

Based on the results of the two doublet tasks, the three groups of participants demonstrated some sensitivity to the position constraints of double letters. They generally preferred double letters at the word-final position than at the word-initial position in the choice task as well as in the production task.

The results obtained from the elementary school children were especially interesting. English L1 children have been found to demonstrate knowledge of position constraints as early as first grade (Wright & Ehri, 2007). The earliest age that was studied in the present study was fourth grade. Although there was three-year difference between L1 learners and the EFL learners in the present study, the finding was still quite impressive. Given that EFL teachers did not normally provide information about position constraints in class, the potential source for the development of knowledge of position constraints in the EFL learners seemed to be the small corpus of English words they had encountered in print, just as L1 children do (Cassar & Treiman, 1997; Treiman & Bourassa, 2000; Treiman & Kessler, 2006; Varnhagen, McCallum, & Burstow, 1997). However, average fourth graders in Taiwan learn approximately 165 words by the fourth grade (Ministry of Education, Republic of China, 2003), whereas English L1 children learn 200 to 300 words per year (Jenkins & Dixon, 1983). Based on this estimation, it is reasonable to assume that the fourth graders in the present study had significantly smaller printed vocabulary than L1 first graders, and yet they started to show some sensitivity to position constraints of English spelling. It seems that the young EFL learners of Chinese, who were drilled predominantly with single letter-sound correspondences in school and who generally did not have extensive reading experience of English, nonetheless attended to the regularities that extended beyond single letter-sound correspondences and abstracted away the orthographic constraints embedded in the small set

of words they had encountered in print.

However, it should be noted that this context-sensitive knowledge is yet to be firmly established. While the three groups of participants took advantage of position constraints to aid in the choice among alternative spellings (e.g., *luss* vs. *llus*), they did not seem to *use* it spontaneously and actively in their spelling. They tended to use a single letter when they spelt novel words (e.g., *lop* for /lap/). It is possible that the participants in the current study were conservative in using doublets to spell consonants at the word-final position. The one-to-one phonics rules might be more familiar to them than the contextualized spelling patterns, which might have prompted the EFL learners to use the conceptually more familiar and thus psychologically safer one-to-one mapping rules in a free-spelling task. To further understand whether and how EFL learners use doublet rules in active spelling, future studies may investigate EFL learner's *functional* knowledge of doublets by examining their performance in a word *learning* task, where the spelling pattern of a new word is dictated and has to be learned as it is, thereby discouraging the use of the conservative strategy.

Knowledge of Sequential Dependencies

In addition to position constraints, the EFL learners in the present study also demonstrated knowledge of sequential dependencies in using vowel information to spell onsets and codas. In spelling words with the onset /k/, the participants preferred *k* spellings than *c* spellings when the stimuli contained vowels /ε, ɪ/ than when the stimuli contained vowels /æ, ʌ, ɑ/. In spelling codas, they showed evidence of using vowel information to determine whether a coda could be spelt with extended letters in a forced-choice spelling test. However, when the participants were required to spell codas in a production task, the fourth graders rarely used extended letters to spell the codas.

A closer look at the participants' performances in the different spelling tasks revealed that they performed better in the onset conditions than in the coda conditions. This was especially the case for the fourth graders, who did not use extended letters in spelling codas. One might attribute the participants' poorer performances in spelling coda to the difficulties in distinguishing short vowels from long vowels, which was critical to the choice between single letters and extended letters in spelling codas. However, this possibility was ruled out as the participants basically had no problem distinguishing long vowels from short ones in sound discrimination.

The poorer performance on coda production in the present study might be related to the salience of the onset /k/ in a word than codas to younger children (Hayes et al., 2006). The younger participants might attend to the initial onset /k/ more carefully because its position was at the beginning of words, compared with codas. Moreover, in onset production, the only onset that appeared in the stimuli was /k/ and the possible spelling form for it was either *c* or *k*, which seemed easy for the participants to choose. On the other hand, in coda production, there were four different codas at the end of the stimuli, each with two possible spellings (extended letters and single letters). The participants had to consider the possible spelling forms for each coda. The complexity in the coda production task might have increased the difficulty of spelling the codas.

In coda production, the participants did not frequently use extended letters to spell codas after short vowels. However, a small number of coda extensions did recur in their spellings. In the elementary school children, codas /l/ and /k/ were spelt by coda extensions most frequently (*-ll*: 42% for the fourth graders and 34% for the sixth graders; *-ck*: 36% for the fourth graders and 55% for the sixth graders), followed by the coda /f/ (*-ff*: 19% for the fourth graders and 12% for the sixth graders). These results were generally consistent with the results in Hayes et al. (2006). In addition, both the children in the current study and those in Hayes et al. (2006) rarely used *-tch* to spell the coda /tʃ/. One reason for the rare use of the extended coda *-tch* might be that trigraphs like *-tch* were considerably rare in English and

thus difficult to younger children (Hayes et al., 2006). In contrast, the extended letters that appeared in the college students were more evenly distributed among the four different codas. Among the 280 coda extensions by the college students, 85 were *-ll* (30%), 71 were *-ff* (25%), 68 were *-ck* (24%), and 56 were *-tch* (20%), indicating that the college students had a rather abundant knowledge of word spellings in English that allowed them to use extended letters more frequently than the elementary school children.

The Development of Context-sensitive Spelling Knowledge

The last question of the present study was whether there was an incremental development in Chinese EFL learners' context-sensitive spelling knowledge. To the best knowledge of the authors, no studies have been conducted to investigate Chinese EFL learners' context-sensitive knowledge across ages. Chinese EFL adult learners have been found to be engaged in intraword analysis less constantly than L1 readers and thus are less sensitive to intraword orthographic constraints (Koda, 1999). The poorer sensitivity to the context-sensitive spelling knowledge in Chinese EFL learners is not inconceivable when they are compared with L1 readers given that EFL learners generally do not have as extensive and intensive reading experience as L1 readers. The more interesting question should be whether Chinese EFL learners' context-sensitive spelling knowledge grows even when they do not read as extensively and intensively as L1 readers. If Chinese EFL learners' context-sensitive spelling knowledge does not grow significantly with years of literacy experience in English, it calls for direct and explicit instruction in context-sensitive spelling rules in classroom settings.

The findings of the present study indicated that there was an incremental development in knowledge of position constraints and knowledge of sequential dependencies in Chinese EFL learners, though the development appeared to vary with task demands and proceeded at different paces from one developmental point to another. The fourth graders were the youngest group that had the context-sensitive spelling knowledge among the three groups of the participants (except for coda production), and the context-sensitive knowledge tended to develop saliently from the sixth grade to the college level. The incremental development of the context-sensitive spelling knowledge seemed to be related to the accumulated experience with the printed vocabulary. However, it remains unclear how the EFL learners' experience with English vocabulary triggered the development of the context-sensitive spelling knowledge. The knowledge might be obtained in a case-by-case manner. For example, seeing the word *small* in print may lead the hypothesis that doublet *-ll* occurs only in the word-final position. The knowledge could also be assembled through statistical induction from the experience of reading words with double letters or extended letters in the word-final position, such as *small*, *watch*, *duck* (Treiman & Kessler, 2006). The fine-tune mechanism behind the development of context-sensitive spelling knowledge and reading experience awaits future efforts that employ a longitudinal design in the Chinese EFL learners.

The results of the study provide some preliminary understanding about the development of context-sensitive spelling knowledge in Chinese EFL learners. Chinese EFL learners have already developed some measurable sensitivity to the contextualized spelling patterns at the elementary level. Knowledge of position constraints in double letters and knowledge of sequential dependencies both emerge no later than the fourth grade, while a full-fledged knowledge of sequential dependencies in using vowels to spell consonants develops from the sixth grade. The contextualized spelling knowledge seems to develop as the learners' English printed vocabulary accumulates through reading. As the learners become more familiar with the spelling conventions of English, they gradually learn to make use of the more sophisticated context-dependent orthographic knowledge in spelling words.

One of the immediate implications from the findings of the present study is that direct instruction on the context-sensitive spelling rules of English can be beneficial, given that the

context-sensitive spelling knowledge, albeit evident, is yet firmly established among elementary school children. However, the implication can be misleading. Direct instruction on the complex spelling rules can be overwhelming and thus discouraging to beginning EFL learners. A reasonable and practical implication should be that young EFL learners, while they are in the stage of mastering the simple letter-sound correspondences, can benefit from extensive reading, which not only allows them to practice and consolidate simple letter-sound correspondences in meaningful texts but also gives them opportunities to generalize about the more complex context-sensitive rules from printed input. Studies have found that extensive reading provides EFL learners with a wide range of learning benefits (e.g., Day & Bamford, 1998; Iwahori, 2008). For example, it helps improve learners' reading rate and language proficiency in the EFL context. One additional benefit of extensive reading can be the facilitation of the development of EFL learners' sensitivity to context-sensitive spelling rules without direct instruction. Nevertheless, the connection between extensive reading and the development of context-sensitive spelling knowledge needs to be confirmed in future studies.

ACKNOWLEDGEMENTS

This study was supported, in part, by a grant from the National Science Council (NSC94-2411-H-133-004; NSC95-2411-H-133-001-MY3).

REFERENCES

- Cassar, M., & Treiman, R. (1997). The beginnings of orthographic knowledge: children's knowledge of double letters in words. *Journal of Educational Psychology*, 89, 631-644.
- Day, P. R., Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge, England: Cambridge University Press.
- Iwahori, Y. (2008). Developing reading fluency: A study of extensive reading in EFL. *Reading in a Foreign Language*, 20, 70-91.
- Jenkins, J., & Dixon, R. (1983). Vocabulary learning. *Contemporary Educational Psychology*, 8, 237-260.
- Kessler, B., & Treiman, R. (2004). Sensitivity to statistical contextual patterns when spelling consonants. In Hayes, H., Treiman, R., Kessler, B. (2006). Children use vowels to help them spell consonants. *Journal of Experimental Child Psychology*, 48, 224-245.
- Koda, K. (1999). Development of L2 intraword orthographic sensitivity and decoding skills. *The Modern Language Journal*, 83(1), 51-64.
- Ministry of Education, Republic of China. (2003). *General guidelines of grade 1-9 curriculum of elementary and junior high school education*. Taiwan, Taipei, Ministry of Education. Retrieved October, 1, 2007, from http://teach.eje.edu.tw/edu_resource/all/.
- Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal*, 24, 237-270.
- Thompson, G. B. (1998). The processes of learning to identify words. In G. B. Thompson & T. Nicholson (eds.), *Learning to Read* (pp. 25-54). New York, NY: Teachers College.
- Treiman, R., & Bourassa, D. C. (2000). The development of spelling skill. *Topics in Language Disorders*, 20(3), 1-18.
- Treiman, R., & Kessler, B. (2006). Spelling as statistical learning: using consonantal context to spell vowels. *Journal of Educational Psychology*, 98, 642-652.
- Treiman, R., Kessler, B., Zevin, J. D., Bick, S., & Davis, M. (2006). Influence of consonantal context on the reading of vowels: Evidence from children. *Journal of Experimental Child Psychology*, 93, 1-24.
- Varnhagen, C. K., McCallum, M., & Burstow, M. (1997). Is children's spelling really stage-like? *Reading and Writing: An interdisciplinary journal*, 9, 451-481.
- Wright, D. A. & Ehri, L. C. (2007). Beginners remember orthography when they learn to read

words: The case of doubled letters. *Applied Psycholinguistics*, 28, 115-133.
 Ziegler, J. C., Stone, O. G., & Jacobs, A. M. (1997). What is the pronunciation for –ough and the spelling for /u/? A database for computing feedforward and feedback consistency in English. *Behavior Research Methods, Instruments, and Computers*, 29, 600-618.

APPENDIX

Sample Stimuli for Doublet Production and Doublet Choice

Pronunciation	Spelling choice
/lʌs/	luss-llus
/lɑp/	lopp-llop
/læʔ/	latt-llat
/tɪs/	tiss-ttis
/tʌl/	tull-ttul
/tʌn/	tunn-ttun

Sample Stimuli for Onset Production and Onset Choice

onset	Pronunciation	Spelling choice
/k/	/kɪft/	kift-cift
	/kæft/	kaft-caft
	/kɪmp/	kimp-cimp
	/kʌmp/	kump-cump

Sample Stimuli for Coda Production and Coda Choice

coda	Short vowel stimuli		Long-vowel stimuli	
	Pronunciation	Spelling choice	Pronunciation	Spelling choice
/f/	/væf/	vaf-vaff	/vef/	vaif-vaiff
	/sæf/	saf-saff	/suf/	soof-sooff
/k/	/gæk/	gak-gack	/gok/	goak-goack
	/mek/	mek-meck	/mok/	moak-moack
/l/	/θʌl/	thul-thull	/θul/	thool-thooll
	/jʌl/	yul-yull	/jil/	yeel-yeell
/tʃ/	/gætʃ/	gach-gatch	/gitʃ/	geech-geetch
	/jætʃ/	yach-yatch	/jotʃ/	yoach-yoatch

Students Perspectives toward the Management of a Self-Access Learning Center

Wen-wen Cheng (鄭文文), Mei-ling Lee (李美玲)

National Kaohsiung First University of Science and Technology

wliebe@gmail.com
meiling@ccms.nkfust.edu.tw

For the past two decades, the most typical approach for educational institutions to experiment with the notion of learner autonomy is through the establishment of self-access resource centers (Benson & Voller, 1997). To catch up with the trend, National Kaohsiung First University of Science and Technology set up its Multimedia English Learning Center (MELC) through grants from the Ministry of Education in 2003. The goal of the Center is to create a facilitating environment for students to learn to plan, monitor and evaluate their own learning beyond the classroom setting. In the literature, a good number of experts suggest that the success of a self-access center depends on good management with support of effective learning resources, responsible staff and learner involvement. The purpose of the current study, therefore, is to examine the management of the MELC from the aspects of learning environment, self-access program, and counseling services. To gather the data, two hundred and seventeen non-English majored students from NKFUST were recruited to respond to the questionnaires designed for this study. Major findings indicate that more than half of the students are satisfied with the learning environment and self-access program. Yet, approximately one third of the students (32.7%) are not content with the weekly allotted learning time since they expect longer hours for training in self-directed learning. Furthermore, the low student involvement (1.8%) in the counseling service shows that the majority of students do not have the habit of requesting help from English counselors. Pedagogical implications are addressed.

INTRODUCTION

The successful management of a self-access learning center depends on the support of advantageous software and hardware, responsible staff and learner involvement (Benson, 2001; Gardner and Miller, 1999; Sturtridge, 1997). Over the past decade in Taiwan, though the number of self-access learning centers has mushroomed in the university communities, very few schools have conducted a thorough and systematic evaluation on the operation of such centers from learners' perspectives. Not surprisingly, most self-access program providers are not clear about the attitudes of learners with regard to the overall learning environment and management of these centers. Therefore, for the present study, we are interested in examining whether the functions of the Multimedia English Learning Center (MELC) at NKFUST fulfill the expectations of student users from the following three aspects of management: the overall learning environment, self-access program, and counseling services.

For the learning environment, the air-conditioned MELC equips 60 multimedia stations with access to the Internet, broadcasting system, and two counseling rooms. Learning in this spacious and modern center, students' reflection on the learning environment and its high-tech equipment will be the major issue we would like to tackle. Other issues pertaining to the management of the Center will also be explored. For instance, we would like to know whether the services offered by the staff are effective in supporting day-to-day learning such

as seat assignment, computer hardware and software operation and consultation. In addition, we would also like to find out whether the opening hours are convenient for the students and if students expect the opening hours to be expanded or changed in the future. As to the self-access program, the current self-access courses have been integrated with the general education curriculum and its purpose is to provide students with learning opportunities outside the classroom. Specifically, we would like to discover if students like this new mode of self-learning and whether they actually improve their English speaking and listening abilities. Lastly, whether the counseling service elaborates its function to assist students during the self-access learning process will also be examined. Based on the purpose of the study, the following questions will be investigated.

1. Are students satisfied with the overall learning environment and management of the MELC?
2. What are the attitudes of students toward the self-access program offered by the MELC?
3. Do students make frequent and regular use of the counseling service provided in the MELC?

METHOD

The research questions mainly concern students' attitudes and viewpoints toward the overall management of the MELC. To assemble the opinions of a large number of students, we selected the questionnaire as the major data collection tool. The questionnaire primarily consisted of two types of questions. The first kind included multiple-choice items designed for eliciting respondents' in-depth reflection related to the issues I would like to know. The second type involved questions that used the five-point Likert rating scale to gauge student responses. During the survey, students were instructed to mark their opinions in the scale to indicate the extent to which they agreed with each of the designed statements.

PARTICIPANTS

The student participants for this study were recruited from College of Foreign languages, College of Management, and College of Engineering at NKFUST in the spring of 2005. Under the school policy, students at Levels 1 and 2, the lower levels of English proficiency, are required to conduct extra 18 hours of self-directed learning in the MELC in addition to 2 hours in-class instruction, while students at Levels 3 and 4, the higher level, have the option to do so. To meet the purpose of this study, students of Levels 1 and 2, the required users of the MELC, were recruited as the target subjects for this study. In total, 217 students were randomly selected to participate in this survey.

STATISTICAL PROCEDURES

In order to analyze and describe the trends of students' attitudes toward the management of the MELC, descriptive statistics were used to summarize the data. Specifically, the gathered data were being computed and tabulated with application of SPSS 10.0 (Statistical Package for the Social Science). Moreover, to present the results in a meaningful way, the numbers and percentages of the data were compiled question-by-question. All the calculations were based on the valid data set and were rounded up to the first place after the decimal point. It is worth noting that if the participants were asked to rank their choices from a list in the questionnaire, their top three choices were assigned different scores as follows: the first choice was assigned three points, the second choice two points, and the third choice one point. The counts for the selected options were then added up and ranked from the most to the least points to determine the priority.

RESULTS

The first research question addresses student satisfaction with the overall learning environment and management of the MELC. The results show that 53.9% of the students like the learning environment of the MELC, and only 11.5% do not like it. When asked about whether they are satisfied with the computer equipment offered by the MELC, 61.3% of the students reported that they are satisfied. As to the staff and services provided in the MELC, only 34.1% of the students stated that the staff in the MELC is friendly and 40.6% of them think the staff offers good services to solve their problems. About the availability of the Center, 47.5% of the participants are satisfied with the opening hours of the MELC. However, 73.7% of the students reported that they often have to wait in line when they visit the MELC. When asked their opinions about the learning passport, a booklet designed for recording learning progress, 36.4% of the students reported that they like to use the learning passport, whereas 21.2% do not. Table 1 summarizes the above findings.

Table 1. *Students' Satisfaction with the Management of the MELC*

ITEMS	AGREE		NC	DISAGREE	
	SA	A		D	SD
1. Overall, I like the learning environment of the MELC.	53.9%		34.6%	11.5%	
	12.0%	41.9%		9.7%	1.8%
2. Overall, I am satisfied with the computer equipment offered by the MELC.	61.3%		28.1%	10.6%	
	10.1%	51.2%		8.3%	2.3%
3. Overall, I think the staff in the MELC is friendly.	34.1%		39.6%	26.3%	
	4.6%	29.5%		19.4%	6.9%
4. I think the staff in the MELC offers good services to students.	40.6%		42.9%	16.5%	
	5.1%	35.5%		12.4%	4.1%
5. Overall, I am satisfied with the opening hours of the MELC.	47.5%		26.7%	25.8%	
	7.4%	40.1%		20.3%	5.5%
6. I often need to line up and wait when I visit the MELC.	73.7%		17.5%	8.8%	
	40.1%	33.6%		6.5%	2.3%
7. Overall, I like to use the learning passport designed by the MELC.	36.4%		42.4%	21.2%	
	4.6%	31.8%		18.4%	2.8%

SA = Strongly Agree, A = Agree, NC = No Comment, D = Disagree, SD = Strongly Disagree

In responding to the services that the staff in the MELC offers or fails to offer, the students rank "Making online reservations and stamping the paper learning passport" as the first efficient service, "Assigning seats" as the second and "Operating computer" as the third. As to the services that the staff fails to offer, the ranking of the top three items is described as follows: "Renting audiovisuals equipment like headphones," "Operating multimedia learning software applications" and "Supervising learning in the MELC" respectively. Table 2 outlines the results.

Table 2. *Ranking of Services Provided in the MELC*

Top Three Efficient Services Offered in the MELC		
Rank	Services	Count
1	Making online reservations and stamping the paper learning passport	64
2	Assigning seats	53
3	Operating computer	36
Top Three Services the MELC Fails to offer		
Rank	Services	Count
1	Renting audiovisuals equipment like headphones	24
2	Operating multimedia learning software applications	13
3	Supervising learning in the MELC	12

The second research questions examines student attitude toward the self-access program offered by the MELC. The statistics outlined in Table 3 show that 50.2% of the participants agreed that it is a good idea to integrate the self-access program into the general English curriculum; 49.7% of the participants reported self-access courses will help them improve listening and speaking abilities; 53.4% responded that it's a good idea to include the results of English self-directed learning as part of the final grade for the general English course; 54.4% of them are satisfied with the policy that English self-directed learning constitutes 20% of the final grade for the general English course; 32.7% of the them are satisfied with the requirement of conducting one-hour self-directed learning per week in the MELC. Finally, only 12.4% of the students reported they will spend extra time learning English in the MELC in addition to the required hours for self-directed learning, while as many as 44.7% of the students said they will not.

As to the study habits of the students in the MELC, the result in Table 4 shows that 52% of the participants usually study weekly and regularly; 38% of them study in a sporadic manner; and only 10% of them study intensively at the end of the semester.

Table 3. *Responses to Self-directed Learning Curriculum*

Items	AGREE		NC	DISAGREE	
	SA	A		D	SD
1. I think integrating self-directed learning into the General English Program is a good idea.	50.2%		29.0%	20.8%	
	10.1%	40.1%		14.3%	6.5%
2. The self-access program helps me improve my listening and speaking abilities.	49.7%		31.3%	18.8%	
	7.8%	41.9%		14.7%	4.1%
3. I think that it's a good idea to include the results of English self-directed learning as part of the GE course's final grade.	53.4		23.5%	23%	
	14.7%	38.7%		16.1%	6.9%
4. I am satisfied with the current school policy that English self-directed learning constitutes 20% of the GE course's final grade.	54.4%		24.0%	21.7%	
	14.3%	40.1%		14.3%	7.4%
5. I am satisfied with the current school policy that requires students to conduct English self-directed learning one hour per week.	32.7%		28.1%	39.2%	
	4.6%	28.1%		28.1%	11.1%

6. In addition to the required hours for self-directed learning, I will spend extra time learning English in the MELC.	12.4%		42.9%	44.7%	
	2.3%	10.1%		33.6%	11.1%

Table 4. *Students' Study Habits*

Response	Frequency	Percentage
Intensively at the end of the semester	22	10%
Weekly and regularly	112	52%
In a sporadic manner	83	38%
Total	215	100.0%

When asked how many hours of English self-directed learning they would like to have if they have a choice, 35.8% of the participants choose “one hour” and almost an equal percentage of students 33.5% select “two hours.” Interestingly, only as few as 11.6 % like to learn more than 3 hours, while 7.4% do not want to study English at all in the MELC in addition to two-hour classroom instruction as shown in Table 5.

Table 5. *Responses to the Ideal Learning Hours for English Self-directed Learning*

Response	Frequency	Percentage
0 hour	16	7.4%
1 hour	77	35.8%
2 hours	72	33.5%
3 hours	25	11.6%
more than 3 hours	25	11.6%
Total	215	100.0%

Regarding the English counseling service probed in the third research question, Table 6 shows that only 34.6% of the participants knew about this service in the MELC. As the participating students were not required to visit the English counselors by the university, only 1.8% of them have ever used this service; among them, 75% of them are satisfied with the counseling service.

Table 6. *Responses to the Effectiveness of the Counseling Service*

Items	Yes	No
1. Do you know the MELC offers English counseling service to the students? (N=217)	34.6%	65.4%
2. Have you ever used the English learning service offered by the MELC? (N=217)	1.8%	98.2%
3. I have used the English counseling service offered by the MELC and I was satisfied with the service. (N=4)	75%	25%

CONCLUSION

As discussed earlier, the success of a self-access learning center heavily relies on support of good management, facilitative environment, responsible staff and learner involvement. The findings of the current study suggest that the MELC should continue to improve its management so that it can better support the ongoing development of students' ability to take charge of their own learning outside the classroom. Though the participating students are generally fond of the overall learning environment in the Center, the future improvements should be particularly made in the areas of learning-related services, quality staff, extension of opening hours, user-friendliness of the learning passport, and promotion of the counseling service. To elaborate, the statistical analysis shows that over half of the

students (53.9%) like the environment of MELC and are satisfied with the computer facilities. Yet, when it comes to the working attitude of the staff, a fairly low percentage of students think the staff is friendly enough (34.1%) and can provide effective services to them (40.6%). Based on the results, though students think the staff is helpful in providing routine services such as making online reservations and stamping the paper learning passport, they do not offer services that are essential in supporting self-directed learning. From the perspective of students these services include renting audiovisual equipment, operating multimedia learning software applications, supervising learning and interacting with students in English. Apparently, students' responses imply that they are eager to have an effective and all-English learning environment where the staff is expected to play an active role in serving the students and fostering real-life communication skills.

As to the open hours of the MELC, a major problem concerning the students is that the MELC should consider extending the service hours so that students can have more opportunities to study in the center. Currently, it is open from 1:30 in the afternoon to 9:30 in the evening from Monday to Friday and 10:30 in the morning to 6:30 in the evening on Saturday. As a result, students who do not have classes in the morning cannot study in the center as the service is not available. Second, a high percentage of the students (73.7%) reported that they often had to wait in a queue when they visited the center. The problem of lining up may be a result of the limited service hours and insufficient capacity of the Center which has only 60 multimedia stations.

With regard to the paper learning passport, the dissatisfaction mainly lies in the issue of convenience. First, standing in a queue in order to receive a stamp record is not an interesting thing to do for most university students. Second, some students consider recording the learning content by hand in the passport as toilsome. Based on the findings, the computer experts in NKFUST should consider creating an electronic version of the learning passport that has the function to automatically record the learning content as well as learning time spent in the Center for the students in the future.

Finally, when it comes to the English counseling service, though as many as 34.6% of the students know that the MELC offers such a service to the students, only 1.8% of the students, an astonishingly low percentage, have ever used the service. We conjecture that the low student involvement in the counseling service might be due to two reasons. First, the ineffective promotion of the service to the students explains why students possibly fail to see the positive relationship between the usefulness of English counseling and their development of self-directed learning skills in the long run. To address the problem, teachers and school authorities need to find effective ways to help students recognize the value of seeking professional advice to facilitate language learning. Secondly, students may not have a habit of visiting a counselor for help to solve their learning problems. As some scholars like Kennedy (2002) and Tusi (1996) have observed, Chinese students tend to be passive learners in the traditional classroom and are shy to request help.

To sum up, the overall learning environment of the MELC should be made in ways to support the opportunities for students to direct their own learning. Based on what we found in this study, we suggest that the Center should continue to improve its management with regard to the working attitude of the staff, a realignment of the opening hours, promotion of the counseling service and improvement on the printed learning passport to increase learner involvement. Furthermore, an effective self-access learning center could not depend on technology alone to stimulate and foster learner autonomy. It should have the power to motivate students to engage in and reflect on their own learning. Gibbons (2002) argued that one of the major principles of motivation for self-directed learning is to create a positive, inviting, and non-threatening climate that will help students nurture productivity. We propose that the MELC needs to further gauge the effectiveness of its management on the

basis of student participation and ongoing development of motivation for self-regulation. In so doing, the MELC will not only promote a technology-based learning approach but also create an optimal environment for students to hone their learning skills and strategies in an autonomous manner.

REFERENCES

- Benson, P. (2001). *Teaching and Researching Autonomy in Language Learning*. London: Longman.
- Benson, P. & Voller, P. (1997). *Autonomy & Independence in Language Learning*. London: Longman.
- Gardner, D. and Miller, L. (1999). *Establishing Self-Access : From Theory to Practice*. Cambridge: Cambridge University Press.
- Gibbons, M. (2002). *The Self-Directed Learning Handbook: Challenging Adolescent Students to Excel*. California: John Wiley & Sons.
- Kennedy, Peter. (2002). Learning cultures and learning styles: myth-understandings about 9adult (Hong Kong) Chinese learners, *International Journal of Lifelong Education*, 21 (5) (September- October): 430-445.
- Sturtridge, G. (1997). Teaching and language learning in self-access centers: changing roles?. In P. Benson and P. Voller (eds) *Autonomy and Independence in Language learning*. London: Longman, 66-78.
- Tsui, A. (1996). Reticence and anxiety in second language teaching. In K. Bailey and D. Nunan (eds) *Voices from the language classroom*. Cambridge: Cambridge University Press: 145–167.

Investigating the Summary Writing Performance of University Students in Taiwan

Shao Joyce Chin (金韶)

The Language Training and Testing Center
joycechin@lttc.ntu.edu.tw

This paper reports the initial findings of an ongoing investigation into the summary writing performance of university students in Taiwan. Forty-six students at the high-intermediate and advanced levels of proficiency participated in this study, in which they were asked to read two argumentative English texts and write a summary in English for each source text. Frameworks from Friend (2001), Rivard (2001), and Yu (2007) were adapted for data analysis. The 92 summaries were analyzed for the number and types of text propositions covered (i.e., main ideas, extraneous ideas and inaccurate statements) and then evaluated for the quality of writing (i.e., paraphrase and integration of ideas, rhetorical features, and language control). The performance data were triangulated with data obtained from post-task questionnaires and follow-up interviews.

The results showed that the high-intermediate level students, compared to their advanced-level counterparts, included fewer main ideas and more extraneous ideas in their summaries. The former group also appeared to be more constrained by lexical and grammatical deficiencies than the latter group did while writing summaries, resulting in more frequent use of verbatim replication and greater difficulty in reformulating and integrating ideas. Both groups needed more training on how to rearrange and reorganize the ideas extracted from a source text to compose a coherent summary text with adequate rhetorical fluency. The paper concludes with suggestions for teaching and further research.

INTRODUCTION

In the field of English for Academic Purposes (EAP), summary writing has been regarded as an integrative task that involves the interplay of two abilities—the ability to understand the main ideas and rhetorical organization of a text and the ability to compose a succinct and coherent restatement of the author's gist (Enright *et al.*, 2000). Research on EAP training suggests that summary tasks can facilitate the holistic development of learners' English reading and writing abilities as well as their content area study skills (Radmacher & Latosi-Sawin, 1995; Friend, 2001). Nonetheless, previous studies (Cohen, 1994; Moore, 1997; Kim, 2001; Yang & Shih, 2003) have also shown that this integrative task is very challenging for ESL/EFL learners, as their deficiencies in reading and writing might lead to breakdowns at certain points in the process of summary production, ranging from identifying key words to integrating ideas into a coherent restatement.

Despite the above-mentioned pedagogical potentials and challenges of summary writing for EAP training, only a limited number of empirical studies have investigated the summary writing performance of university students in Taiwan. Some of these studies have explored the role of summarization in facilitating the development of writing ability. Hsu (2003), for instance, investigated the summary writing performance of 12 first-year English majors participating in web-based integrated reading and writing activities. She reported that summary writing helped the students conceive how to present the ideas in an organized manner and thus served as a “springboard” for them “to embark on composing tasks” (p. 18). Several other studies examined the effectiveness of summary writing as an instructional

strategy. Pan (2003), who experimented with two ways of using model essays in writing instruction, found that students who were instructed to summarize model essays regularly performed better in the post-course argumentative writing task than those instructed to read model essays extensively. In contrast to Pan's (2003) product-oriented approach, Chen (2002) analyzed students' think-aloud protocols to compare the strategies they used when completing different writing tasks. The results indicated that when writing a summary, the students employed relatively fewer writing strategies and high-level reading strategies than they did in writing a response essay.

These studies, though limited in number, have provided some glimpses of how university students in Taiwan deal with summary tasks and how this task type might foster reading and writing connection. However, in order to better understand how learners' developing reading and writing abilities affect their summary writing performances, there is an imperative need for empirical research which analyzes learners' written summaries in a more comprehensive manner, covering both the dimensions of content coverage and writing quality, and provides diagnostic information about their strengths and weaknesses in both dimensions.

Some recent empirical studies have provided insight into the development of a framework for analyzing learners' summaries and profiling their performances. Rivard (2001), through the collaborative effort of curriculum and language experts, identified ten "variables" for analyzing summaries. Four variables (*main ideas*, *secondary ideas*, *integration of ideas*, and *fidelity to the text*) focus on content-related issues; five variables (*organization*, *style*, *language usage*, *objectivity*, and *holistic writing score*) focus on language-related issues; and the last variable, *efficiency*, is related to both content and language. More recently, Yu (2007) developed a similar, but less complicated, evaluation framework incorporating two scoring schemes. The first scheme assigns points to each summary according to its adequacy and accuracy of content coverage. The second scheme evaluates the overall quality of the summaries based on a holistic scale that incorporates four key indicators: (1) *faithfulness to the source text*, (2) *summary and source text relationships*, (3) *conciseness and coherence*, (4) *ease of understanding*.

The above two researchers have identified some useful criteria for profiling students' summary writing performance. However, neither has provided an adequate treatment of learners' difficulties in paraphrasing and integrating ideas. Concerning these difficulties, Friend (2001), following the work done by Winograd (1984), developed a scheme for coding each sentence in a summary according to the amount of text from the original subsumed and the quality of paraphrase or integration. The scheme includes five categories: (1) *reproduction* (a single sentence rewritten), (2) *run-on* (two or more sentences poorly integrated), (3) *combination* (two or more sentence well integrated), (4) *low-level invention* (more than one paragraph poorly integrated) and (5) *high-level invention* (more than one paragraph well integrated). This coding scheme, nevertheless, does not include verbatim replication, a problem commonly observed in earlier empirical research on summary writing (Campbell, 1990; Johns & Mayes, 1990; Moore, 1997).

In the present study, a framework for analyzing and profiling written summaries was constructed on the insights provided by the above-mentioned studies. The framework includes three aspects: (1) a set of content-related criteria for analyzing the number and types of text propositions abstracted from the source text (i.e., main ideas, extraneous ideas and inaccurate statements), (2) a set of writing-related criteria (i.e., paraphrase and integration of ideas, rhetorical features, and language control), and (3) a holistic scale for evaluating the overall quality of the summaries, taking both content coverage and writing quality into consideration. The framework was employed to compare the summary writing performances of university students in Taiwan at two proficiency levels: the high-intermediate and

advanced levels. The study was designed to address the following research questions:

- (1) In terms of content coverage (i.e., the number and types of text propositions abstracted from the source text), are there any significant differences in the summaries produced by Taiwanese university students at the high-intermediate and advanced levels of proficiency?
- (2) In terms of the quality of writing (i.e., paraphrase and integration of ideas, rhetorical features, and language control), are there any significant differences in the summaries produced by Taiwanese university students at the high-intermediate and advanced levels of proficiency?
- (3) When both content coverage and writing quality are considered, are there any significant differences in the summaries produced by Taiwanese university students at the high-intermediate and advanced levels of proficiency?

RESEARCH DESIGN

Participants

Forty-six EFL university students in Taiwan at the high-intermediate and advanced levels of proficiency participated in this study. The high-intermediate level students (N=23) were recruited from a pool of candidates who had scored between 80 and 90 on the LTTC-GEPT High-Intermediate Level Reading Test, and the advanced-level students (N=23) were recruited from a pool of candidates who had scored 90 or above on the LTTC-GEPT Advanced Level Reading Test¹⁰. The decision to recruit participants from those two pools was made on the basis of two considerations: (1) these two levels of tests are intended to measure the English abilities of university students in Taiwan, with the High-intermediate Level Test targeting non-English majors and the Advanced Level Test targeting English majors; (2) university students account for the largest proportion (approximately 60%) of the candidature of these two levels of tests.

Data Collection Procedures

The participants had to complete two summary writing tasks. For each task, they were given 60 minutes to read an argumentative text in English (approximately 800 words) and write a summary in English (approximately 180-200 words). A total of 92 summaries were produced by the 46 participants. While writing each summary, they had access to the source text but were not allowed to use dictionaries. Immediately after completing each task, they filled out a post-task questionnaire. In addition, before the first task session began, the participants were asked to fill out learner background questionnaires and then given a 10-minute orientation on summary writing task requirements. An overview of the data collection procedures is provided in Table 1.

To control text effects, counterbalancing procedures, illustrated in the shadowed parts in Table 1, were developed. Within each proficiency group, roughly half of the students were randomly assigned to Sequence 1, and the other half to Sequence 2.

¹⁰ The cut-off scores for passing the LTTC-GEPT High-Intermediate and Advanced Reading Tests are both 80 (out of 120 points).

Table 1. *Data collection procedures*

Sequence 1 (11 HI* participants and 12 AD* participants)		Sequence 2 (12 HI participants and 11 AD participants)	
Procedures	Time (min.)	Procedures	Time (min.)
• Learner background questionnaire	15	• Learner background questionnaire	15
• Orientation on summary writing task requirements		• Orientation on summary writing task requirements	
• Summary Writing Task A -summarizing the source text entitled "The Biopharming Controversy"	60	• Summary Writing Task B -summarizing the source text entitled "The Aquarium Controversy"	60
• Post-task questionnaire on Task A	10	• Post-task questionnaire on Task B	10
<i>Break</i>	10	<i>Break</i>	10
• Summary Writing Task B -summarizing the source text entitled "The Aquarium Controversy"	60	• Summary Writing Task A -summarizing the source text entitled "The Biopharming Controversy"	60
• Post-task questionnaire on Task B	10	• Post-task questionnaire on Task A	10

*The abbreviations "HI" and "AD" are used to refer to "high-intermediate level" and "advanced level", respectively.

In addition to collecting data from the two groups of university students using the above instruments, the researcher also conducted post-evaluation interviews with the raters to elicit their comments about the participants' summaries and the evaluation framework.

Materials and Instruments

1. Summary writing tasks

Alderson (2000) suggests that when designing a summary task, if the summarizers are informed of a purpose of summarization and provided with clearly-defined task requirements, the subjectivity of scoring can be reduced. In the present study, orientation sheets were provided before each task session to explain the requirements of the summary task. For each source argumentative text, the participants had to summarize the author's position on the issue, major arguments that oppose the author's position, and the author's rebuttals of the opposing arguments. They were also reminded that their summaries should be complete and coherent texts, restating the key arguments of the source texts and excluding personal opinions¹¹.

2. Source texts:

Two argumentative texts, one on the biopharming controversy and the other on the aquarium controversy, were developed as the source texts. The two texts were composed by a native speaker of English who is an experienced EFL writing instructor and had taught summary writing courses prior to this study. Efforts were made to maintain the comparability of the two source texts. Both texts were about 800 words in length and consisted of 6 paragraphs. The syntactical complexity of the two texts was measured by the Flesh Reading Ease Score (39.1 and 39.2) and the Flesch-Kincaid Grade Level (both at Level 12), and the

¹¹ Summaries fall into different categories according to the classification schemes proposed by different researchers. In this study, the participants were required to write "reader-based" summaries (i.e., a stand-alone text written for an external audience, as defined in Hidi & Anderson, 1986), and as they had only to summarize the gist without adding comments, the summaries produced belong to the genre of "précis", following Bleck's (2003) classification (cited in Liu, 2004).

lexical diversity was measured by the type/token ratios (.47 and .43) and the percentages of words appearing in A. Coxhead's *Academic Word List* (8.15% and 5.38%)¹².

To determine the main ideas in each source text, five EFL writing instructors, consisting of three native speakers of English (hereafter abbreviated as NS) and two non-native speakers (hereafter abbreviated as NNS), were asked to write a summary of each source text within the same time limit as specified in Table 1. The eight most frequently occurring arguments (including the author's stated position, opposing arguments and the author's rebuttals of those arguments) were regarded as the "prespecified main ideas" of each source text. One of the source texts used in this study is provided in Appendix 1.

Data Analysis Procedures

As mentioned in the previous section, the participants' performances on the summary writing tasks were compared according to three sets of criteria: (1) content-related criteria for analyzing the number and types of text propositions abstracted, (2) writing-related criteria, and (3) a holistic scale for rating overall quality. The three sets of criteria were operationally defined as follows:

(1)-(3): Content-related criteria

- (1) **Main Ideas (MI):** the coverage rate of "prespecified main ideas" identified by the five EFL writing instructors. The following unweighted partial credit system (cf. Yu, 2007) was adopted to assign Main Ideas scores of students' summaries:
 - 2 points for a main idea adequately restated
 - 1 point for a main idea included but inadequately restated or misplaced
 - 0 point for no inclusion of the main idea(The maximum score per summary was 16 points; Recurring statements of a main idea could be credited only once.)
- (2) **Extraneous Ideas (EI):** the percentage of T-units¹³ in a summary that should be excluded as unimportant ideas (with reference to the "prespecified main ideas"). A high percentage of extraneous ideas is an indicator of inadequate ability to summarize.
- (3) **Inaccurate Statements (IS):** the percentage of T-units in a summary which were factual inaccuracies or over-generalizations. As in the case of extraneous ideas, a high percentage suggests unsatisfactory performance.

(4)-(6): Writing-related criteria

- (4) **Paraphrase & Integration (P&I):** Friend's (2001) coding scheme was refined to assign points to each T-unit in a summary according to the amount of the text from the original subsumed and the quality of paraphrase or integration. The scheme included seven coding categories, as operationally defined in Table 2.

¹² Coxhead, A. (2000). A new academic word list. *TESOL Quarterly* 34(2): 213-238.

¹³ A T-unit is generally defined as "one main clause with all subordinate clauses attached to it" (Hunt, 1965, p.20).

Table 2. *Coding Scheme for Paraphrase & Integration (P&I)*

P&I Coding Categories	Definition	Points
Verbatim Replication	Copying the whole sentence or more than five consecutive words from a sentence in the source text	0
Unsuccessful Reproduction	A single sentence poorly rewritten	1
Successful Reproduction	A single sentence well rewritten	4
Unsuccessful Combination	Two or more sentences (from one paragraph) poorly integrated	2
Successful Combination	Two or more sentences (from one paragraph) well integrated	5
Unsuccessful Integration	Ideas from more than one paragraph poorly integrated	3
Successful Integration	Ideas from more than one paragraph well integrated	6

To correct for the length of the summary, the sum of the points was divided by the *T* score of the number of T-units in the summary (excluding T-units judged as extraneous ideas).

- (5) **Rhetorical Features (RF)**: an analytical rating of the organization, coherence and cohesion of a summary using a six-band scale (1-6) adapted from the Michigan Writing Assessment Scoring Guide (Hamp-Lyons, 1991).
- (6) **Language Control (LC)**: an analytical rating of the lexical and grammatical usage in a summary using a six-band scale (1-6) adapted from the Michigan Writing Assessment Scoring Guide (Hamp-Lyons, 1991).
- (7) **Holistic Summary Quality (HSQ)**: Yu's (2007) holistic scale was adopted for rating the overall quality of the summaries, relating to both content and language. The scale contains six bands (A-F) and a three-level argumentation method (e.g., A-, A=, A+) was used, so the maximum score is 18.

The 92 summaries produced by the 46 students were typed and word processed; no attempts were made to correct spelling, punctuation, or grammatical mistakes. Two teams of raters were involved in the evaluation process. The first team of raters, including an EFL writing instructor (NS) and a research assistant (NNS), scored and coded the summaries according to the criteria of Main Ideas, Extraneous Ideas, Inaccurate Statements and Paraphrase & Integration. About 25% of the summaries were scored by both raters to establish an inter-rater reliability of .80 for each criterion. The remaining portion of the summaries were scored by one rater. The second team of raters, including two EFL writing instructors (one NS and one NNS), evaluated the overall quality of the summaries (HSQ scores) and provided analytical ratings on Rhetorical Features and Language Control. These two raters scored all of the 92 summaries. The inter-rater reliability (Spearman correlation coefficient) varied from .79 for the criterion of Holistic Summary Quality to .85 for the criterion of Language Control. Discrepancies were resolved by the researcher as the third rater¹⁴.

¹⁴ When the raters' band scores differed by one band, the scores were averaged. When they differed by two or more bands, the researcher determined the final score.

RESULTS AND DISCUSSION

Effects of proficiency level: an overview

The independent variable investigated in this study was the proficiency level (advanced vs. high-intermediate). Nine dependent variables were used to measure various aspects of the summaries written by university students at the two proficiency levels. The nine variables included the seven evaluation criteria operationally defined in the previous section, as well as the average numbers of words and T-units per summary. Statistical treatment of the data involved a series of independent-samples *t*-tests using “proficiency level” as the grouping variable. Table 3 presents the descriptive statistics by proficiency level, and Table 4 summarizes the results of independent-samples *t*-tests.

Table 3. *Means and standard deviations by proficiency level*

Dependent Variables	Advanced Level (N=46)		High-Intermediate Level (N=46)	
	Mean	SD	Mean	SD
Average no. of words per summary	179.800	22.469	183.630	18.452
Average no. of T-units per summary	14.890	3.025	15.540	2.641
Main Ideas	13.848	1.548	11.739	2.662
Extraneous Ideas	.365	.491	.914	.841
Inaccurate Statements	.666	.914	.912	.886
Paraphrase & Integration	.874	.184	.583	.197
Rhetorical Features	3.935	.523	3.315	.355
Language Control	3.924	.447	3.098	.442
Holistic Summary Quality	12.848	1.053	10.413	1.514

Table 4. *T-test summary: effects of proficiency level on dependent variables characterizing the written summaries*

Dependent Variables	<i>t</i>	Sig. (2-tailed)	η^2
Average no. of words per summary	-.893	.374	.009
Average no. of T-units per summary	-1.441	.153	.023
Main Ideas	4.644**	.000	.193
Extraneous Ideas	-3.815**	.000	.139
Inaccurate Statements	-1.308	.194	.019
Paraphrase & Integration	7.342**	.000	.375
Rhetorical Features	6.647**	.000	.329
Language Control	8.910**	.000	.469
Holistic Summary Quality	8.953**	.000	.471

**significant at $p < .01$

As can be seen in Tables 3 and 4, the summaries written by the high-intermediate level participants were longer and contained more T-units than those written by the advanced-level participants, although the differences were not statistically significant. Furthermore, gauging the participants' performances against the seven criteria in the evaluation framework, the advanced-level participants significantly outperformed ($p \leq .0005$) the high-intermediate level participants on six out of the seven criteria, except for the criterion of Inaccurate Statements ($p = .194$). These results suggest that compared to their high-intermediate level counterparts, the advanced-level participants tended to write summaries which were slightly

shorter, but generally more accurate, concise and coherent.

This initial finding awaits further fine-grained analyses. However, before continuing to comparison of the performances of the two proficiency groups in the aforementioned aspects of summary writing, we have to examine whether the use of two source texts, a moderator variable in this study, resulted in any significant performance differences or interactions between the variables, both of which might complicate the interpretation of the results.

Effects of source texts

Due to the limited number of participants who fit the sampling requirements, this study required each participant to write two summaries based on different source texts, with the intention of eliciting more data for analysis. As described in the previous section, several measures were adopted to ensure that the two source texts were comparable. In order to examine the effects of source texts on the dependent variables characterizing the written summaries, a series of 2 x 2 Mixed Design ANOVAs were conducted with “proficiency level” as the between-subjects variable and “source text” as the within-subjects variable. Only a moderate effect¹⁵ of source texts was observed for the dependent variable of Extraneous Ideas ($F = 4.739, p = .035, \eta^2 = .097$).

As another measure of the effects of source texts, the post-task questionnaires included questions on how familiar the participants were with the topics of each source text. Three options were provided for them to indicate the degree of familiarity: (1) I am quite familiar with the topic; (2) I have heard or read about the topic once or twice; (3) I am not familiar with the topic. The percentages of participants who selected each option are presented in Table 5. A series of one-way ANOVAs were conducted to test the effects of topic familiarity on the dependent variables characterizing summary performances. No significant effects were found.

Table 5. *Participants' familiarity with the topics of the source texts*

	biopharming	aquarium
Quite familiar	2.17%	2.17%
Have heard/read about it once or twice	54.35%	63.04%
Not familiar	43.48%	34.78%

The above ANOVA results indicate that the use of different texts had little intervening effect on the participants' performances.

Profiling the summary writing performances of Taiwanese university students at the high-intermediate and advanced levels of proficiency

This section reports the findings from a series of comparative analyses on the performances of the two groups of participants, focusing on the content- and writing-related aspects of their summaries, as well as the overall quality of summarization.

1. Level differences in the content-related aspects of summaries

Among the three content-related dependent variables, statistically significant differences were found on two variables: Main Ideas ($t = 4.644, p \leq .0005$) and Extraneous Ideas ($t = -3.815, p \leq .0005$). The high-intermediate participants generally included fewer main ideas and more extraneous ideas in their summaries than did the advanced-level participants, as shown in the differences in their average Main Ideas Scores (11.739 and

¹⁵ According to Cohen's (1988) criteria, an η^2 value ranged from .01 to .058 indicates a small effect; an η^2 value ranged from .059 to .138 indicates a moderate effect; and an η^2 value higher than .138 indicates a large effect.

13.848 for the high-intermediate and advanced groups, respectively) and in the percentages of T-units judged as extraneous ideas (.914 and .365). No significant level difference was detected on the dependent variable of Inaccurate Statements ($t = -1.308, p = .194$).

With respect to the variable of Main Ideas, an item analysis was conducted to further examine what types of main ideas were more likely to be omitted or misrepresented by the high-intermediate level participants. As mentioned in the research design section, eight “prespecified main ideas” were identified through textmapping procedures for each source text. These main ideas were classified into three types according to their argumentative functions:

- Type I: the author's overall position on the issue
- Type II: an opposing argument
- Type III: the author's rebuttal of an opposing argument

Based on the performance data on the high-intermediate group, mean item scores were computed for the eight "prespecified main ideas" in each source text, and the results are presented in Table 6.

Table 6. *Mean item scores on Main Ideas obtained by high-intermediate participants*

	MI 1	MI2	MI 3	MI 4	MI 5	MI 6	MI 7	MI 8
Text A (biopharming)	1.830	1.650	1.703	1.782	1.265	1.617	1.096	.967
Type	<i>I</i>	<i>II</i>	<i>III</i>	<i>II</i>	<i>III</i>	<i>II</i>	<i>III</i>	<i>III</i>
Text B (aquarium)	1.887	1.651	1.744	1.745	.930	.893	1.587	1.131
Type	<i>I</i>	<i>II</i>	<i>III</i>	<i>II</i>	<i>III</i>	III	<i>II</i>	<i>III</i>

An interesting finding derived from the data in Table 6 is that for both source texts, the three main ideas with the lowest mean item scores were all related to the author's rebuttals of the opposing arguments, indicating that the high-intermediate participants were more likely to omit or misrepresent the rebuttals than to do with the author's overall position and major opposing arguments. This finding corresponded to the results reported in Kintsch (1990) and Cohen (1994), that less proficient summary writers, who might be able to identify the topical information and overall thesis, often fail to differentiate finer levels of importance in the information gleaned, an indicator of inadequate competence in recognizing the hierarchical structure of text propositions.

Regarding the variable of Extraneous Ideas, Friend (2000) points out that novice summarizers tend to be distracted by what Garner & Gillingham (1989) labeled as “seductive details”—unimportant details that have the effect of enlivening an essay. In this study, several instances of this phenomenon were also found in the summaries written by the high-intermediate level participants.

2. Level differences in the writing-related aspects of summaries

2.1 Paraphrase and integration of ideas

As described in the section on research design, each T-unit in the participants' summaries was coded and scored according to the Coding Scheme for Paraphrase & Integration (P&I) summarized in Table 2. A total P&I score was computed for each summary to provide quantitative information about how well the ideas from the source text were paraphrased and integrated. An independent samples *t*-test was conducted to compare the mean P&I scores obtained by the high-intermediate and advanced groups. The results, as

presented in Table 4, showed a significant difference by proficiency level ($t = 7.342$, $p \leq .0005$) with a fairly large magnitude of effect ($\eta^2 = .375$).

To further compare how the two groups of university students processed and combined ideas from a source text, the percentages of different P&I types in each summary were computed¹⁶; the group means of the seven P&I types are listed in Table 7.

Table 7. *Mean percentages of P&I types by proficiency level*

P&I Types	Level	
	Advanced (N=46) Mean	High-Intermediate (N=46) Mean
Verbatim Replication	.047	.133
Unsuccessful Reproduction	.015	.064
Successful Reproduction	.301	.262
Unsuccessful Combination	.064	.106
Successful Combination	.363	.229
Unsuccessful Integration	.017	.029
Successful Integration	.116	.021

A 2 x 7 Mixed Design ANOVA was conducted with “proficiency level” as the between-subjects variable and “P&I type” as the within-subjects variable to test differences in the students' use of P&I types. The results showed significant effects of both the between-subjects variable (proficiency level, $F = 14.474$, $p \leq .0005$, $\eta^2 = .139$) and the within-subjects variable (P&I type, $F = 86.07$, $p \leq .0005$, $\eta^2 = .489$). A significant interaction effect between the two variables was also observed ($F = 10.921$, $p \leq .0005$, $\eta^2 = .108$). A further analysis of between-subjects simple effects pointed to significant performance differences between the two proficiency groups on the use of three P&I types: Verbatim Replication, Successful Combination and Successful Integration. These differences are discussed below.

First, with respect to the category of Verbatim Replication, the coding criterion, as described in Table 2, was whether a T-unit contained more than five consecutive words from a sentence in the source text. It was found that the high-intermediate level participants relied significantly more on copying than did their advanced-level counterparts ($F = 8.80$, $p = .03$). This problem was frequently reported in the summary research in the 1990's (Campbell, 1990; Johns & Mayes, 1990; Moore, 1997), and as the above result suggests, still deserves the continuing attention of EAP instructors.

Significant differences between the two proficiency groups were also observed on the percentages of T-units marked as Successful Combination ($F = 21.56$, $p \leq .0005$) and Successful Integration ($F = 10.78$, $p = .001$). The former involves combining ideas from sentences within a paragraph, and the latter involves across-paragraph integration of ideas. At both levels of processing, the advanced group outperformed the high-intermediate group. The following two pairs of samples demonstrate the qualitative differences in their performances.

¹⁶ To compute the percentage of a particular P&I type in a summary, the number of T-units coded as that particular P&I type was counted and then divided by the total number of T-units in the summary, with the T-units judged as extraneous ideas excluded.

Examples 1 & 2: Unsuccessful and successful combinations of ideas within a paragraph

<p>Source text: If this technology works as intended, it could save an astounding amount of lives. Only 200 acres would be needed to produce enough vaccines to vaccinate every baby on Earth against hepatitis B, a disease which kills more than a million people annually. Even greater benefits could be had by the developing world. Because the plants could possibly be dried and shipped without the need of cold storage, the process of getting the vaccines to those in need could be drastically simplified. Furthermore, some scientists claim that plant-based vaccines which could be taken orally could be created, which would further reduce the cost of administering the vaccines.</p>
<p>Example 1 (Writer's code: HI-03) <i>Only smaller growing fields and less costs can those genetic modified plants easily conveyed and produced, unlike most of the corps in market.</i></p>
<p>Example 2 (Writer's code: AD-02) <i>If the technology succeeds, it can help millions of people by manufacturing at low cost medical products that are easily processed & administered.</i></p>

Examples 3 & 4: Unsuccessful and successful integrations of ideas across paragraphs

<p>Source text: Appendix 1</p>
<p>Example 3 (Writer's code: HI-18) <i>Although it is very economical and convenient [para. 2], biopharming has the great risk of cross pollination, which lets a biopharm crop and a non-biopharm crop <u>produce seeds</u> and their crops <u>containing doses of a drug but identical to regular ones</u>. [para. 3]</i></p>
<p>Example 4 (Writer's code: AD-01) <i>Biopharming is a new method to produce medical products that could help human efficiently [para. 2], yet the author disapproves by stating that it causes cross-pollination, which might pollute other plants [para. 3] and affect animals and food chain [para. 4].</i></p>

The writers of these four examples all attempted to generate one-sentence summaries of ideas either from within a paragraph or from more than one paragraph. However, in the case of high-intermediate level writers, these attempts were often undermined by inadequate lexical and grammatical competence, as demonstrated in Examples 1 and 3. In Example 1, the writer seemed to understand the global meaning of the paragraph, but syntactical errors (e.g., inaccurate passive forms), inappropriate word choice (e.g., *conveyed*) and collocation errors (e.g., *growing fields*) resulted in a restatement which cannot be understood with ease. In Example 3, the difficulties in reformulating ideas in his/her own words probably led to the writer's heavy use of verbatim replications, marked with underlining. These two examples highlight the possible effects of lexical and syntactical deficiencies on the summary writing performance of the Taiwanese university students at the high-intermediate level.

In comparison with the two high-intermediate examples, Examples 2 and 4, written by two advanced-level participants, demonstrate better paraphrasing and integration skills. In Example 2, all of the main points in the source paragraph have been incorporated into a well-written complex sentence and expressed predominantly in the summarizer's own words. In Example 4, the main ideas from three paragraphs in the source text have been condensed to form the thesis statement of the summary. Errors in grammatical usage and collocation do exist (e.g., *help human efficiently*) but do not obscure or distort the meaning.

As shown in the above discussions, statistical analysis found significant differences between the two proficiency groups in the amount of verbatim replications and in their abilities to integrate ideas from different parts of the text. An analysis of summaries written by the high-intermediate level participants indicated that lexical and grammatical deficiencies might hamper their attempts to synthesize ideas. Interestingly, the post-task questionnaires revealed that the participants at both proficiency levels were aware of the constraint of limited vocabulary, particularly the difficulties in using synonyms, during the process of summarization, but few of them identified problems with syntactical structures as a source of difficulty. The discrepancy between performance and self-report data suggests an avenue for future research.

2.2 Rhetorical Features and Language Control

In addition to Paraphrase and Integration, this study investigated two more aspects of students' summaries, Rhetorical Features and Language Control. The participants' performances in these two aspects were measured by two separate 6-band analytical scales adapted from the Michigan Writing Assessment Scoring Guide (Hamp-Lyons, 1991). The scale for evaluating Rhetorical Features incorporated key indicators such as cohesion, coherence, organization, and rhetorical fluency, and the scale for rating Language Control focused on how well the grammatical structures and vocabulary chosen could express the ideas.

The results of independent samples *t*-tests, as summarized in Table 4, indicated that the advanced-level participants scored significantly higher than their high-intermediate counterparts on both Rhetorical Features ($t = 6.647, p \leq .0005$) and Language Control ($t = 8.910, p \leq .0005$). The values of η^2 (.329 and .469) indicated a fairly large magnitude of level effect on both dependent variables.

As shown in Table 3, the mean scores of the advanced group on both scales were close to 4, and the mean scores of the high-intermediate group on the two scales were either slightly above or close to 3. Based on the above patterns of score distribution, the following two summaries were chosen to illustrate typical summarization performances of the advanced and high-intermediate groups, as analyzed with reference to the scales for rating Rhetorical Features and Language Control.

Example 5

(Writer's code: HI-22; Score on Rhetorical Features: 3; Score on Language Control: 3)

Nowadays, people had argued over the biopharming, which is a project of United States to produce medicine by farming corn, rice, and soybeans. If biopharming success, it could saved more lives in the developing world. *Otherwise*, every baby on Earth against hepatitis B will be saved because of this new technology. The scientists also claim that the cost of administering the medicine can decrease a lot.

However, this new technology run the risks. The first significant risk is cross-pollination. If the testing plants contact with the normal plants, the seeds they produce might be got by a consumer and he (or she) might inject it without doctors' permission. The second risk is the animals who eat the testing corns might influence other animals through the food chain. The last but not the least risk is the effect of the micro-organisms, which could be killed by contacting the testing medicine.

[missing transition]

The US government have set strict regulations to avoid cross-pollination, but the US government still sometimes failed. The regulations can prevent human's behavior, but it can't control animals' behavior from eating crops. Therefore, regulations cannot totally avoid any risks happening; it only can lower the risks.

Example 6

(Writer's code: AD-14; Score on Rhetorical Features: 4; Score on Language Control: 4)

Biopharming is a promising technology that uses genetically modified crops to produce the proteins we need for medical purposes. It seems like a blessing for those who need vaccines to vaccinate their descents.

Though it seems so wonderful, some problems are reported. First, cross-pollination makes it difficult to control the dosage of vaccines in biopharm crops. Second, the biosystem may be disturbed after animals ingest these crops. Because of all these uncertainties of nature, the author opposes the idea of biopharming.

The pharmaceutical companies think up lots of ways to regulate the operations of biopharming, and feel that they've done enough. However, the failure in inspecting some test crops shows that the strict regulatory methods don't work. Furthermore, the tracking of crop development is conducted poorly, and can't give definitive answer for those who want to know what happened.

The mechanism of nature is too complex for us to regulate by man-made rules. The risk of hybrid crops and destruction of food chain can't be eliminated by the regulations. Our author think

the drug-making companies should think seriously about the danger before they conduct the open-field testing.

In terms of rhetorical features, Example 5 received a score of 3, which was characterized in the rating scale as having "uncertain rhetorical control" (Hamp-Lyons, 1991: 274). This example reflects two problems frequently found in the summaries produced by the high-intermediate group. The first problem is unbalanced development of ideas. The writer used two long paragraphs (149 words) to summarize the first half of the source text and only one short paragraph (46 words) to abstract the second half, making the first two paragraphs of the summary too detailed and the last paragraph too vague. The second problem is illogical links among ideas, usually resulting from inaccurate use of connectors (e.g., "otherwise" in the first paragraph) or missing transitions (e.g., between the second and third paragraphs).

Example 6, written by an advanced-level participant, received a score of 4 on Rhetorical Features. Compared to the previous example, the summary was stronger in organization, coherence, and cohesion. Its organization was appropriate to the task and most of the textual elements were logically sequenced. Nevertheless, there were still instances of redundancy (e.g., "seems like a blessing" in the first paragraph and "seems so wonderful" in the second paragraph) and missing transitions across paragraphs (e.g., between the second and third paragraphs). Hamp-Lyons (1991: 274) considers occurrences of redundancy, repetition and missing transitions to be indicators of weak "rhetorical fluency." This example suggests that, in terms of rhetorical features, the advanced-level participants still had plenty of room for improvement.

With regard to the criterion of Language Control, the two raters unanimously gave Example 5 a band score of 3 and Example 6 a band score of 4. These two bands were distinct from each other in the amount of errors and the extent to which the errors interfered with the clarity of the summary. Example 5 contains more lexical and grammatical errors which obscure or alter the meanings of the source text. This result demonstrates again that lexical and grammatical deficiencies can affect the summary writing performance of high-intermediate university students in Taiwan.

3. Level differences in the overall quality of summaries

Thus far, we have compared the content- and writing-related aspects of the summaries written by the high-intermediate and advanced level university students participating in this study. This section will complete the comparative analysis by focusing on the overall quality of their summaries.

As mentioned in the research design section, Yu's (2007) holistic scale, with six bands (A-F) and a maximum score of 18, was adopted for assigning scores on Holistic Summary Quality (HSQ). The statistical data, as summarized in Tables 3 and 4, once again revealed a significant difference by proficiency level. The mean HSQ score obtained by the advanced group was 12.848, and the mean HSQ score of the high-intermediate group was 10.413. *T*-test results indicated that the difference was statistically significant ($t = 8.953, p \leq .0005$) and the magnitude of level effect was fairly large ($\eta^2 = .471$).

Yu's (2007) holistic scoring scheme, as reviewed earlier in this paper, incorporates both the content and language dimensions of written summaries. The general guidelines are as follows (citing Yu, 2007, p.567):

- faithfulness to the source text, in terms of the percentages of right and wrong statements in the summary,
- the topological relationships between the summary and the source text, with emphasis on the use of the summarizer's own language and the integration and connections of the statements,

- the conciseness, coherence and logicity of communication of meaning,
- the overall difficulty that the raters might encounter.

Using the above general guidelines and the detailed rubrics provided by Yu (2007: 568-572), the raters gave an average HSQ score of 10 (equivalent to a C-) to Example 5 and a score of 13 (equivalent to a B-) to Example 6. According to the raters' comments, collected in follow-up interviews, both Examples 5 and 6 covered most of the prespecified main ideas, but had a couple of misinterpretations and omissions. Example 6 included more extraneous ideas than Example 5 did. However, it was the writer's inadequate control over rhetorical features and grammatical/lexical usage, as discussed on the previous page, that resulted in the lower score given to Example 5.

Another issue of interest is how the raters utilized the scoring rubrics to evaluate overall summary quality. During the post-evaluation interviews, the two raters were asked how they applied the criteria specified to assign a holistic score. One of the raters reported that she first compared the summary with the source text and the accompanying list of prespecified main ideas to check the accuracy of content coverage and spot instances of verbatim replication. After the first reading, she decided an initial score range. To finalize the rating, she then evaluated the summary for its writing quality. The other rater adopted a different approach. After the first reading, he gave an impressionistic score based on the writing quality of the summary, paying particular attention to its comprehensibility and clarity. Then he read the summary again to examine its faithfulness to the source text and the extent to which ideas were reformulated and integrated.

Although these two raters approached the evaluation task in distinctive ways, both of them endorsed the idea of analyzing summaries based on a framework covering both the dimensions of content coverage and writing quality. They noted that diagnostic information about the learners' strengths and weaknesses in different aspects of summary writing can help course designers and instructors to develop materials and pedagogies tailored to the needs of their target students.

CONCLUSION

This study was designed to compare the summary writing performances of university students in Taiwan at the high-intermediate and advanced levels of proficiency. The results of statistical analysis indicated that the summaries written by the advanced-level university students participating in this study were generally more accurate, concise, and coherent than those written by the participants at the high-intermediate level. More fine-grained analyses showed the following performance characteristics:

- (1) In the dimension of content coverage, the high-intermediate level students, compared to the advanced-level students, included fewer main ideas and more extraneous ideas in their summaries. Although they were able to identify the topical information and overall thesis of the source text, they experienced difficulty in reconstructing a complete hierarchical structure of the ideas in the source text.
- (2) In the dimension of writing quality, the high-intermediate level students resorted to verbatim replications more frequently and were less skillful at reformulating and integrating ideas, as compared to their advanced-level counterparts. However, both groups needed more training on how to rearrange and reorganize the ideas extracted to compose a coherent summary text with adequate rhetorical fluency.
- (3) The high-intermediate level students were more likely to be constrained by deficiencies in lexical and grammatical competence, which affected the accurate conveyance of main ideas, the coherence and organization of summary texts, and thus the results of overall evaluation.

Based on these performance profiles, the following pedagogical suggestions should be

considered by those who intend to incorporate summary writing into EAP training for university students in Taiwan. With regard to high-intermediate level students, the instructors may wish to provide continued guidance on the use of synonyms and other paraphrasing strategies. Students at this level also need further practice in combining sentences within, as well as across, paragraphs, and the importance of syntactical skills should not be neglected. Furthermore, students at both the high-intermediate and advanced levels need to work on recognizing the logical and rhetorical relationships of ideas in a source text and using appropriate connectors or transitions to represent the macrostructure established in that text.

Finally, it is necessary to note the limitations of this study and suggest avenues for future research. First, since a limited number of university students at only two proficiency levels were sampled, the generalizability of the results is contingent upon validation with more empirical evidence. It is suggested that further comparative analyses of this nature should be conducted with a larger sample of students at a wider range of proficiency levels. Second, this study focused only on the products of summary writing. To gain further insight into the processes and strategies involved in summarization, introspective data should be collected through mentalistic measures (e.g., concurrent or retrospective verbal reports). It is hoped that these lines of research can help generate more research-based materials and pedagogies to enhance the summarization performance of university students in Taiwan.

REFERENCES

- Alderson, J. C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.
- Campbell, C. (1990). Writing with others' words: using background reading text in academic compositions. In B. Kroll (Ed.) *Second language writing* (pp. 211-230). Cambridge: Cambridge University Press.
- Chen, Y. C. (2002). *Writing from Sources: Summary and response writing by four EFL graduate students*. Unpublished master's thesis. Providence University, Taiwan.
- Cohen, J. (1988). *Statistical power analysis for the behavior science*. Hillsdale, NJ: Lawrence Erlbaum.
- Cohen, A. D. (1994). English for academic purposes in Brazil: The use of summary tasks. In C. Hill & K. Parry (Eds.) *From testing to assessment: English as an international language* (pp. 174-204). London: Longman.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly* 34(2): 213-238.
- Enright, M. K., Grabe, W., Koda, K., Mosenthal, P., Mulcahy-Ernt, P., & Schedl, M. A. (2000). *TOEFL 2000 reading framework: A working paper*. Princeton, NJ: Educational Testing Service.
- Friend, R. (2000). Teaching summarization as a content area reading strategy. *Journal of Adolescent and Adult Literacy*, 44(4), 320-329.
- Friend, R. (2001). Effects of strategy instruction on summary writing of college students. *Contemporary Educational Psychology*, 26, 3-24.
- Garner, R., & Gillingham, M.G. (1989). Effects of "seductive details" on macroprocessing and microprocessing in adults and children. *Cognition and Instruction*, 6, 41-57.
- Hamp-Lyons, L. (1991). *Assessing Second Language Writing in Academic Contexts*. Norwood, NJ: Ablex Publishing Corporation.
- Hidi, S., & Anderson, V. (1986). Producing written summaries: Task demands, cognitive operations, and implications for instruction. *Review of Educational Research*, 56(4), 473-493.
- Hsu, H. C. (2003). *A case study of the process of web-based reading and summary writing for EFL college novice writers in Taiwan*. Unpublished doctoral dissertation. National Taiwan Normal University, Taiwan.
- Hunt, K. W. (1965). *Grammatical structures written at three grade levels*.

- Urbana-Champaign, IL: National Council of Teachers of English.
- Johns, A.M. & Mayes, P. (1990). An analysis of summary protocols of university ESL students. *Applied Linguistics 11*, 253-271.
- Kim, S. A. (2001). Characteristics of EFL readers' summary writing: A study with Korean university students. *Foreign Language Annals*, 34(6), 569-581.
- Kintsch, E. (1990). Macroprocesses and microprocesses in the development of summarization skill. *Cognition and Instruction*, 7(3), 161-195.
- Liu, S. H. (2004). *A study on the effects of English summary writing instruction on Taiwanese senior high school students' reading and writing abilities*. Unpublished master's thesis. National Taiwan Normal University, Taiwan.
- Moore, T. (1997). From text to note: cultural variation in summarization practices. *Prospect*, 12, 54-63.
- Pan, Y. C. (2003). *A study of the effects of reading and summarizing model essays on EFL writing*. Unpublished master's thesis. National Tsing Hua University, Taiwan.
- Radmacher, S. A., & Latosi-Sawin, E. (1995). Summary writing: A tool to improve student comprehension and writing in psychology. *Teaching of Psychology*, 22(2), 113-115.
- Rivard, L. P. (2001). Summary writing: A multi-grade study of French-immersion and Francophone secondary students. *Language, Culture and Curriculum*, 14, 171-186.
- van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Winograd, P. N. (1984). Strategic difficulties in summarizing texts. *Reading Research Quarterly*, 19, 404-425.
- Weir, C. J. (2005). *Language testing and validation: An evidence-based approach*. Basingstoke: Palgrave Macmillan.
- Yang, L., & Shih, L. (2003). Exploring six MBA students' summary writing by introspection. *Journal of English for Academic Purposes*, 2(3), 165-192.
- Yu, G. (2005). *Towards a model of using summarization tasks as a measure of reading comprehension*. Unpublished PhD thesis, University of Bristol, Bristol.
- Yu, G. (2007). Students' voices in the evaluation of their written summaries: Empowerment and democracy for test takers? *Language Testing*, 24(4): 539-572.

APPENDIX: Source text (The Biopharming Controversy)

Biopharming is the genetic modification of plants to produce proteins which can be extracted and turned into medicines and vaccines targeted at a host of diseases and conditions which plague mankind today. In the United States, there are hundreds of open-field test crops of corn, rice, soybeans, which promise to yield hundreds of new medical products. While the technology looks promising, closer inspection reveals that not only are these promises far from guaranteed, but biopharming could have disturbing consequences.

If this technology works as intended, it could save an astounding amount of lives. Only 200 acres would be needed to produce enough vaccines to vaccinate every baby on Earth against hepatitis B, a disease which kills more than a million people annually. Even greater benefits could be had by the developing world. Because the plants could possibly be dried and shipped without the need of cold storage, the process of getting the vaccines to those in need could be drastically simplified. Furthermore, some scientists claim that plant-based vaccines which could be taken orally could be created, which would further reduce the cost of administering the vaccines.

Unfortunately, with these potential benefits are great risks, chief among them is the risk of cross-pollination. Cross-pollination occurs when pollen from the flower of one plant comes into contact with another to produce seeds which possess characteristics from both parent plants. The danger is that if cross-pollination occurred between a biopharm crop and a non-biopharm crop, the seeds produced might carry doses of a drug. Those seeds would be identical to regular seeds and would produce grains, fruits, and vegetables which would not show any sign that they contained a drug. Neither the farmer who plants the seeds nor the consumer who later ingests the produce will detect the difference. As a result, the consumer of these

plants could be ingesting medicine that should only be administered in controlled amounts under a doctor's orders. Though pharmaceutical companies claim the quantity would be too small to affect a human, they have not provided evidence supporting this.

Even if a human may be insusceptible, what about smaller animals which ingest these crops? The effects could certainly be greater for animals the size of a rabbit or bird. It is possible that these drugs could build up in the liver of these animals, and later these animals might be eaten by humans, providing a mega-dose of the drug. It is also possible that the dose might kill these animals. If enough animals were affected in a given species, the food chain itself could be disturbed. Other at-risk life forms are micro-organisms, which could be harmed or even killed by even the smallest exposure. Though micro-organisms are small, they play a mighty role in the life of every plant and animal on Earth. Not even the simplest bodily function could occur without them. No one knows with certainty whether micro-organisms would be affected, but if so, the effects could be catastrophic.

Proponents of biopharming feel controlling the risks is within our means. They presume that we can prevent cross-pollination with strict regulations. Indeed, many strict regulations have been created by the US government, covering all aspects of biopharming. The Animal and Plant Health Inspection Service (APHIS), under the jurisdiction of the United States Department of Agriculture (USDA), has been charged with enforcing them. One of the regulations is that every biopharm test crop must be inspected by APHIS. Nevertheless, in December 2005, the USDA released a report that revealed APHIS had not only failed to inspect some test crops but were not even aware of the locations of others. Furthermore, though the tracking of each crop through to their harvest is also required, APHIS did not make any inspections beyond an initial one on several crops. As a result, APHIS can not definitively say what happened to them. Pharmaceutical companies, on the other hand, claim that despite the APHIS' failures, they have complied with the regulations. However, records indicate otherwise, as several companies have already been cited for allowing cross-pollinations.

In fact, cross-pollination as a result of poorly enforced regulations is not the central issue for critics of biopharming. For them, what the argument boils down to is that no man-made regulation can possibly account for the unpredictability of nature. Regulations cannot prevent birds, animals, insects, or the wind from spreading pollen and seeds. They cannot prevent birds or animals from eating seeds. They cannot ensure that seeds are not left behind to sprout long after experiments are over or prevent seeds from falling off trucks during transport. Regulations can not eliminate these risks. For this reason alone, the biopharm industry needs to seriously rethink the open-field testing of biopharm crops.

The Application of Summarization Technology in an EFL Classroom

Chia-hui Chiu (邱嘉慧)

Tunghai University
chiahuichiu@thu.edu.tw

The WriteToLearn web-based program has been developed to provide an electronic environment, which allows students to write multiple drafts of online summaries of reading assignments and receive an immediate automated feedback of how well they have captured the main idea of the reading passage in their summaries. The feedback helps students to form an appropriate macrostructure of the text they read. In order to learn knowledge in various content areas, most Taiwanese college students are more or less required by their content-area course teachers to read English textbooks. While summarization has been identified as one of the most important skills in reading texts, most Taiwanese college students have not yet developed adequate summarization skills. Seeing the need to emphasize the teaching of summarization skills in EFL classrooms, the present study looks at the implementation of WriteToLearn in a college-level EFL classroom. The aim is to examine the effects of WriteToLearn on EFL learners' reading comprehension, summarizing skills and self-concept as readers.

INTRODUCTION

Many EFL learners have great difficulty with comprehension even when the percentage of new vocabulary in the passages is low. It is often the case that poor EFL readers tend to focus on vocabulary at the expense of comprehending meaning when reading (Koda, 2004). Deficits in reading processes that are fundamental to comprehension, such as the ability to locate main ideas from textual materials, are commonly seen in poor EFL readers. Problems in identifying main ideas are especially evident when expository or informational text is considered because poor readers are generally less aware of text structure and text-processing strategies (Jitendra, Cole, Hoppes & Wilson, 1998).

In Taiwan, most college students are required to read English textbooks to gain knowledge in their content-area courses (e.g., History of Art, Biology). In a needs analysis carried out at the end of the fall semester of 2008 in the researcher's non-English-major Freshman English class, students revealed that they had difficulty comprehending the material assigned by their content-area course teachers. While they in general liked the communicatively oriented instructional method in which they had been provided with ample opportunities in their EFL class to develop and use skills while completing assigned tasks (e.g., oral presentations, dialogue and play contests, reading novels), they considered themselves poor readers when required to read English textbooks. While researchers (e.g., Caccamis & Snyder, 2005) have recognized the importance of summarization in promoting reading comprehension, most Taiwanese college students have never received sufficient training nor developed adequate ability in summarization. Seeing the need to emphasize the teaching of summarization skills in EFL classes, the present study looks at the implementation of WriteToLearn, an on-line tutoring software designed by Pearson Technologies in an EFL classroom in Taiwan.

RELATED LITERATURE

In his theory of text comprehension, Kintsch (1998) assumes that text information is represented in multiple levels including surface linguistic forms, meaning-based propositional networks, and conceptually driven situation models. The basic unit of text representation is the proposition, defined as the smallest semantic element. Each proposition contains a relational term, that is, the predicate, and one or more arguments. An example from Koda (2004) is “Bill bought a beautiful sweater”. This sentence consists of two propositions: “Bill bought a sweater” and “The sweater is beautiful.” To construct the meaning conveyed by a text, individual propositions must be interlinked on the basis of argument overlap as well as other clues signaling their semantic connections. The two propositions just mentioned are linked through their shared argument “sweater”. The propositional textbase, moreover, has a hierarchical structure: microstructure and macrostructure. According to Kintsch, the microstructure is a network of propositions representing meaning at the sentence level of a text; it includes both text-based propositions (also referred to as micropropositions) and those propositions generated from long-term memory to create a local understanding at the sentence level. The macrostructure of a text, on the other hand, contains summary propositions (also referred to as macropropositions) and organizes the micropropositions according to their relative relevance to the main text theme, and in so doing denotes the global structure of the text. In short, micropropositions are generated through analysis of the text’s surface; whereas macropropositions are formed through the utilization of reduction principles or summarization rules. Kintsch further contends that for comprehension and memory, the gist of a text expressed by the macrostructure is usually what matters most.

Several studies (e.g., Jitendra, Cole, Hoppes & Wilson, 1998; Jitendra, Hoppes & Xin, 2000; Rogevich & Perin, 2008) have shown that explicit teacher-mediated summarization instruction can effectively help students with learning disabilities identify main ideas and promote their reading comprehension. This kind of intervention is largely based on Kintsch’s theory of text comprehension (1998), which asserts that a good summary of a text contains main ideas and hence comprises the text macrostructure. Thus, the reader’s summarization ability of a text can be seen as a reliable indicator of his/her reading ability. For example, in their study on the effectiveness of a summarization strategy and self-monitoring instructional procedure for improving comprehension of textual material in students with high-incidence disabilities, Jitendra, Hoppes and Xin (2000) randomly assigned thirty-three middle school students with disabilities to experimental and control groups. Students in the experimental condition were trained to identify main idea statements using summarization strategy and a self-monitoring procedure. Results indicated that the instructional procedures led to increased reading comprehension by students in the experimental group on the training measure, which was maintained over time. On near and far transfer measures, the experimental group statistically outperformed students in the control group on posttest and delayed posttest items requiring selection responses. Students in the experimental group maintained strategy usage 6 weeks later on selection type responses on the near transfer measure but not on the far transfer measure.

THE WRITETOLEARN PROGRAM AND THE PRESENT STUDY

Based on reading comprehension theory, the WriteToLearn program, which was designed to support the key learning strategy, constructing a written summary of the instructional text, was implemented in the present study. The precursor of WriteToLearn is Summary Street. An article by Wade-Stein and Kintsch (2004) provided detailed explanation on how Summary Street works. According to them, the mechanism underlying Summary Street was based on latent semantic analysis (LSA), a powerful method that enables computers to evaluate the content similarity of texts in ways that resemble human judgments

(Landauer & Dumais, 1997). Wade-Stein and Kintsch (2004) explain that LSA represents the meaning of words, sentences, and whole texts mathematically, as vectors in high-dimensional semantic space. This semantic space makes it possible to calculate the semantic distance between the meanings of two different words or texts. Texts with similar meanings are located near one another in the space, and words that occur in similar contexts are grouped together as semantically related. The measure of semantic distance is the cosine, which can be interpreted as a kind of correlation. The major drawback of LSA is that it does not take into account syntax and does not model precise, logical thinking. When LSA applied to Summary Street, a cosine can be computed between the summary a student has written and the original text. If the cosine is below a certain empirically determined threshold value, Summary Street tells the students that his or her summary does not yet cover the content sufficiently well at the levels of section and the text as a whole. In addition, the redundancy and relevance checks help reduce the length of a too long summary. The sentences with a cosine that exceeds a redundancy threshold are flagged with the same colors so that the summary writer can locate portions of their summaries that are redundant and/or unrelated to the text and make appropriate changes. The software hence is useful in providing students with extensive summarizing practice without increasing the teacher's workload. The results of the classroom trials as demonstrated in Wade-Stein and Kintsch's study (2004) indicated that thousands of elementary and high school students in the United States have used *Summary Street* and found using *Summary Street* motivating, rewarding and fun. Test results also indicated significant improvement in their reading comprehension.

The use of technology in facilitating language learning has received a lot of attention from researchers. Seeing the great promise of WriteToLearn and the lack of discussion on summarization in promoting EFL learners' reading comprehension, the present study aims at investigating whether and how WriteToLearn can be useful in promoting EFL learners' reading comprehension and summarization skills, and also in improving their self-perception as readers. More specifically, the present study is designed to answer the following research questions.

1. How much do the EFL learners improve their summarization skill after a four-week summarization training session using WriteToLearn?
2. Do the EFL learners improve their reading comprehension after a four-week summarization training session using WriteToLearn?
3. Do the EFL learners change their self-perception readers after a four-week summarization training session using WriteToLearn?
4. How satisfied are the EFL learners with the implementation of WriteToLearn?

The following sections describe the participants of the study, the implementation of WriteToLearn in the EFL class, instruments used to collect data, data collection procedures, results and findings and conclusion.

Participants

Thirty-five college-level EFL learners participated in this study. They were students from the College of Management majoring in International Trade, Business Administration, Finance, Accounting, and Information Management in a university located in central Taiwan. They were placed in a high-level Freshman English class based on their performance on the placement exam they took when entering the university. They took Freshman English as a required course. In a needs analysis conducted at the end of the fall semester of 2008, all of them reported that they were required to read English textbooks (i.e., expository reading material) for at least one of their content-area courses. They also understood that content area reading they did for school differs from reading for pleasure in the Freshman English class because, in the former, they have to understand and remember new information that seems to have been decided by the author and the teacher. Over 88% of them indicated that they had

difficulties understanding the texts they were required to read. Seeing the need to help students become better readers of expository material, the teacher-researcher attempted to remedy the situation by implementing an on-line tutoring program in the Freshman English class in the spring semester of 2009 to promote students' reading comprehension of the expository material, something similar to what they read for their content area courses.

Implementation of the WriteToLearn Program

Educational technologies are most effective when they go hand-in-hand with classroom lessons and activities. The EFL learners participating in the present study had used an English-learning textbook in the fall semester that partially focuses on strategies essential to successful reading comprehension. The useful strategies taught in class included predicting the content of the article to be read through reading the title, the headings, and the pictures; locating the main ideas of each paragraph by reading the topic sentences, supporting sentences/details, and concluding sentences; using graphic organizers to sketch how main ideas are interlinked; and using context clues to identify the meanings of new words. In the spring semester before the implementation of WriteToLearn, a summary writing instruction was given to the EFL learners in the first and the second week with four hours of instruction in total. The focus of the instruction was to briefly review skills they had learned in the past and introduce summary writing. The EFL learners were guided to think about things to do before, during, and after reading; and were told the importance of summary writing for reading comprehension and as a study skill. The teacher-researcher discussed with the EFL learners what they had to know to be able to write a good summary. The EFL learners reviewed the comprehension strategies and gained a basic understanding that they needed to be able to reproduce the basic argument or logic of the text using their own words so that someone else can understand it. Then the EFL learners read a short text independently, practiced highlighting the main ideas and using their own words to rewrite the main ideas and form a summary. Lastly, they compared their summaries with those of their classmates and discussed the example summary in groups and as a class.

After the introductory lessons, the EFL learners were introduced to start to use WriteToLearn in the third week of the spring semester. Students read one of the passages from WriteToLearn, prepared their summary drafts beforehand and then typed their summary drafts into WriteToLearn when they had their first WriteToLearn lesson in the computer lab. During the two-hour lab lesson, they were directed to acquaint themselves with using WriteToLearn. They learned how to log in to the system, view the passages to be read, look up unfamiliar terms using the dictionary tool, enter the summary writing window, use the editing tools, submit their summary drafts and receive the automated feedback (i.e., content coverage, percentage of copying, spelling, redundancy and relevance check). The EFL learners were told that they could have a maximum of six summary writing trials for each passage they read. They would receive and could view the teacher's comments in the summary writing window of WriteToLearn one or two days after their first draft was submitted. They were expected to refine their summary writing after they received the automated feedback as well as the teacher's comments. They were required to read one passage each week (Week 4 through Week 7). The passages were selected by the teacher/researcher from WriteToLearn with half on science-related topics (e.g., The Application of DNA technology) and the other half on social-science-related topics (e.g., Globalization). The selection criteria included the length and the grade-level. The passages selected were between 1000 and 1200 words long (similar to the length of a section of the English textbook the EFL learners need to read for their content-area course) and were considered suitable for American students of the 9-12 grade-level to read (a level that is close to the material the EFL learners read). The EFL learners were required to turn in their first draft to the teacher/researcher along with the automated feedback from WriteToLearn at the

beginning of the week and then a final draft along with the feedback at the end of the week. They were told periodically in class that their goal was to move the feedback bars into the green zone (instead of staying in the red and yellow zones if that was the case). Their summary work is counted as homework for this class and weighs 10% of the final semester grade.

Instruments

Two equivalent forms of the TOEFL reading subtest, a reading self-perception survey and a student satisfaction questionnaire were used to collect data for this study.

Reading Test. To evaluate students' reading improvement after using WriteToLearn, two equivalent forms of the retired TOEFL reading subtest were used as pretest and posttest. The format of the test was multiple-choice with four options. Each form included two passages, one related to science and the other social science. Each form contained 10 macrostructure items with 5 on the science passage and the other 5 on the social science passage; and 10 microstructure items with 5 on the science passage and the other 5 on the social science passage. Macrostructure items refer to questions that require students to choose the best title to the passage, decide what can be inferred from the passage, etc. Microstructure items refer to questions that require students to choose the word that has a meaning close to a specific word in one line of the passage or to choose the word that "it" or "which" in certain lines of the passage refers to. Each item was worth one point. The maximum of the total points earned on each form is 20. The pretest was administered in Week 3 of the semester, and posttest in Week 8.

The classroom lessons and activities between Week 3 and Week 7 were mainly communicatively oriented. Students were instructed and given ample opportunities to complete tasks that focus on speaking and listening skills (e.g., giving individual oral presentations, watching short clips of video and conducting group discussions, writing scripts for in-class play contest, etc.). A limited amount of class time was devoted to reading. Moreover, the material they read was relatively short (550-650 words). Thus, it seems safe to assume students' reading comprehension improvement during the intervention is mainly due to their use of WriteToLearn.

Reading Self-Perception Survey. The reading self-perception survey used in the present study was a modified version of those used in studies by Henk and Melnick (1995) and Chapman and Tummer (1995). It was used to measure how the EFL learners feel about themselves as readers. Self-perception is a multidimensional construct and can be generally defined as a person's judgments of his or her ability to perform an activity, and the effect this perception has on the on-going and/or future conduct of the activity. The measure developed by Henk and Melnick (1995) was called the Reader Self-Perception Scale (RSPS), a tool for measuring how children feel about themselves as readers. Based on the characteristics of this age group, they took four basic factors into consideration when developing the scale: progress (how one perceives his/her present reading performance to compare with past performance), observational comparison (how one perceives his/her reading performance to compare with the performance of classmates), social feedback (how one perceives direct or indirect feedback from teachers and classmates on his/her reading performance), and physiological states (internal feelings one experiences during reading). In their development of reading self-concept in young children, Chapman and Tummer (1995) identified three factors: competence (beliefs regarding ability and proficiency in reading tasks), attitude (feelings toward and affinity for reading) and difficulty (beliefs that reading activities are hard or problematic). Considering the characteristics of the EFL learners participating in the present study, the researcher developed a reading self-perception survey that tapped the following four dimensions: physiological states, progress, attitude and competence (see Appendix A for sample items). There was a total of twenty 5-point likert scale items with 1 indicating

“strongly disagree”, 2 “disagree”, 3 “undecided”, 4 “agree”, and 5 “strongly agree” to the statements. The survey was distributed to the learners in Week 3, before their use of WriteToLearn, and in Week 8. Again, it is assumed that students’ changes in their self-perception as readers during the intervention are mainly due to their reading and summary writing practices through WriteToLearn because a limited amount of class time was devoted to reading during this period of time.

Student Satisfaction Questionnaire. The student satisfaction questionnaire designed by the researcher included two 5-point likert scale items (with 1 indicating “strongly disagree”, 2 “disagree”, 3 “not sure”, 4 “agree”, and 5 “strongly agree” to the statements) and three open-ended questions (see Appendix B for the questionnaire). The two likert-scale items were to measure the participants’ satisfaction regarding the usefulness of WriteToLearn in terms of helping them improve their reading comprehension and summary writing skills. The open-ended questions were to elicit their opinions on the strengths and the weaknesses of the use of WriteToLearn in the Freshman English class and what the teacher could do to facilitate their use of WriteToLearn.

Data Collection Procedures

Pretests of the TOEFL reading subtest and the reading self-perception survey were administered at the same time in class during the third week of the spring semester of 2009 (before the summarization instructions began). Posttests were administered in the eighth week. The EFL learners were asked to complete the reading test and the reading self-perception survey within 30 minutes. The student satisfaction questionnaire was also distributed in the eighth week right after the EFL learners completed the reading test and the reading self-perception survey. There was no time limit when the students were writing the questionnaire.

RESULTS AND FINDINGS

The internal consistency estimate, Cronbach’s alpha, computed on the performance of all participants who took the TOEFL reading subtest was .79 and .81 for the pretest and posttest respectively. And the internal consistency estimate computed on the responses of all participants who completed the reading self-perception survey was .87 and .85 for the pretest and posttest respectively.

To answer the first research question: “How much do the EFL learners improve their summarization skill after a four-week summarization training session?”, the researcher examined the first summary drafts the EFL learners wrote every week for weeks 4 through 7 and calculated the number and the percentages of the drafts that reached the threshold levels (i.e., the green zone, the good zone) in four different aspects when evaluated by WriteToLearn in terms of the quality of the content. Table 1 shows the results. The percentages of first summary drafts that reached the threshold level of content coverage of a text as a whole increased from 37.1 to 60.6. Clearly, the EFL learners made substantial improvement in this aspect of their summary writing but there were still around 40% of the EFL learners struggling to write more in their summaries. The percentages of first summary drafts that reached the threshold level of no copying slightly increased from 68.6 to 78.8. Most of the EFL knew that when writing their summaries, they needed to restate the main ideas using their own words and avoid copying the sentences of the text they read right at the beginning. The percentages of first summary drafts that reached the threshold level of no redundancy remained almost the same around 90, which means that most of the EFL learners did not write repeated information in their summaries. This is probably because they were still struggling to write more sufficient information to cover the text. Lastly, the percentages of first summary drafts that reached the threshold level of no irrelevance remained low ranging from 14 and 30. A closer look at the EFL learners’ summaries revealed that the EFL

learners tended to include in their summaries their background knowledge or life experiences related to the topic of the text read and digressed from the main ideas of the text.

Table 1. *Numbers (and percentages) of the EFL learners' first summary drafts that reached the threshold levels in four aspects*

	Summary 1 (Week 4)	Summary 2 (Week 5)	Summary 3 (Week 6)	Summary 4 (Week 7)
Number of summary drafts submitted	35	35	34	33
Content coverage of the text as a whole	13 (37.1%)	17 (48.6%)	19 (55.9%)	20 (60.6%)
No Copying	24 (68.6%)	27 (77.1%)	26 (76.5%)	26 (78.8%)
No redundancy	31 (88.6%)	32 (91.4%)	30 (88.2%)	31 (93.9%)
No irrelevance	10 (28.6%)	5 (14.3%)	8 (23.5%)	10 (30.3%)

Note: Thirty-five EFL learners participated in the study. Not all submitted their summary drafts for each week during the implementation of WriteToLearn.

Table 2 shows the means, standard deviations and ranges of the EFL learners' pretest and posttest scores on the reading test as a whole and their subscores of the macrostructure and the microstructure testing components. As can be seen, the posttest score means were 10.57 for the reading test as a whole, 5.54 for the macrostructure items, and 3.93 for the microstructure items, all higher than the pretest score means.

Table 2. *Means (and standard deviations) and [ranges] of the total reading scores, macrostructure subscores and microstructure subscores of the Reading Test*

	Pretest (n= 35)	Posttest (n=35)
Total reading scores	8.23 (2.07) [4-11]	10.57 (2.29) [6-17]
Macrostructure subscores	4.3 (1.44) [1-7]	5.54 (1.44) [2-10]
Microstructure subscores	3.93 (1.39) [1-7]	5.03 (1.34) [2-8]

To answer the second research question: "Do the EFL learners improve their reading comprehension after a four-week summarization training session (as measured by the TOEFL reading test)?", a hypothesis testing of the mean score differences between the EFL learners' pretest scores and posttest scores was conducted. The researcher employed a repeated-measure dependent t-test using SPSS to see if there was a statistically significant mean score differences between the pretest and posttest scores of the reading test. The results of the test, as can be seen in Table 3, revealed that the mean differences between the posttest and the pretest scores of the reading test as a whole, of the macrostructure items and of the

microstructure items were significant ($p < .001$, $p < .001$, $p < .004$ respectively). The results implied that the EFL learners did improve their reading comprehension after a four-week summarization training session.

Table 3. *Dependent t-test: hypothesis testing of the mean differences of the pretest and posttest total scores and subscores of the Reading Test*

	Mean difference (posttest- pretest)	Std. Error	t-values	df	Sig.
Total reading scores	2.34	0.54	4.32*	34	.001
Macrostructure subscores	1.24	0.31	3.94*	34	.001
Microstructure subscores	1.10	0.36	3.06*	34	.004

* The mean difference was significant at the .05 level.

To gain a basic understanding of students' self-perception as readers, descriptive statistics of the results of the pretest and the posttest of the Reading Self-Perception Survey were computed and shown in Table 4. The mean of the pretest scores was 3.41, the same as that of the posttest scores. Given that the mid-point of the scale is 3, the EFL learners as a whole considered themselves as better-than-average readers before and after the implementation of WriteToLearn.

Table 4. *Means (and standard deviations) and [ranges] of the Reading Self-Perception Survey*

Pretest (n= 35)	Posttest (n=35)
3.41 (0.31)	3.41 (0.29)
[2.75-4.05]	[2.80-4.05]

To gain an overview of students' changes in different subscales of their self-perception as readers, descriptive statistics of subscale scores were computed. Table 5 shows the means, standard deviations and ranges of the EFL learners' pretest and posttest total scores on the reading self-perception survey and their total subscale scores on physiological states, progress, attitude and competence. As can be seen, the posttest score mean was 68.21 for the reading self-perception survey as a whole, similar to the pretest score mean (i.e., 68.26). The posttest score mean was 16.57 for the physiological states subscale, a little lower than the pretest score mean of this subscale (i.e., 17.34). The posttest score mean was 18.82 for the progress subscale, slightly higher than the pretest score mean of this subscale (i.e., 18.74). The posttest score mean was 16.25 for the attitude subscale, slightly lower than the pretest score mean of this subscale (i.e., 16.43). The posttest score mean was 16.57 for the competence subscale, higher than the pretest score mean of this subscale (i.e., 15.74).

Table 5 Means (and standard deviations) and [ranges] of the total reading self-perception scores, total physiological states subscale scores, total progress subscale scores total attitude subscale scores and total competence subscale scores of the Reading Self-Perception Survey

	Pretest (n= 35)	Posttest (n=35)
Total reading self-perception scores	68.26 (6.10) [55-81]	68.21 (5.88) [56-81]
Total physiological States subscale scores	17.34 (2.93) [9-21]	16.57 (2.65) [10-21]
Total progress subscale scores	18.74 (2.23) [13-23]	18.82 (1.42) [15-21]
Total attitude subscale scores	16.43 (2.60) [12-20]	16.25 (2.43) [9-22]
Total competence subscale scores	15.74 (2.24) [11-20]	16.57 (2.08) [11-20]

To answer the third research question: “Do the EFL learners change their self-perception as readers after a four-week summarization training session (as measured by the reader self-perception test)?”, a hypothesis testing of the mean score differences between the EFL learners’ pretest scores and posttest scores of the reading self-perception survey was conducted. Again the researcher employed a repeated-measure dependent t-test using SPSS to see if there was a statistically significant mean score differences between the pretest and posttest scores of the reading self-perception survey. The results, as can be seen in Table 6, revealed that the mean differences between the posttest and the pretest scores of the reading self-perception survey as a whole, the progress subscale and attitude subscales were not significant ($p < .964$, $p < .855$, $p < .670$ respectively). The mean differences between the posttest and the pretest scores of the progress subscale and the attitude subscales were significant ($p < .046$, $p < .016$ respectively). The results implied that the EFL learners as a whole did not improve their self-perception as readers after participating in the four-week summarization training session. They did make some improvement in terms of the internal feelings they experienced during reading (i.e., their physiological states) and in terms of their beliefs regarding their ability and proficiency in reading in the four-week summarization training session.

Table 6. *Dependent t-test: hypothesis testing of the mean differences of the pretest and posttest total scores and subscale scores of the Reading Self-Perception Survey*

	Mean difference (posttest- pretest)	Std. Error	t-values	df	Sig.
Total reading self-perception scores	-0.05	0.955	-0.045	34	.964
Physiological States subscale scores	-0.77	0.373	-2.068*	34	.046
Progress subscale scores	0.08	0.427	0.184	34	.855
Attitude subscale scores	-0.18	0.415	-0.430	34	.670
Competence subscale scores	0.83	0.326	2.542*	34	.016

* The mean difference was significant at the .05 level.

To answer the fourth research question: “How satisfied are the EFL learners with the use of the WriteToLearn program?”, EFL learners’ responses to the student satisfaction questionnaire were analyzed. Table 7 shows the distribution of students’ responses to the two 5-point likert scale items. Approximately 71% (i.e., 24 out of 34) of the EFL learners chose “agree” or “strongly agree” as their responses to the first 5-point likert scale item and 68% (i.e., 23 out of 34) of the EFL learners chose “agree” or “strongly agree” as their responses to the second 5-point likert scale item. The results indicated that most of the EFL learners felt satisfied with WriteToLearn in terms of its effectiveness of promoting reading comprehension and summary writing skills.

Table 7. *Number (and percentage) of EFL learners’ rating responses to the 5-point likert-scale items in the Student Satisfaction Questionnaire (n= 34)*

	1	2	3	4	5
1. I think the WriteToLearn program implemented in my class effectively improved my reading ability of expository texts.	0 (0.0%)	4 (11.8%)	6 (17.6%)	20 (58.8%)	4 (11.8%)
2. I think the WriteToLearn program implemented in my class effectively improved my summarization skills of expository texts.	1 (2.9%)	1 (2.9%)	9 (26.5%)	17 (50.0%)	6 (17.6%)

Note: One student was absent the day the questionnaire was distributed.

The written responses to the first open-ended question confirmed the EFL learners’ positive attitude toward the implementation of WriteToLearn. Students reported that the on-line learning tool has helped them to stay focused because they wanted to find the main ideas of the passages while reading. Although the summary writing tasks were very challenging in the beginning, they found the editing tools including checks on copying, relevance and redundancy very useful in terms of refining their summary drafts. The automated feedback was interesting and they were motivated to improve their summary by moving the bar from the red or yellow zone to the green zone. To them, the summary writing activity was somewhat like playing a video game. The EFL learners also reported that they

gained substantial knowledge and learned new words from reading the passages. The dictionary tool provided by WriteToLearn saved them much time from looking up words in a paper-print dictionary. The EFL learners also felt that they wrote better summaries with the system's guidance than without it and felt a great sense of achievement after completing the summary writing tasks. Lastly, they appreciated the teacher's comments, which provided encouragement and helped them refine their summaries. The EFL learners' written responses to the second open-ended question showed their complaints about the use of WriteToLearn. The greatest complaint was the amount of time they spent on reading each passage and writing a summary. They were frustrated when they encountered too many new words in a section. They also complained that WriteToLearn did not provide enough grammatical feedback. They did not know whether all the sentences they wrote were grammatically correct. They also had difficulty using their own words to restate the main ideas of the passages. In addition, they had their doubts on how powerful the system is in terms of judging the quality of their summary writing. The EFL learners' written responses to the third open-ended question revealed that they needed more help from the teacher to effectively improve their writing skills during the implementation of WriteToLearn. They complained that the passages provided by WriteToLearn were not very appealing to them. They expressed their interests in reading things related to fashion and entertainment, which were not included in WriteToLearn.

CONCLUSION

In sum, the four-week summary writing training provided by WriteToLearn helped the EFL learners improve their summary writing skills in terms of content coverage but the desired effect of WriteToLearn on the EFL learners' improvement in other aspects of summary writing was not yet significant. In regard to their reading comprehension development, the EFL learners did improve after the summarization training, both in macrostructure and microstructure level. WriteToLearn did not have great impact on the EFL learners' self-perception as readers. However, the EFL learners did make some improvement in their internal feelings toward reading (i.e., their physiological states) and in their belief regarding their ability and proficiency in reading. The EFL learners' numerical responses to the satisfaction survey showed their positive attitude toward WriteToLearn in terms of its effectiveness of promoting reading comprehension and summary writing skills. Moreover, their written responses showed they were motivated to improve their summary writing skills and to learn new knowledge and new words from reading the passages. The researcher believes that we can see a considerably positive effect of WriteToLearn on the EFL learners' summarization skills, reading comprehension and self-perception as readers if the learners continue to use WriteToLearn for a longer period of time (e.g., 3 months). The EFL learners' complaints about the insufficient grammatical feedback from WriteToLearn, their lack of great interests in the passages and their doubts on the judgment of WriteToLearn were something the program developers of WriteToLearn should take into consideration when they are to improve the service by WriteToLearn in the future.

REFERENCES

- Koda, K. (2004). *Insights into second language reading: A cross-linguistic approach*. Cambridge: University Press.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: University Press.
- Caccamise, D., & Synder, L. (2005). Theory and pedagogical practices of text comprehension. *Topics in Language Disorders*, 25(1), 5-20.
- Chapman, J., & Tummer, W. (1995). Development of young children's reading self-concepts: An examination of emerging subcomponents and their relationship with reading

- achievement. *Journal of Educational Psychology*, 87(1), 154-167.
- Henk, W., & Melnick, S. (1995). The reader self-perception scale (RSPS): A new tool for measuring how children feel about themselves as readers. *The Reading Teacher*, 48(6), 470-482.
- Jitendra, A., Cole, C., Hoppes, M. & Wilson, B. (1998). Effects of a direct instruction main idea summarization program and self-monitoring on reading comprehension of middle school students with learning disabilities. *Reading and Writing Quarterly*, 14(4), 379-396.
- Jitendra, A., Hoppes, M. & Xin, Y. (2000). Enhancing main idea comprehension for students with learning problems: The role of a summarization strategy and self-monitoring instruction. *The Journal of Special Education*, 34(3), 127-139.
- Landauer, T., & Dumis, S. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104(2), 211-240.
- Rogevich, M. & Perin, D. (2008). Effects on science summarization of a reading comprehension intervention for adolescents with behavior and attention disorders. *Exceptional Children*, 74(2), 135-154.
- Wade-Stein, E. (2004). Summary street: Interactive computer support for writing. *Cognition and Instruction*, 22(3), 333-362.

APPENDIX A: Sample Items from the Reading Self-Perception Survey

Subscales	Sample items
Physiological Subscale	3. I enjoy reading in general.
Progress Subscale	12. I read faster than I could before.
Attitude Subscale	13. I like to read as much as I can.
Competence Subscale	6. I can figure out the meanings of hard words in a story.

APPENDIX B: Student Satisfaction Questionnaire

Please give a rating to Questions 1 and 2 by circling one of the five numbers on the scale beside it. 1 = strongly disagree, 2 = disagree, 3 = not sure, 4 = agree, 5 = strongly agree
1. I think the WriteToLearn program implemented in my class effectively improved my reading ability of expository texts. 1 2 3 4 5
2. I think the WriteToLearn program implemented in my class effectively improved my summarization skills of expository texts. 1 2 3 4 5
Please say as much as you want to answer Questions 3, 4, and 5.
3. What did you like most about the WriteToLearn program implemented in your class (most useful/ interesting)?
4. What did you find most troublesome about the WriteToLearn program implemented in your class (most troublesome/ difficult)?
5. What do you suggest your teacher do to facilitate your learning using the WriteToLearn program?

The Effects of Interactive Reading-Writing Journal in EFL College Writing

Chu-yao Chiu (邱筑瑤)

National Kaohsiung Normal University
h9612101@stu96.nknu.edu.tw

Journal writing has become an increasingly effective tool used to improve L2 learners' writing proficiency and fluency. Most previous studies investigated the teacher-student dialogue or reflective journals. Little research has tried to look at the effects of language learning journals or peer-to-peer journals. This study therefore examined the interactive reading-writing journals on EFL college students' writing. The subjects were twenty-nine freshmen from a national university in southern Taiwan. In the project, the subjects were asked to write four pieces of journal writings related to the topics of the four reading selections. After finishing each piece of journal writing, students had to exchange their writing with peers, who played the role as the readers for classmates' writing.

Four major findings were shown from the study. First, the interactive reading-writing journal could be an effective technique in improving students' writing proficiency from the pre-study and post-study writing tests. Second, the interactive reading-writing journal project also affected the students' writing fluency. The subjects' total writing words in the pre-study writing test and the post-study writing test revealed a significant difference. Third, a great majority of students reacted positively to the interactive reading-writing journals. They found that English reading provided them with ideas for English writing, facilitated their English writing processes, and helped expand their use of English vocabulary and syntactic usage. Finally, most students appreciated the peer interaction in the project.

INTRODUCTION

Reading and writing are two sides of a coin, which are interdependent and reinforce each other. Reading provides inputs and contexts for writers, while writing activates readers' prior knowledge and organizational abilities. Combining reading and writing has been shown to prompt more valuable thinking for learners (Lapp, Flood & Farnan, 1989). Rubin (1992) further indicated that integrating reading and writing could develop better reading and writing skills. However, connecting reading and writing in EFL instruction in Taiwan is far. For many EFL students in Taiwan, reading is regarded as an accumulation of new vocabularies, and writing as an uninteresting task of sentences making. Therefore, it is of critical need for EFL teachers to explore how to connect reading and writing in order to facilitate EFL learning.

Journal writing is a written log of one's thoughts, feelings, reactions, assessments, ideas, or progress, which focuses on learning process and language development (Brown, 2004). Through journal writing, students can express their ideas freely and communicate directly with their readers. In particular, journal writing is regarded as one effective strategy for developing the reading and writing skills of learners (Montgomery, 2001). In light of the use of journal writing, the current study aimed to explore the effects of interactive reading-writing journals with peer-to-peer interaction.

LITERATURE REVIEW

Theoretically and empirically reviewed, interactive reading-writing journal writing could be an effective technique used to facilitate EFL learners' language performance. From the theoretical perspectives, several learning principles could support the notion. First, in

schema theory, Noyce and Christie (1989) claimed that reading comprehension usually occurred in two conditions. One condition is that readers assimilate texts into an existing schema and the other condition is that the readers alter their schema to accommodate new information from the text. Specifically, writers can shape their writing from the input of reading and also apply their reading schemata to their writing. Therefore, reading-writing connection is positively encouraged. Second, journal writing is a process-oriented approach and is widely used in both L1 and L2 contexts as a means to improve writers' fluency and further reduce their writing anxiety (Hyland, 2002). By using process-oriented learning, students will easily involve in the learning process and are seen as whole persons in learning contexts. Weigle (2002) encourages teachers of writing to consider the process in writing instruction since it will help students become more involved and enjoy the learning. Third, scaffolding is acknowledged as supportive and effective in enhancing the beginning learner's progress to a higher level of language development (Vygotsky, 1978). In language instruction, journal writing is employed as a productive pedagogy because the regular writing behavior and interaction between readers are beneficial for the construction of scaffolding. Journal writing offers students more opportunities for experience the target language, as a result, it would improve the students' writing fluency and increase their motivation to write.

In exploring the empirical studies of journal writing, many researchers have investigated teachers' (Nassaji & Cumming, 2000; O'Connell & Dymont, 2006) and students' perceptions (Bocarro, 2003; Harris, 2008; Myers, 2001; Nassaji & Cumming, 2000) towards the implementation of teacher-student dialogue or reflective journals, which aimed to understand learners' learning progress and affective states. In spite of the efforts on the topic of journal writings, some issues are still unanswered. First, few studies tried to investigate alternative forms of journal writings to facilitate EFL language learning. It is obviously assumed, nevertheless, various forms of journal writing have been applied for educational practice. For example, Montgomery (2001) indicated combing reading and writing in journals not only gave important insights to teachers about students' learning, but also advanced learners' reading and writing skills since it is an authentic means of connecting reading and writing in language learning. According to Brown (2004), the results of journal writings could be evaluated from a number overlapping categories such as language-learning logs, grammar journals, responses to readings, strategies-based learning logs, self-assessment reflections etc. Thus, conducting more studies on alternative forms of journal writing, other than dialogue journals or reflective journals, is necessary. Second, while most studies reported the implementation of teacher-student dialogue journals, surprisingly little research has been conducted to explore peer-to-peer journal writing. In spite of the promising benefits of teacher-student dialogue journals, teachers may not provide clear and detailed feedback to students' journals due to the big-size classroom settings and time limitation. Thus, some scholars proposed peer-oriented journal exchange for sharing and discussing responses with peers. For instance, Leu and Kinzer (1999) called "buddy journals" as an effective literacy instruction, which invite pairs of students to write back and forth and maintain a written exchange about topics.

Since journal writings have been found to be of positive values for learners, researchers were encouraged to extend the benefits of journal writings to further improve learners' language skills. Concerning reading-writing into journals with peer interaction may be an effective way to advance students' language performance. By means of the peer-to-peer reading-writing journals, students (writers) not only directly communicate with the reading texts, but also share their comments with their peers (readers). It is hoped that the study will provide language teachers with an alternative activity as well as a reference to design some effective writing instructions. It is also hoped that the study will offer students with less-threatening learning environments in a community discourse to improve their writing

motivation.

Drawing on this line of reasons mentioned above, the current study, therefore, explored the effectiveness of interactive reading-writing journal (IRWJ)¹⁷ on EFL university students' writing. It examines the students' writing performance, writing fluency and attitudes before and after the project. The following research questions are addressed in the study.

1. Is there any significant difference in the students' English writing proficiency in terms of content, organization, grammar, diction, and mechanics before and after the study?
2. Is there any significant difference in the students' length of writing before and after the study?
3. Are there any significant differences in the students' responses to reading, writing, and reading-writing connection before and after the study?
4. What are the students' responses to peer interaction in the study?

METHODOLOGY

Subjects

The participants of the study were 29 freshmen from a national university of education in southern Taiwan. Altogether there were 22 males and 7 females. These students were from the course of *Freshmen English*, which meet two hours a week. All the subjects in this study were at a similar level of proficiency since they were placed into low-intermediate level according to their English scores in the Joint College Entrance Exam (JCEE). They were all non-English major students and have learned English for at least six years. Among these subjects, 22 students had ever received English writing instruction in senior high school. The other rest of 7 students did not receive any writing instruction before. In other words, nearly 80% of the students learned how to write in English before the study.

Instruments

In order to investigate the effectiveness of interactive reading-writing journal on English writing, the following instruments were employed in this study – the English reading selections, the peer evaluation checklist, the students' journals, the English writing tests, and the pre- and post-study questionnaires.

Reading Selections

Four English reading selections were adopted in this study, including “The Reasons of Getting Tattoos”, “The Weather in Chicago and Miami”, “Describing a Person”, and “Describing My Past Experiences”. The four English reading selections were all from the two writing textbooks: *Refining Composition Skills: Rhetoric and Grammar* (Smalley, Ruetten, & Kozyrev, 2001) and *From Great Paragraphs to Great Essays* (Folse, Solomon, & Clabeaux, 2007). In particular, the four reading selections represented four different writing genres. The four genres are cause-effect paragraph, comparison paragraph, descriptive paragraph, and narrative paragraph. The four genres aimed to investigate to what extent the subjects apply their acquisition in reading to their performances in writing. Moreover, the choice of the four different genres was to avoid being partial to a specific field of writing. In this way, the knowledge and skills the subjects learn in their writing could be more holistic, which could further create higher reliability in analyzing the data.

The selection of the four English articles in this study is based on the three considerations. First, the four English articles are related to the topics that the subjects are about to write, therefore, the subjects could apply the English reading selections to their English compositions. Second, the four reading selections were ordered from easy to difficult, which could effectively help the subjects develop better reading and writing skills. Third, the

¹⁷ In this paper, the acronym of IRWJ represents interactive reading-writing journal and both will be used interchangeably.

readability of these four reading selections was on the subjects' level according to Fry's formula (1991), which was suitable for the subjects in university in Taiwan. The words, phrases and sentences structures used in these four reading selections were readable and correspond with those the subjects learned in their textbooks.

Students' Journals

In this study, the subjects were asked to write four pieces of journal writings related to the topics of the four reading selections. The genres and the topics of the students' journals were shown in Table 1. The genres in the journals topics were identical to those in the reading selections. The four topics of journals were chosen in terms of strong correlation with the genres of the reading selections, which could serve as an effective stimulus to students' idea generation and writing guidance.

Peer Evaluation Checklist

The peer evaluation checklist was made up of two parts. The first part was on the discourse level, which contained 9 items of features regarding the three elements of a good paragraph, namely, the topic sentence, the supporting sentences, and the concluding sentence. The 9 checking items of the discourse level were adopted from the following books: Huang and Tzen's (2001) *English Writing Practice 1*, Lin's (2000) *English Composition for Senior High School*, Mullaney, Letkeman and Nelson's (2001) *Writing in Action*, and Folse, Solomon, & Clabeaux's (2007) *Great Paragraph to Great Essay*. In the first part of peer evaluation checklist, the student reader would make a check next to an item when the specified feature was found in their peers' writing. The second part of the evaluation checklist was on the grammar level, which covers five types of grammatical errors: fragments, run-on sentences, lack of verb, disagreements between the subject and the verb, and articles. These five types of grammatical errors were chosen because they are among the most common errors in students' writing tasks according to the researcher's teaching experiences. In the second part of peer evaluation checklist, students were told to underline these five types of errors and encode them with a symbol in their peers' writing. Students were also requested to give a frequency count of each type of errors and mark the frequency on the checklist.

The students were offered to complete the peer evaluation checklist in Chinese version. As the subjects were non-English major students and not advanced language learners, reading the statements in English might be a little difficult for them. Using their native language, Mandarin, would be much easier for the subjects to express their comments on their peers' writings.

Writing Tests

In this study, all the subjects had to complete two writing tests in class, a pre-study and a post-study writing test. The subjects had to write each writing task within one class period, 50 minutes. They were not allowed to use dictionaries for their writing.

Before writing the pre-study test, all subjects would receive the explicit writing instruction. The writing instruction in this study was to focus on paragraph writing. Since the subjects' language proficiency was not very high, paragraph writing would be a comprehensible input for their learning. After receiving explicit English paragraph writing instruction, the subjects had to write a pre-study writing test of the topic "Are you for or against making friends online?". After the study, the subjects would write a post-study writing test. In this study, the topic in post-study writing test was identical with that of the pre-study writing test. The same topics were used in order to measure the improvement or differences before and after the implementation of the interactive reading-writing journal.

In this study, the subjects' writings were scored based on Joint College Entrance Exam (JCEE) analytic scoring for the following two reasons. First, the JCEE writing scoring would be a standardized and appropriate scoring system to measure students' English writing proficiency. Second, analytic scoring was preferred in this study. Brown (2004) indicated

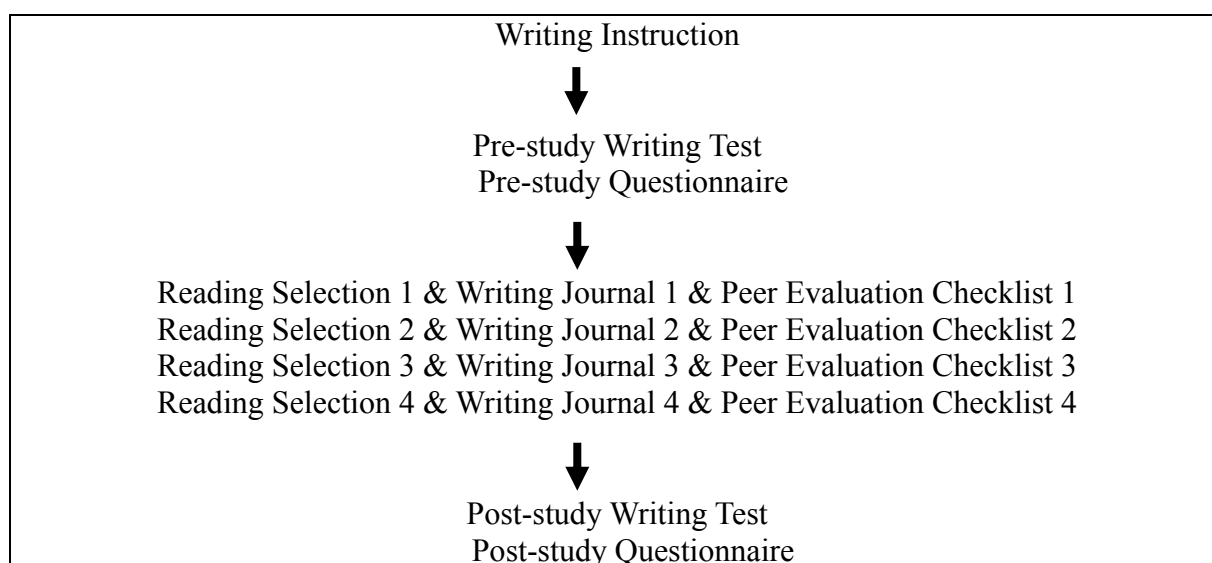
classroom evaluation of learning was best through analytic scoring since holistic scoring provided little washback for writers. Analytic scoring used in this study could function as a diagnostic purpose to offer the teacher further information about the students' strengths and weaknesses. The scoring used in the study included five analytic parts: content (25%), organization (25%), grammar (20%), diction (20%) and mechanics (10%), and the total score was 100.

Table 1. *Four Writing Genres and Four Writing Topics in the Study*

	Genres	Reading Selections	Writing Topics
1	Cause-effect	The Reasons of Getting Tattoos	The Effects of Improved Technology on Life
2	Comparison	The Weather in Chicago and Miami	The Life in High School and University
3	Descriptive	Describing a Person	Describing Your Class Teacher
4	Narrative	Describing My Past Experiences	Describing Your Experiences: Seeing a Doctor

Pre- and Post-study Questionnaires

There were two kinds of questionnaires used in this study: a pre-study and a post-study questionnaire. Before the study, the subjects had to fill in a pre-study questionnaire for a better understanding about their attitude toward interactive reading-writing journal. In the pre-study questionnaire, the researcher designed 15 items to explore the subjects' attitudes towards reading, writing, and reading-writing connection before the study. After the study, the students would fill in a post-study questionnaire to express their feelings about interactive reading-writing journal. In the post-study questionnaire, the researcher included the 15 identical items in the pre-study questionnaire to see the subjects' improvement and differences towards reading, writing, and reading-writing connection before and after the study. Additional 4 items towards peer interaction were covered in the post-study questionnaire. All the statements in both the pre- and post-study questionnaires were attitudinal in nature accompanied by a Likert-type scale ranging from *5 strongly agree* to *1 strongly disagree*. Also, the subjects were offered to complete the two questionnaires in Chinese version. The whole procedure was sequenced as follows (see Figure. 1).



RESULTS AND DISCUSSION

Results of Writing Proficiency

According to Table 2, the total mean scores and five analytic scores in the post-study writing were better than those in the pre-study writing test. In particular, significant differences were found in the subjects' total scores, and the analytic criteria of content, organization. However, there were no significant differences in the other three criteria of grammar, diction, and mechanics. The results, therefore, indicated that the interactive reading-writing journal helped enhance the subjects' English writing proficiency. The results confirmed Connor-Greene's (2000) and Hyers' (2001) notions that students made progress in their English writing performance through the implementation of interactive journal writing. Thus, it could be concluded that interactive reading-writing journal project promoted students' writing proficiency. Moreover, the students' writing proficiency improved particularly in the aspects of content, and organization after the interactive reading-writing journal project.

Table 2. *T-test Results of Students' Writing Proficiency Before and After the IRWJ*

		N	Mean	Mean Difference	SD	t	Sig. (2-tailed)
Content	Pre-study	29	16.86	1.172	3.032	-3.779	.001*
	Post-study	29	18.03				
Organization	Pre-study	29	17.14	1.828	3.102	-5.257	.000*
	Post-study	29	18.97				
Grammar	Pre-study	29	14.24	.379	1.806	-1.652	.110
	Post-study	29	14.62				
Diction	Pre-study	29	13.97	.172	1.239	-.542	.592
	Post-study	29	14.14				
Mechanics	Pre-study	29	8.83	.103	.539	-1.797	.083
	Post-study	29	8.93				
Total Writing Scores	Pre-study	29	71.03	3.76	7.794	-4.603	.000*
	Post-study	29	74.79				

Note. The asterisk * indicated the mean differences were significant at the .05 level.

Results of Writing Fluency

In order to investigate the effect of interactive reading-writing journal on the students' writing fluency, the subjects' length of pre- and post-study writing tests was compared. As presented in Table 3, the subjects' mean length of word numbers in the post-study writing test exceeded those in the pre-study writing test ($M = 101.45 > 97.93$). With the significant difference of the length between the pre- and post-study writing tests, it could be inferred that the interactive reading-writing journal had positive influence on students' writing fluency. The results were consistent with previous studies announcing that journal writings could benefit students in their development of writing fluency (Holmes & Moulton, 1997; Liao & Wong, 2008).

Table 3. *T-test Results of Students' Writing Fluency Before and After the IRWJ*

		N	Mean	Mean Difference	SD	t	Sig. (2-tailed)
Length of Writing	Pre-study	29	97.93	3.517	21.063	-3.335	.002*
	Post-study	29	101.45		20.094		

Note. The asterisk * indicated the mean differences were significant at the .05 level.

Results of Students' Attitudes towards Reading

According to Table 4, there were three significant differences in the students' responses to Item 1, Item 2, and Item 3. The results indicated that most subjects would like reading English articles, and take reading as a way to communicate with the authors. It implied that the interactive reading-writing helped bridge the gap between readers and writers. The study findings support Clouse's (1994) notion that readers carried on a conversation with the writers when they read in order to comprehend the reading articles better. Thus, the implementation of the interactive reading-writing journal helped promote the interaction between readers and writers. However, no significant differences were found in Item 4 and 5. It would indicate that students did not turn to their classmates for better understanding of the reading articles. Moreover, the students did not like to write down their ideas while reading English articles.

Table 4. *T-test Results of Students' Attitudes toward Reading Before and After the IRWJ (N=29)*

Item		Mean	SD	t	Sig. (2-tailed)
1. I like reading English articles.	Pre-study	2.45	.632	-6.769	.000*
	Post-study	3.28	.591		
2. I take English reading as a way to communicate with the author.	Pre-study	2.69	.806	-5.270	.000*
	Post-study	3.34	.670		
3. I will find English outside reading materials to enrich my knowledge.	Pre-study	2.38	.775	-2.368	.025*
	Post-study	2.76	.786		
4. I will discuss with my peers to know the main ideas of the article.	Pre-study	3.07	.998	-1.095	.283
	Post-study	3.24	.872		
5. In English reading, I like to write down my thoughts in English.	Pre-study	2.48	.785	-1.316	.199
	Post-study	2.72	.797		

Note. The asterisk * indicated the mean differences were significant at the .05 level.

Results of Students' Attitudes towards Writing

In order to further understand the students' attitudes towards writing before and after the study, Table 5 showed the statistical results in the pre- and post-study questionnaires. According to Table 5, significant differences existed among those five paired-items. The results indicated that students liked English writing more and enjoyed sharing their ideas with their peers in the process writing. Thus, the interactive reading-writing journal not only improved the students' writing habit, but also motivated the students' motivation for English writing.

Table 5. *T-test Results of Students' Attitudes toward Writing Before and After the IRWJ (N=29)*

Item		Mean	SD	t	Sig. (2-tailed)
6. I like to write down my ideas in English.	Pre-study	2.38	.728	-6.298	.000*
	Post-study	2.97	.778		
7. I take English writing as a way to communicate with the readers.	Pre-study	2.55	.783	-2.768	.010*
	Post-study	2.86	.789		
8. I will discuss with my peers to make clear what to write about.	Pre-study	3.24	.951	-4.137	.000*
	Post-study	4.00	.598		
9. In English writing, I like to share my writing with my classmates.	Pre-study	2.90	.860	-6.699	.000*
	Post-study	3.76	.786		
10. I will write down my thoughts in English.	Pre-study	2.07	.884	-6.009	.000*
	Post-study	2.79	.620		

Note. The asterisk * indicated the mean differences were significant at the .05 level.

Results of Students' Attitudes towards Reading-writing Connection

Table 6 further showed the results of students' attitudes towards reading-writing connection after the study. As shown in Table 6, significant differences existed among those five paired-items in terms of reading-writing connection. The interactive reading-writing journal could stimulate the students' writing motivation. Moreover, students could apply the new phrases or sentences patterns used in the reading selections into their own writing tasks. In particular, most students agreed that writing after reading could make their writing process easier and more organized. Those results mentioned above showed that the interactive reading-writing journal could be regarded as an effective and systematic way of facilitating students' writing through reading.

Table 6. *T-test Results of Students' Attitudes toward Reading-Writing Connection Before and After the IRWJ (N=29)*

Item		Mean	SD	t	Sig. (2-tailed)
11. Reading articles can inspire my motivation in writing.	Pre-study	2.76	.786	-6.250	.000*
	Post-study	3.62	.775		
12. My English writing ability improved when reading a relevant article before writing.	Pre-study	3.17	.928	-5.073	.000*
	Post-study	4.03	.499		
13. I can apply the new English words, phrases, and sentence patterns in the reading selections to my English writing.	Pre-study	3.38	.862	-3.134	.004*
	Post-study	3.93	.593		
14. My English writing became easier after reading articles.	Pre-study	3.31	.806	-3.415	.002*
	Post-study	3.93	.458		

15. Writing after reading makes the writing process more organized.	Pre-study	3.21	.940	-5.525	.000*
	Post-study	3.97	.626		

Note. The asterisk * indicated the mean differences were significant at the .05 level.

Results of Students' Responses to Peer Interaction

Peer interaction played a vital role in the study. During the interactive reading-writing journal project, students kept interacting with peers by means of journal writing exchange, in order to give their peers' useful feedback for better English writing performance. Table 7 showed that more than half students took positive attitudes towards peer interaction after the study. Those results implied that the interactive reading-writing journal project served as an effective means to help increase peer interaction. On the whole, the students' positive responses to peer interaction in the interactive reading-writing journal manifested that the journal writing succeeded in facilitating the students' English reading and writing by offering them opportunities to engage in peer interaction. As well as in English learning, the students developed friendships and promote their communication skills with peers. These study findings confirmed the notion of social interaction, which was recognized as prerequisite for learning writing by Vygotsky (1978). Fathman and Kessler (1993) addressed that cooperative peer interaction provided a learning social context where students use the target in a meaningful way. De Guerreto & Villamil (2000) further promoted the importance of peer interaction in L2 learning. Hyland (2003) also pointed out that giving feedback also helped the writers to comprehend the writing context, providing a sense of audience and an understanding of the expectations of the communities they were writing for. On account of the critical importance of peer interaction in language learning, teachers should try their best to construct a learning environment involving efficient peer interaction.

Table 7. *Frequencies and Percentage of the Students' Responses to Peer Interaction in the IRWJ*

Item	SD	D	N	A	SA	Total
16. My classmates' journal writing can enhance my inspiration and interest in my writing.	0	3	7	17	2	29
	0%	10%	24%	59%	7%	100%
17. Exchanging the journals with my classmates makes it easier for me to write.	0	3	8	14	4	29
	0%	10%	28%	48%	14%	100%
18. Exchanging the journals with my classmates helps shorten my time on writing.	0	1	5	20	3	29
	0%	3%	17%	69%	10%	100%
19. Suggestions from my classmates can help solve my writing problems.	0	0	13	14	2	29
	0%	0%	45%	48%	7%	100%

Note. SD = Strongly Disagree; D = Disagree; N = Neutral; A = Agree; SA = Strongly Agree.

CONCLUSION

Four major conclusions can be drawn from the findings of the study. First, interactive reading-writing journal could be an effective and useful technique in improving students' writing proficiency. Second, the interactive reading-writing journal project also affected the students' writing fluency. Third, a great majority of students reacted positively to the interactive reading-writing journal project after they study. Finally, most students appreciated the peer interaction in the project.

On the basis on the study findings, several pedagogical implications would be provided. First, EFL teachers can improve students' English writing proficiency as well as writing fluency through interactive reading-writing journal exchange. As shown in the study, students expressed that the reading selections and journal exchange had positive effects on their writing performance. Reading selections, here, served as role models for students to learn the essentials of writing, such as the development of ideas, organization, mechanics in English writing. On the other hand, journal exchange provided students with authentic readers to give useful feedback and suggestions on their writing. Second, EFL teachers can take advantage of peer interaction to promote students' English learning efficacy and lower their learning anxiety. In the study, through interaction with classmates, the students figured out the gist and gained inspiration for their peers' English writing. Furthermore, peer interaction facilitated the students' English writing processes and shortened their English writing time. With the help and support from peers, students can feel less burdened and less resistant in English writing. Third, from the results of the study, reading-writing journal project indeed improves students' writing abilities and leads the students into more positive attitudes toward reading, writing and reading-writing connection. With the incorporation of journal writing into writing instruction, teachers will feel satisfied as more time is saved and students will find writing instruction to be more fun than work.

REFERENCES

- Bocarro, J. (2003). Maintaining the balance between service and learning: The use of journals in promoting critical thinking. *Scholar: A Journal of Leisure Studies and Recreation Education*, 18, 7-22.
- Brown, H. D. (2004). *Language Assessment Principles and Classroom Practices*. NY: Longman.
- Connor-Greene, P. A. (2000). Making connections: evaluating the effectiveness of journal writing in enhancing student learning. *Teaching of Psychology*, 27(1), 44-46.
- Clouse, B. M. (1994). *Transitions: From reading to writing*. New York: McGraw-Hill, Inc.
- De Cuerrero, M. C. M., & Villamil, O. S. (2000). Activating the ZPD: Mutual scaffolding in L2 peer revision. *Modern Language Journal*, 84, 51-68.
- Fathman, A. K. & Kessler, C. (1993). Cooperative language learning in school context. *Annual Review of Applied Linguistics*, 13, 127-140.
- Folse, K. S., Solomon, E. V. & Clabeaux, D. (2007). *From Great Paragraphs to Great Essays*. Boston: Houghton Mifflin.
- Fry, E. (1991). *Ten best ideas for reading teachers*. New York: Adios-Wesley.
- Harris, M. (2008). Scaffolding reflective journal writing – Negotiating power, play and position. *Nurse Education Today*, 28, 314-326.
- Holmes, V. L., & Moulton, M. R. (1997). Dialogue journals as an ESL learning strategy. *Journal of Adolescent & Adult Literacy*, 40(8), 616-621.
- Hyers, A. D. (2001). Predictable achievement patterns for student journals in introductory earth science courses. *Journal of Geography in Higher Education*, 25(1), 53-56.
- Hyland, K. (2002). *Teaching and Researching Writing*. Harlow: Pearson Education.
- Hyland, K. (2003). *Second Language Writing*. Cambridge: Cambridge University Press.

- Lapp, D., Flood, J., & Farnan, N. (1989). *Content area reading and learning: Instructional strategies*. Englewood, NJ: Prentice-Hall.
- Leu, D. & Kinzer, C. K. (1999). *Effective literacy instruction, K-8*. NJ: Merrill.
- Liao, M. T., & Wong C. T. (2008). Effects of dialogue journals on students' writing fluency, motivations, anxiety, and reflection. *Proceedings of the 17th International Symposium and Book Fair on English Teaching*, pp. 144-156. Taipei: Crane.
- Montgomery, W. (2001). Journal writing: Connecting reading and writing in mainstream educational settings. *Reading and Writing Quarterly*, 17, 93-98.
- Mullaney, M., Letkeman T., & Nelson S. A. (2001). *Writing in action – A style guide for writing*. Taipei: AMC publishers.
- Myers, J. L. (2001). Self-evaluation of the “stream of thought” in journal writing. *System*, 29, 481-488
- Nassaji, H. & Cumming, A. (2000). What's in a ZPD? A case study of a young ESL student and teacher interacting through dialogue journals. *Language Teaching Research*, 4(2), 95-121.
- Noyce, R. M., & Christie, J. (1989). *Integrating reading and writing instruction in grades K-8*. Boston: Allyn & Bacon.
- O'Connell, T. & Dymont, J. (2006). Reflections on using journals in higher education: A focus group discussion with faculty. *Assessment and Evaluation in Higher Education*, 31(6), 671-691.
- Rubin, D. (1992). *Teaching reading and study skills in content areas*. Massachusetts: Allyn and Bacor, a Division of Simon & Schuster, Inc.
- Smalley, R. L., Ruetten, M. K. & Kozyrev, J. R. (2001). *Refining Composition Skills: Rhetoric and Grammar*. Boston: Heinle & Heinle.
- Weigle, S. C. (2002). *Assessing Writing*. Cambridge: Cambridge University Press.
- Vygotsky, L. S. (1978). *Mind in society: the development of higher psychological processes*. Cambridge, Mass.: Harvard University Press.
- 林茂松 (2000)。遠東高中英文作文。遠東圖書公司。
- 黃素月，曾瑾玲 (2001)。英文寫作練習一。三民書局。

Effects of Reciprocal Teaching on Reading Comprehension in an EFL Classroom

Shiao-ting Chou (周筱婷), Yu-ling You (游毓玲)

National Changua University of Education

cathyhtchou@yahoo.com.tw
youyl@cc.ncue.edu.tw

The purpose of this study is to investigate if reciprocal teaching can assist EFL junior high school students to improve their reading comprehension, and to probe the students' response to reciprocal teaching. The subjects were 70 ninth-grade junior high school students from two mixed-level classes, one class is the experimental group and the other the control group. A ten-week reciprocal teaching instruction, which emphasized teacher modeling, guided practice and independent application, was conducted to the experimental group to teach four metacognitive strategies, i.e., predicting, summarizing, questioning, and clarifying. The instruments of the study included reading comprehension pre-test and post-test to assess the subjects' reading comprehension, and a response questionnaire. The findings show firstly that the experiment group gained more significant improvement in reading comprehension than the control group, and that most subjects in the experiment group were affirmative about the efficiency of reciprocal teaching on promoting and enhancing reading comprehension. The results suggest that reciprocal teaching can be an alternative for teaching reading in EFL junior high school classrooms.

INTRODUCTION

Reading is an active process of meaning making, which needs various types of knowledge and involves complex cognitive activities. Among them, effective reading strategies are essential in achieving successful reading comprehension (Johnson, 1982; Mikulecky, 1990). Herrmann (1988) also mentioned that "reading was not the mindless application of isolated skills; rather, it was a strategic, meaning-getting process, which required awareness and control of complex reasoning process" (p. 24). Additionally, in examining the reading process of a foreign language in his study, Phillips (1984) indicated that the complex process of reading comprehension might be even more difficult for EFL learners because their information processing was not developed completely and not firmly established in the their mind. Therefore, he suggested that instructing EFL learners reading strategies might be helpful in promoting their reading comprehension. However, students in Taiwan are seldom instructed with strategic behavior (Hsu, 2003). Traditional English reading instruction in Taiwan is teacher-directed. What teachers do in traditional reading classrooms is to measure comprehension instead of teaching students strategies for comprehension. While instructing, teachers focus on word and grammar explanation. Consequently, students tend to pay their attention on details instead of the main idea of their reading (Chen, 2005; Chen, 2005; Hsu, 2004; Hsu, 2003). Their poor reading competence is believed to result from insufficient employment of reading strategies (Hsu, 2003).

With regard to instructing students' reading strategies, Clark (2003) considered reciprocal teaching as an effective method in improving reading comprehension in school-aged students. The four metacognitive strategies in reciprocal teaching are the strategies of summarizing, questioning, clarifying, and predicting. In addition to instructing students the four strategies to help them build the meaning from the reading text as well as

monitor their reading process to promote their comprehension, reciprocal teaching also provides students a learning environment, where they can share their feelings and knowledge with their classmates. According to Rosenshine and Meister (1994), reciprocal teaching is to use dialogue, a kind of group discussion, as a way to practice and learn these four strategies. Helping and sharing knowledge through group discussion allow students to focus on meaning construction about the text actively while reading.

Paris and Winograd (1990) indicated that reciprocal teaching contributed to students' learning because it could enhance students' metacognition. Carter (1997) once took "a sentry" as a metaphor of metacognition, a thinking and monitoring process. She thought the "sentry" could inform all readers when understanding took place and when it did not, and reciprocal teaching could help students develop their metacognitive "sentries" (p. 66). The four strategies in reciprocal teaching can help students not only foster their reading comprehension but also monitor their reading process (Palincsar & Brown, 1984). Furthermore, group discussion (dialogue) offers students a natural context for learning. In group discussion, students share their rationale about how to use the strategies, when these strategies are helpful, and why they should use these strategies. In other words, reciprocal teaching helps students to monitor and control their own learning; as a result, reciprocal teaching can enhance learners' metacognition. Then what is metacognition?

Flavell (1979) defined metacognition as thinking about one's own thinking. It is also broadly defined as awareness and control of one's own learning (Baker, 1989; Baker & Brown, 1984; Brown, 1985; Schraw, 2001). According to Carrell (1988), as readers apply metacognition in their reading, they get aware of their cognitive resources, and then this awareness helps readers to control their own cognitive process. Namely, after getting aware of what strategies to use, as well as when, why, and how to use them appropriately and effectively, these readers can apply it to their future reading independently (Duffy, Roehler, & Herrmann, 1988). Moreover, many researches also indicated that metacognition is what makes good readers more proficient in reading (Baker & Brown, 1984; Byrd & Gholson, 1985; Clark, 2003; Garner & Kraus, 1981-1982). They stated that good readers were more able to monitor their comprehension and more aware of the strategies to use than poor readers. Good readers also used strategies more flexibly (Anderson, 1980; Block, 1986; Garner & Kraus, 1981-1982; Paris & Jacob, 1984; Paris & Myers, 1981). In a word, readers' metacognitive awareness of strategies is a critical component of effective strategy use and reading competence.

Several studies have shown the effect of reciprocal teaching in promoting L1 learners' reading comprehension, especially those who are adequate decoders but poor comprehenders (Alfassi, 2004; Al-Hilawani, Marchant, & Poteet, 1993; Bruce & Robinson, 2001; Guerlene, 2002; Palincsar & Brown, 1984). Researchers in the field have also recommended reciprocal teaching in increasing second-language learners' reading proficiency (Clark, 2003; Cotterall, 1990; O'Malley & Chamot, 1990). However, few studies were conducted in examining the effectiveness of the reciprocal teaching for EFL junior high school students, especially in Taiwan. Therefore, the purpose of the present study is to investigate the efficacy of using reciprocal teaching in promoting EFL junior high school students' reading comprehension and its applicability in a classroom setting. Additionally, students' responses on the reciprocal teaching instruction are also probed. The following research questions are addressed:

1. Is the reciprocal teaching instruction more effective in promoting reading comprehension of the ninth-grade EFL students in Taiwan than the traditional reading instruction?
2. What is EFL students' response towards the reciprocal teaching instruction in promoting English reading comprehension in Taiwan?

METHODOLOGY

This section consists of five parts. The first part describes the background of the subjects in this study. The second one illustrates the selection of instruments used in this study. The third one introduces the whole procedure. The fourth one illustrates the reciprocal teaching instruction. The last part explains the statistic method of data analysis.

Subjects

The subjects are 70 junior high school students from two different classes. They are ninth graders in a junior high school in northern Taiwan. This study conducted convenient sampling. The researcher at the same time is their English teacher. All the subjects in both classes are native Chinese who have learned English for at least four years and none of them ever lived in English-speaking countries. Before the instructions, these two classes took a pretest in the reading comprehension assessment. The means scores in these two classes were 38.0 and 38.94 respectively. According to an independent samples t-test, no significant difference was found between these two classes, $p=.817$. Thus, the English reading proficiency in these two classes was equivalent. Based on the similar learning background and English reading proficiency level, one class was assigned as the experimental group, while the other was the control group (Table 1). There were 36 students in the experimental group and 34 in the control group.

Table 1. Comparison of Subjects' Homogeneity in English Reading Proficiency

Group	Number	Pre-test Means	Independent Samples T-test P Value
Experimental group	36	38	.817
Control group	34	38.94	

* $p < .05$

To conduct the reciprocal teaching instruction, the subjects in the experimental group were formed into seven small groups according to the scores of the pretest in the reading comprehension assessment. Based on the scores of the pretest, the subjects of the experimental group were divided into five different levels, "High," "Medium-high," "Medium," "Medium-low," and "Low." Each small group included each of the five levels, so the English reading proficiency in each group was similar. As for the subjects in the control group, they received the traditional reading instruction, so there was no need for them to form into small groups.

Instruments

Two instruments were used in this study to collect data. One was a reading comprehension test; the other was a response questionnaire.

A Reading Comprehension Test. Before the instructions, a pretest of reading comprehension was conducted in both groups. The purpose of the pretest was to evaluate the reading comprehension proficiency of the subjects in both groups before the instructions. Additionally, the pretest was also used to examine the subjects' homogeneity in English reading proficiency in both groups and be the criterion of forming into small groups in the experimental group. The pretest included five articles, and each article was followed by five multiple-choice questions. There were twenty-five multiple-choice questions in total and the total score was 100. The questions of the pretest were to tap understanding of main idea as well as whether students could integrate information in the passages and solve the problems mentioned in these articles. All the articles were selected from *Fluency Kits Independent Practice* (2004) and *Reading Builder* (2007).

Table 2. *Five Articles in the Reading Comprehension Test*

Article	Length	Readability	Genre
The Talking Dogfish	157	5.8	descriptive
The Garage Sale	164	5.8	descriptive
The Fire Fighters	165	5.7	descriptive
The Experiment	177	5.0	descriptive
A Penny Saved	244	5.5	descriptive

As shown in Table 2, the five articles in the pretest were described with respect to the length, readability, and genre. The length in the Talking Dogfish was 157, which meant there were 157 words in this article. The length in these articles ranged from 157 to 244. The readability in these articles was from 5.0 to 5.8. The genre of these articles was all descriptive. Additionally, the Chinese translation of the words not included in the 2000 Essential Words in English for beginners was put in the word bank in each article except those words designed to have the subjects apply the strategy of clarifying.

After the instructions, a posttest of reading comprehension was conducted in both groups. The purpose of the posttest was to probe the effects of the reciprocal teaching instruction in the experimental group and the traditional reading instruction in the control group. The posttest shared the same items with the pretest. Thus, there was no explanation about these five articles after the pretest. The subjects in both groups were not told in advance that they would have a posttest; nor were they told that they would take the same exams. In addition, after the subjects took the pretest, the test sheets were retrieved immediately. The interval between the pretest and the posttest was twelve weeks. The scores in the pre-and posttest would be compared to evaluate the effects of these two instructions in these two groups.

A Response Questionnaire about the Reciprocal Teaching Instruction. To realize the subjects' responses toward the reciprocal teaching instruction, a response questionnaire was administered to the subjects in the experimental group. The purpose of the questionnaire was to explore their perceptions and feedbacks of the whole experiment. There were sixteen questions in the response questionnaire. These questions were based on the following four sections. They were the belief in the effect of these four strategies (questioning, summarizing, predicting, and clarifying), the belief in the effect of the group discussion in applying these four strategies, the evaluation of reciprocal teaching in promoting and monitoring comprehension, and the attitude toward the application of reciprocal teaching in the future (Table 3). Additionally, all the items in the questionnaire were constructed based on the scales of "strongly agree," "agree," "neutral," "disagree," and "strongly disagree." All the statements in the questionnaire were presented to the subjects in their first language, Mandarin Chinese.

Table 3. *The Response Questionnaire on the Reciprocal Teaching Training*

Sections	Questions
1. Subjects belief in the effect of the four strategies	1, 2, 3, 4
2. Subjects belief in the effect of group discussion in the application of the four strategies	5, 6, 7, 8
3. Subjects evaluation of reciprocal teaching in promoting and monitoring comprehension	9, 10, 11, 12, 13
4. Subjects attitude toward applying the experience of the activity in the future	14, 15, 16

With regard to the design of the questionnaire, questions numbered one, two, three, and four were related to the subjects' belief in the effect of the four strategies. Questions numbered five, six, seven, and eight were about the subjects' belief in the effect of group discussion in the application of the four strategies. Questions numbered nine, ten, eleven, twelve, and thirteen aimed to examine the evaluation of reciprocal teaching in promoting and monitoring comprehension. Questions numbered fourteen, fifteen, and sixteen concerned the attitude toward applying the experience of the activity in the future. To sum up, the four main sections were explored by the response questionnaire with the sixteen questions.

Procedure

The design of the present study included four parts. The first part was a pretest of reading comprehension in both groups. The second one involved a traditional reading instruction in the control group and a reciprocal teaching instruction in the experimental group. The third one included a posttest of reading comprehension in both groups. The last one was a response questionnaire about the reciprocal teaching instruction in the experimental group only.

The Reciprocal Teaching Instruction

All the subjects in the experimental group participated in the reciprocal teaching instruction that was incorporated into an English reading course. The instruction, extending into ten sessions, included the following three phases: direct instruction of the reading strategies, guided practice, and independent use of the reading strategies.

Direct Instruction of the Reading Strategies. Pressley (2000) and Snow (2002) indicated that explicit instruction in cognitive strategies allowed students to monitor and adjust their processing of information. For students to develop higher levels of understanding, explicit or direct instruction in metacognitive strategies that regulate self-awareness, self-control, and self-monitoring are necessary (Feldman, 2000). Therefore, for the first two sessions of the reciprocal teaching instruction, the strategies of questioning, summarizing, clarifying, and predicting were explained explicitly and directly. Afterwards, the teacher modeled how to apply these four strategies to comprehend the articles on the worksheet by thinking aloud.

Guided Practice. In scaffolding, experts model the task for the novices. Then the novices are ceded more and more responsibility for completing the reading task on their own (Pearson & Fielding, 1991). Thus, during the guided practice phase of the instruction, students gradually took increasing responsibility for completing the reading task. For the following four sessions of the reciprocal teaching instruction, the four strategies were practiced on the training articles with cued scaffolding from the teacher each week. The training articles were selected from *Reading Comprehension* (2005) and *Reading Builder* (2007). After students practiced applying each strategy in comprehending the articles, they needed to answer the questions at the end of each article. Additionally, in this phase, the subjects in the experimental group were formed into small groups for the sake of group discussion.

In the third session of the instruction, the teacher reviewed how, why, and when to use the strategy of *predicting*, and then each group practiced applying it to comprehend the first training article through group discussion. In order to help students structure the process of reciprocal teaching until students could do it on their own, each group was given a weekly worksheet. These weekly worksheets were used in guided practice phase only. After practicing the strategy of predicting, students needed to fill in their predictions on the worksheet. Afterwards, the teacher modeled how to apply the strategies of summarizing, clarifying, and questioning in comprehending the first training article by thinking aloud. Afterwards, students needed to answer the questions below the article. In the end of the session, each group shared what they wrote aloud, and then other groups gave comments or

feedback.

In the fourth session of the instruction, the teacher reviewed how, why, and when to use the strategies of *predicting* and *summarizing*, and then each group practiced applying them on the second training article. After the practice, each group needed to fill in their predictions and summaries on the weekly worksheet, and then the teacher modeled the remaining strategies by thinking aloud. In the fifth session of the instruction, the teacher reviewed how, why, and when to use the strategies of *predicting*, *summarizing*, and *clarifying*, and then each group practiced applying them on the third training article. After the practice, each group needed to fill in their application of the strategies on the weekly worksheet, and then the teacher modeled how to use the strategy of questioning to promote comprehension by thinking aloud. In the last session of the guided practice phase, the teacher reviewed how, why, and when to use *the four strategies* and observed each group who took full responsibility for the reciprocal teaching instruction. The time for teacher modeling reduced as students got more control about the guided practice.

Independent Use of the Reading Strategies. In the beginning of the phase, a group leader was decided on in each small group. The role of the group leader was to encourage discussion and call on other group members to apply the strategies of predicting, clarifying, and questioning, but the group leader needed to summarize the text. All the group members took turns being the group leader during the last four sessions.

In order for students to apply the four strategies fluently in comprehending the training articles, they were cued to read the title and the picture first. Then the leader called on one group member to predict the whole text. Next, all of the members read the first paragraph together to confirm or revise their prediction. Students could read it silently or aloud to one another in small groups. After that, the group leader summarized the first paragraph with one or two sentences and then asked if there was any comment or feedback. The other members in the same group took turns talking, addressing to each other. Next, the leader asked if anything needed to be clarified in the first paragraph. If there was, all the members could use context clues or prior knowledge to determine the meaning of the words. Last, the leader asked if anyone could pose a question, which could indicate the most important information of the first paragraph. All the other members took turns generating questions, and answering them. After applying all the four strategies on the first paragraph, the leader chose one member to predict what would happen in the second paragraph, and then all the group members read the second paragraph aloud to confirm or revise their prediction. The same format was repeated on the following paragraphs until the end of the article.

In the independent use of the reading strategies phase, the teacher circulated among the groups. Sometimes the teacher might answer questions the group asked her, or get a group or group member on task. Sometimes the teacher might provide help when the leader had a particularly difficult time controlling the group. In short, the role of the teacher was a facilitator who was to keep discussion on track and participate only when needed

Data Analysis

To analyze the first research question, the researcher applied the Statistical Package for Social Science (SPSS) to deal with the collected data. Two paired samples t-tests by courtesy of SPSS 13.0 were used to measure the scores of the pre-and posttest in reading comprehension in both groups. Descriptive statistics such as means and standard deviations were also computed. The purpose of this analysis was to measure if there was significant differences in the subjects' reading comprehension in each group individually.

As to analyze the second research question, the questionnaire about the response to the reciprocal teaching instruction would be employed to examine the perceptions of the subjects about the experiment toward this reading activity. The subjects' response to the multiple-choice questions would be displayed as frequency counts to find out how many

subjects would answer the item on the questionnaire. Descriptive statistics would be used to interpret the results from the questionnaire responses.

RESULTS

The study was designed, firstly, to evaluate the effects of the reciprocal teaching instruction in promoting the reading comprehension of the ninth-grade EFL students in Taiwan. Secondly, it was also designed to examine EFL students' response towards the reciprocal teaching instruction in promoting English reading comprehension in Taiwan. Therefore, in this section the following two results are presented. The first one is the results of reading comprehension test; the second one presents the responses toward the reciprocal teaching instruction.

Results of Reading Comprehension Test

Table 4 reported the mean scores of the experimental group and the control group in the reading comprehension test. As the statistics indicated, the experimental group's means of the pre-test and post-test scores were 38 and 53.67 respectively, which was a gain of 15.67 on the reading comprehension test. A paired samples t-test indicated a significant difference between the pre-test and post-test scores performed by thirty-six subjects in the study, $p=.000$.

As far as the control group was concerned, the means of the pre-test and post-test scores were 38.94 and 46.35 respectively, which was a gain of 7.41 on the reading comprehension test. Analysis of a paired samples t-test revealed that there was a significant difference in the progress of the control group, $p=.007$.

Table 4. *Subjects' Pre-and Posttest in Reading Comprehension in Both Groups*

Group (N)	Pre-test Means (SD) (A)	Post-test Means (SD) (B)	Difference between (A) and (B)	Paired Samples T-test P Value
Experimental group (36)	38 (16.74)	53.67 (21.26)	15.67	.000*
Control group (34)	38.94 (17.09)	46.35 (20.28)	7.41	.007*

Note. N = the number of the students and SD = standard deviations

* $p < .05$

Results of the Response Questionnaire

The results were illustrated based on the following four parts. They are subjects' belief in the effects of the four strategies, the belief in the effect of the group discussion in applying these four strategies, the evaluation of reciprocal teaching in promotion and monitoring comprehension, and the attitude toward the application of reciprocal teaching in the future.

Subjects' Belief in the Effects of the Four Strategies. From Table 5, most subjects in the experimental group responded positively on the effect of applying the four strategies in comprehending the text. With regard to the first question, thirty-three subjects (91.6 %) believed that "the strategy of clarifying is effective" whereas only one subject (2.8 %) did not think that case. As for the second question, twenty-two subjects (61.2%) believed that "the strategy of questioning is effective" in contrast to three subjects (8.5 %) who did not think so. As regards the third question, thirty-five (97.3%) of them believed that "the strategy of summarizing is effective." No one held the negative response. As for the fourth question, twenty-five (69.5 %) of them believed that "the strategy of predicting is effective" with only two subjects (5.7%) held negative attitudes. In short, most of the subjects believed the four

strategies in reciprocal teaching were effective on promoting comprehension. Among these four strategies, they considered the strategies of summarizing and clarifying the most effective. This is not completely consistent with the result of Palincsar and Brown's (1984) study. In Palincsar and Brown's study, they thought that the strategies of questioning and summarizing were more effective in that these two strategies served a comprehension-monitoring function.

Table 5. *Subjects' Belief in the Effect of the Four Strategies*

Questions	Frequencies (Percentages)		
	Positive	Neutral	Negative
	Strongly agree/ Agree		Disagree/ Strongly disagree
1. I think the strategy of clarifying is effective in promoting reading comprehension.	33 (91.6%)	2 (5.7%)	1 (2.8%)
2. I think the strategy of questioning is effective in promoting reading comprehension.	22 (61.2%)	11 (30.6%)	3 (8.5%)
3. I think the strategy of summarizing is effective in promoting reading comprehension.	35 (97.3%)	1 (2.8%)	0
4. I think the strategy of predicting is effective in promoting reading comprehension.	25 (69.5%)	9 (25%)	2 (5.7%)

Subjects' Belief in the Effect of Group Discussion in Applying the Four Strategies.

The subjects' opinions about the effects of group discussion in the application of the four strategies were clearly illustrated and analyzed in Table 6. Of thirty-six subjects, thirty-two subjects (88.9%) considered group discussion beneficial "in applying the strategy of summarizing," twenty-seven (75%) "in applying the strategy of questioning," twenty-five (69.5%) "in applying the strategy of predicting," and thirty-four (94.4%) "in applying the strategy of clarifying." As for the negative point of view, only one subject (2.8%) did not believe in the effect of group discussion "in applying the strategy of summarizing," one (2.8%) "in applying the strategy of questioning," two (5.7%) "in applying the strategy of predicting," and one (2.8%) "in applying the strategy of clarifying." Overall, the subjects believed that group discussion was effective in applying these strategies to increase comprehension. This result is consistent with those of Johnson and Johnson (1989) in that oral discussion among group members helps individuals restructure their thinking process.

Table 6. *Subjects' Belief in the Effect of Group Discussion in the Application of the Four Strategies*

Questions	Frequencies (Percentages)		
	Positive	Neutral	Negative
	Strongly agree / Agree		Disagree / Strongly disagree
5. I think group discussion is very effective in applying the strategy of summarizing.	32 (88.9%)	3 (8.3%)	1 (2.8%)
6. I think group discussion is very effective in applying the strategy of questioning.	27 (75%)	8 (22.2%)	1 (2.8%)
7. I think group discussion is very effective in applying the strategy of predicting.	25 (69.5%)	9 (25%)	2 (5.7%)
8. I think group discussion is very effective in applying the strategy of clarifying.	34 (94.4%)	1 (2.8%)	1 (2.8%)

Subjects' Evaluation of Reciprocal Teaching in Promoting and Monitoring Comprehension. Concerning if the subjects could evaluate their learning and consider reciprocal teaching beneficial in promoting and monitoring their comprehension, the subjects' opinions were clearly illustrated and analyzed in Table 7. With regard to the ninth question, twenty-nine subjects (80.5%) considered that reciprocal teaching helped them realize their own reading process whereas three (8.5%) disagreed with its efficiency. As for the tenth question, thirty-one subjects (86.1%) thought reciprocal teaching was beneficial in realizing others' reading process in contrast to one (2.8%) with opposite viewpoint. As regard to the eleventh question, thirty subjects (83.3%) considered that reciprocal teaching promoted their reading comprehension in contrast to two (5.7%) who disagreed with this idea. When it came to if reciprocal teaching could modify their thinking process, thirty-three (91.6%) held positive attitudes, whereas only one subject (2.8%) held negative attitudes. Finally, as for if reciprocal teaching could help them read English with a faster speed than before, twenty-seven subjects (75%) were positive, whereas two (5.7%) held negative attitudes. In a word, this finding is accord with those of Lederer (2000), Palincsar, Brown, and Martin (1987) and Palincsar (1986) in that reciprocal teaching can not only help them to enhance comprehension but also help them to monitor and modify their thinking process through peer interaction.

Table 7. *Subjects' Evaluation of Reciprocal Teaching in Promoting and Monitoring Comprehension*

Questions	Frequencies (Percentages)		
	Positive	Neutral	Negative
	Strongly agree /Agree		Disagree /Strongly disagree
9. Reciprocal teaching helps me realize my own reading process.	29 (80.5%)	4 (11.1%)	3 (8.5%)
10. Reciprocal teaching helps me realize others' reading process.	31 (86.1%)	4 (11.1%)	1 (2.8%)
11. Reciprocal teaching promotes my reading comprehension.	30 (83.3%)	4 (11.1%)	2 (5.7%)
12. Reciprocal teaching modifies my thinking process.	33 (91.6%)	2 (5.7%)	1 (2.8%)
13. Reciprocal teaching helps me read English with a faster speed than before.	27 (75%)	7 (19.4%)	2 (5.7%)

Subjects' Attitude toward Applying the Experience of the Activity in the Future.

The subjects' attitude toward applying this activity in their future English reading was shown in Table 8. With regard to if they would apply these four strategies in the following-up English reading classes, twenty-five subjects (69.4%) agreed or strongly agreed, whereas only two of them (5.7%) were negative. As for if they would like to apply these four strategies in reading other subjects like history or social studies, eighteen of them (50%) agreed or strongly agreed, whereas only three (8.3%) held negative attitudes. Last, as for if they would apply these four strategies while reading alone, twenty of them (55.6%) agreed or strongly agreed, whereas only one subject (2.8%) did not want to apply the strategies while reading alone.

Table 8. *Subjects' Attitude toward Applying the Experience of the Activity in the Future*

Questions	Frequencies (Percentages)		
	Positive	Neutral	Negative
	Strongly agree /Agree		Disagree /Strongly disagree
14. I will try to apply these four reciprocal teaching strategies in the following-up English reading classes.	25 (69.4%)	9 (25%)	2 (5.7%)
15. I will try to apply these four reciprocal teaching strategies in reading other subjects like history and social studies.	18 (50%)	15 (41.7%)	3 (8.3%)
16. I will try to apply these four reciprocal teaching strategies while reading alone.	20 (55.6%)	15 (41.7%)	1 (2.8%)

DISCUSSIONS

According to the analysis of the collected data and the comparison of the statistical results, the research findings are discussed as follows.

Effects of Reciprocal Teaching on EFL Learners' Reading Comprehension

Reciprocal teaching is effective on promoting EFL learners' reading comprehension because it is more beneficial in enhancing EFL students' general reading comprehension in Taiwan than the traditional reading instruction. Through comparing the reading

comprehension tests between the pre-and posttest, a significant difference was found in the experimental group. The thirty-six subjects in this study showed a great improvement in their reading comprehension. As for the effects of the traditional reading instruction on promoting EFL reading comprehension, there was also a significant difference in the control group's progress. However, even though the subjects in both the experimental group and the control group made progress in their reading comprehension, the differences of mean scores of the pre-and posttest in the experimental group and the control group were 15.67 and 7.41 respectively (Table 4). Obviously, the experimental group made much more improvement than the control group. Consequently, the reciprocal teaching instruction is more effective in promoting EFL learners' reading comprehension than the traditional reading instruction.

One reason for the experimental group's promotion in reading comprehension might lie in the fact that the four metacognitive strategies in reciprocal teaching help students to monitor and enhance their reading comprehension. In reciprocal teaching, the four metacognitive strategies students need to learn and practice are the strategies of predicting, questioning, clarifying and summarizing. The strategy of predicting helps students activate their prior knowledge to recall what they know about the topic and hypothesize about what the author might discuss in the following paragraphs. The strategy of questioning requires students to comment on the significance of the content and question the information in the text. This way, students learn to monitor their own thinking process. The strategy of clarifying involves searching for nearby relevant information to determine the meaning of a word or a portion of the text, and using the students' knowledge and experience to explain, extend, and clarify content or unknown words. While applying this strategy, students use their metacognitive process to monitor their reading comprehension (King & Johnson, 1999). While learning the strategy of summarizing, students are required to pay attention to the most important information or the gist of a text. Hashey and Connors (2003) stated that summarizing was an efficient strategy to enhance students' comprehension because it required them to focus on key points, not to restate everything. In short, these students haven't learned any reading strategy before; therefore, after the reciprocal teaching training, they learned how to monitor and enhance their reading comprehension. The result concurs with Eilers and Pinkley's study (2006) in that metacognitive reading strategies could help students to improve students' ability in comprehending text.

Another reason might be the fact that reciprocal teaching provides a more meaningful social activity for promoting language use and comprehension than a traditional whole-class reading instruction. Junior high school teachers seldom use group discussion in a traditional EFL reading classroom and they tend to dominate classroom activities. Students have few chances to express their feelings and their thoughts about the text as well. Reciprocal teaching can not only help students to learn language more actively and confidently through group discussion, but also reduce students' anxiety in reading a foreign language.

Applying Reciprocal Teaching in EFL Reading Classrooms in Taiwan

As shown in the results, the reading comprehension in the experimental group was enhanced significantly, and the experimental group even made much more improvement than the control group. Furthermore, most of the subjects in the experimental group were affirmative about the efficiency of reciprocal teaching on promoting and monitoring comprehension in the response questionnaire. Therefore, the findings of the present study might suggest that reciprocal teaching is applicable to be embedded within the regular curriculum and be delivered by English teachers in Taiwanese junior high school reading classrooms.

PEDAGOGICAL IMPLICATIONS

While conducting reciprocal teaching, teachers should be patient about students'

learning. Reciprocal teaching puts emphasis not only on teachers' scaffolding but also on students' teaching and learning by interacting with each other. The role of teachers in reciprocal teaching is a facilitator or a consultant but not an authority, which is quite different from what they used to be in traditional reading classrooms. Traditionally, teachers take control of their class learning, so they might feel that group discussion is loud and unruly. However, teachers need to keep in mind that reciprocal teaching is to provide a less teacher-centered environment, promote students' thinking, and increase their involvement with text. In order for reciprocal teaching training to be successful, teachers should be patient about students' learning and give them many opportunities and much time to practice and learn the four metacognitive strategies.

CONCLUSION

This study explored the effects of reciprocal teaching on EFL junior high school students' reading comprehension. The study also probed the subjects' response to the reciprocal teaching instruction. The findings were as follows. First, the reciprocal teaching instruction was more beneficial in enhancing EFL junior high school students' reading comprehension in Taiwan than the traditional reading instruction. Even though significant differences were found in both groups, the reading comprehension was improved much more in the experimental group than in the control group. Second, most subjects were affirmative about the efficiency of reciprocal teaching on promoting and monitoring comprehension. Among the four strategies, they considered the strategies of clarifying and summarizing most effective on increasing their comprehension. The results suggest that the reciprocal teaching instruction can be an alternative for teaching reading in an EFL junior high school classroom. Moreover, for metacognitive training in reading to be successful, teachers should give students many opportunities and much time to practice applying the strategies.

REFERENCES

- Alfassi, M. (2004). Reading to learn: Effects of combined strategy instruction on high school students. *The Journal of Educational Research*, 97 (4), 171-184.
- Al-Hilawani, Y. A., Marchant, G. J., & Poteet, J. A. (1993). Implementing reciprocal teaching: Was it effective? Paper presented at the 23rd annual meeting of Midwest Association of Teachers Educational Psychology, Anderson, IN.
- Anderson, T. H. (1980). Study strategies adjunct aids. In R. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension*, Hillsdale, NJ: Erlbaum.
- Baker, L. (1989). Metacognition, comprehension monitoring, and the adult reader. *Educational Psychology Review* 1 (1), 3-38.
- Baker, L., & Brown, A. (1984). *Metacognitive skills and reading*. In P. D. Pearson (Eds.), *Handbook of reading research* (pp. 353-394). New York: Longman.
- Block, E. L. (1986). The comprehension strategies of second language readers. *TESOL Quarterly*, 20 (3), 463-494.
- Brown, A. L. (1985). Metacognition: the development of selective attention strategies for learning from texts. In H. Singer & R. B. Ruddel (Eds.), *Theoretical models and processes of reading* (pp. 501-526). Newark, DE: International Reading Association.
- Bruce, M. E., & Robinson, G. L. (2001). The clever kid's reading program: Metacognition and reciprocal teaching. Paper presented at the 12th European Conference on Reading, Dublin University.
- Byrd, D. M., & Gholson, B. (1985). Reading, memory, and metacognition. *Journal of Educational Psychology*, 77, 428-436.
- Carrell, P. L. (1988). Metacognitive awareness and second language reading. *Modern Language Journal*, 73 (2), 121-134.

- Carter, C. (1997). Why reciprocal teaching? *Educational Leadership*, 54 (6), 64.
- Chen, C. C. (2005). *A study of the effects of expository texts with structure guidelines on Taiwanese EFL senior high school students' reading comprehension*. Unpublished master's thesis, National Kaohsiung Normal University: Kaohsiung.
- Chen, M. C. (2005). *The reading comprehension effectiveness research of the mutual instruction based on the content of the text book towards mild mentally retarded students*. Unpublished master's thesis, National Taichung University: Taichung.
- Clark, L. (2003). Reciprocal teaching strategy and adult high school students. Paper presented in partial fulfillment of the requirements for the Masters of Arts Degree in Reading Specialization.
- Cotterall, S. (1990). Reading strategy training in second language contexts: Some caveats. *Aral*, 16(1), 71-82.
- Duffy, G., Roehler, L. & Herrmann, B. A. (1988). Modeling mental processes helps poor readers become strategic readers. *The Reading Teacher*, 41, 762-767.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring. *American Psychologist*, 34, 906-911.
- Garner, R., & Kraus, C. (1981-1982). Good and poor comprehenders differences in knowing and regulating reading behaviors. *Educational Research Quarterly*, 6, 5-12.
- Guerlene, S. (2002). Improving reading comprehension: A comparative study of metacognitive strategies. Paper presented in Partial Fulfillment of the Requirements for the Master of Arts Degree in Reading Specialization Kean University.
- Herrmann, B. A. (1988). Two approaches for helping poor readers become more strategic. *The Reading Teacher*, 42 (1), 24-28.
- Hirschfeld, D. (Ed.). (2004). *Fluency kits for independent practice*. NY: Benchmark Education Company.
- Hus, L. R. (2003). *A study of relationships between feeling of knowing about English reading strategy use and reading comprehension of Taiwanese college students*. Unpublished master's thesis. National Changhua University, Changhua.
- Hus, W. L. (2004). *An Analysis of the Reading Comprehension Questions in the JCEE English Tests*. Unpublished master's thesis, National Kaohsiung Normal University: Kaohsiung.
- Janzen, A., & Thomas, J. (2007). *Reading builder*. New York: Compass Publishing.
- Johnson, P. (1982). Effects on reading comprehension of language complexity and culture background of text. *TESOL Quarterly*, 15, 169-181.
- Johnson, D. W., & Johnson, R. T. (1989). Cooperative learning: What special educators need to know. *The Pointer*, 33, 5-10.
- Lederer, J. M. (2000). Reciprocal teaching of social studies in inclusive elementary classrooms. *Journal of Learning Disabilities*, 33, 91-106.
- Mikulecky, B. S. (1990). *A short course in teaching reading skills*. Reading, MA: Addison-Wesley.
- O'Malley, J. M., & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.
- Palincsar, A. S. (1986). The role of dialogue in scaffolded instruction. *Educational Psychologist*, 21, 73-98.
- Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and monitoring strategies. *Cognition and Instruction*, 1, 117-175.
- Palincsar, A. S., Brown, A. S., & Martin, S. M. (1987). Peer interaction in reading comprehension instruction. *Educational Psychologist*, 22, 231-253.
- Paris, S. G., & Jacobs, J. E. (1984). The benefits of informed instruction for children's reading awareness and comprehension skills. *Child Development*, 55, 2083-2093.

- Paris, S. G., Lipson, M. Y., & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology*, 8, 293-316. In R. Ruddell, M. Ruddell, & Singer, H. (Eds.), *Theoretical models and process of reading*. Newark Lawrence: International Reading Association.
- Paris, S. G., & Myers, M. (1981). Comprehension monitoring, memory, and study strategies of good and poor readers. *Journal of Reading Behavior*, 13, 5-22.
- Paris, S. G., & Winograd, P. (1990). How metacognition can promote academic learning and instruction. In B. F. Jones & L. Idol (Eds.), *Dimensions of thinking and cognitive instruction* (pp. 15-51). Hillsdale, NJ: Erlbaum.
- Phillips, J. K. (1984). Practical implications of recent research reading. *Foreign Language Annual*, 17 (4), 309-312.
- Pressley, M. (2000). What should comprehension instruction be the instruction of? In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.). *Handbook of reading research*, 3, 545-563, Mahwah, NJ: Erlbaum.
- Rosenshine, B., & Meister, C. (1994). Reciprocal teaching : A review of the research. *Review of Educational Research*, 64, 479-530.
- Schraw, G. (2001). Promoting general metacognitive awareness. In H. J. Hartman (Ed.), *Metacognition in learning and instruction: Theory, research and practice* (pp. 3-16). Boston: Kluwer Academic Publisher.
- Snow, C. (2002). *Reading for understanding: Toward an R & D program in reading comprehension*. Retrieved October 18, 2003, from <http://www.rand.org/publicaions/MR/MR4465/index.html>
- Yang, Y. F. (2002). Reassessing readers' comprehension monitoring. *Reading in a Foreign Language*, 14 (1), 1-26.

A Study of Cloze Test Items in Scholastic Aptitude English Test and Department Required English Test

Su-yu Chou (周思余), Yuh-mei Chen (陳玉美)

National Chung Cheng University

ssuyuchou@gmail.com
folcym@ccu.edu.tw

This study attempts to analyze items in Zong-He-Ce-Yian section in Scholastic Aptitude English Test (SAET) and Department Required English Test (DRET) from 2004 to 2008 to find out what kind of language processing skills, discrete-point or integrative skill, does the cloze items in both tests measure.

The results indicate that both SAET and DRET contain a higher ratio of type 1 and type 2 items, measuring examinees' lower-order language skills. The distribution of the four categories is consistent in SAET from 2004 to 2008, whereas in DRET, it is found that there is a higher ratio of the global items from 2006 to 2008. As for examinee's performance on each type in both exams, the average passing rate of all of the Zong-He-Ce-Yian items in SAET (51%) is slightly higher than that in DRET (47%).

INTRODUCTION

"Cloze" was first used by Taylor (1953) to measure the readability of a reading passage. According to Taylor, the standard procedure to generate a cloze passage is to select a passage, delete every n^{th} word (also known as fixed-ratio deletion), provide a blank for the missing word, and then subjects are asked to fill in the exact word in the blanks. The use of cloze passage introduces a crucial concept into the field of language testing, that is, the idea of total language context, indicating that language behavior depends on "total context." Therefore, it is claimed that different from the discrete-point test, which is designed to measure students' knowledge or performance in a restricted area of target language, the cloze test is able to evaluate students' overall language proficiency, and is regarded as integrative test. As such, cloze passage has been widely used in language testing.

Cloze tests have long been used in various testing contexts in Taiwan for decades, especially in middle school levels. In nationwide English tests for senior high school students such as Scholastic Aptitude English Test (SAET) and Department Required English Test (DRET), there is a section containing a passage from which words or phrases are deleted. It is called "Zong-He-Ce-Yian" as "cloze test." Zong-He-Ce-Yian, translated as comprehensive test, includes vocabulary, grammar and other linguistic features. Yang (1996) indicates that Zong-He-Ce-Yian is a modified cloze test rather than the traditional standard cloze procedure. In both SAET and DRET, Zong-He-Ce-Yian is around one third of multiple choice items; examinee's performance on this part plays a crucial role in determining their English grades, which might also have a lifelong effect on whether examinees are able to attend their ideal college or not.

In Taiwan, both Scholastic Aptitude English Test (SAET) and Department Required English Test (DRET) are administered to senior high school students who are going to attend universities or colleges. According to Ministry of Education, even though both of the tests are designed for high school students, the difficulty levels of the two examinations are different

because the purposes of the two tests are different. Jeng et al. (1999, 2001) propose that items in SAET are supposed to be easier than those in DRET. They identify that SAET is designed to evaluate students' basic knowledge learned from senior high school; therefore, it is categorized as an achievement test. Whereas DRET aims to screen students to select superior students to attend college, it is more like a proficiency test.

Yang (1996) points out that the cloze test has been extensively used in EFL testing in Taiwan; however, in practice, it has been considered the equivalent of the fill-in-the-blank or sentence-completion task. Through analyzing cloze items in secondary school monthly English tests and interviewing local English teachers, Yang found that in most cases, the deleted words in a cloze passage can be recovered by using clausal or sentential cues, and teachers are not aware of the true characteristics of cloze procedure. He concludes, "the types of test can at best be called a series of discrete-point completion items addressing the same topic." (p. 62)

Even though the cloze test is claimed to be an integrative test, which is different from discrete-point or sentence completion tasks, little research has been conducted to analyze cloze items in local settings, to examine whether cloze items in Zong-He-Ce-Yian section is able to function as an integrative test, or the section can at best be regarded as a series of discrete-point completion items addressing the same topic.

Bachman (1985) demonstrates that the cloze test taps complex skills ranging from lower- to higher-order abilities of language processing. Based on the level of context required for recovering the missing word, Bachman (1982, 1985) categorizes cloze items into four types, including (1) within clause, (2) across clause, within sentence, (3) across sentence, and (4) extra-textual. He proposes that the way to examine whether a cloze test is a discrete-point test is through item-type analysis.

Yang (1996) points out that generally, type 1 and type 4 are associated with bottom-up or discrete-point skills, whereas type 2 and type 3 deal with cross-sentential and integrative skills. Based on Bachman's taxonomy, Jonz (1990) develops a four-part categorization decision to concisely categorize cloze items. Jonz (1991) further modifies Bachman's taxonomy, and introduces a five-category item-type taxonomy in his study. In the study, he makes the "within clause" category more sensitive to the distinction between syntactic and non-syntactic items. Therefore, there are two within-clause subcategories, syntax and local lexis. Jonz assumes that lexical choices represent higher-order knowledge, and he wanted to segregate the locally constrained higher-order choices from choices associated essentially with syntax.

The present study attempts to take Jonz's taxonomy to categorize items in Zong-He-Ce-Yian in SAET and DRET to find out what kinds of language processing skill, discrete-point or cross-sentential skill, both tests measure.

RESEARCH QUESTIONS

Based on what has been mentioned above, the current study aims to address the following three research questions:

1. Since the purposes of SAET and DRET are different, are the item types in Zong-He-Ce-Yian section different in both tests?
2. Based on the amount of the text content required to answer an item, the items are categorized into global and local items. What kind of language skill, global or discrete-point skill, does the modified cloze test measure?
3. How do the examinees in general perform on items in Zong-He-Ce-Yian section?

METHODOLOGY

Material

Data in the present study include the reading passages and test items in Zong-He-Ce-Yian section in both Scholastic Aptitude English Test (SAET) and Department Required English Test (DRET) from 2004 to 2008. The material was obtained from the website of College Entrance Examination Center (CEEC) (<http://www.ceec.edu.tw>). In general, Zong-He-Ce-Yian section consists of three to four passages, and each passage contains five deleted words. A total of 34 passages and 170 items in SAET and DRET from 2004 to 2008 were collected.

Instrument

According to the level of context required for recovering the missing word, Bachman (1982, 1985) categorizes cloze items into four types, including (1) within clause, (2) across clause, within sentence, (3) across sentence, and (4) extra-textual. Based on Bachman's categorization, Jonz (1990) develops a four-part categorization decision to concisely categorize cloze items.

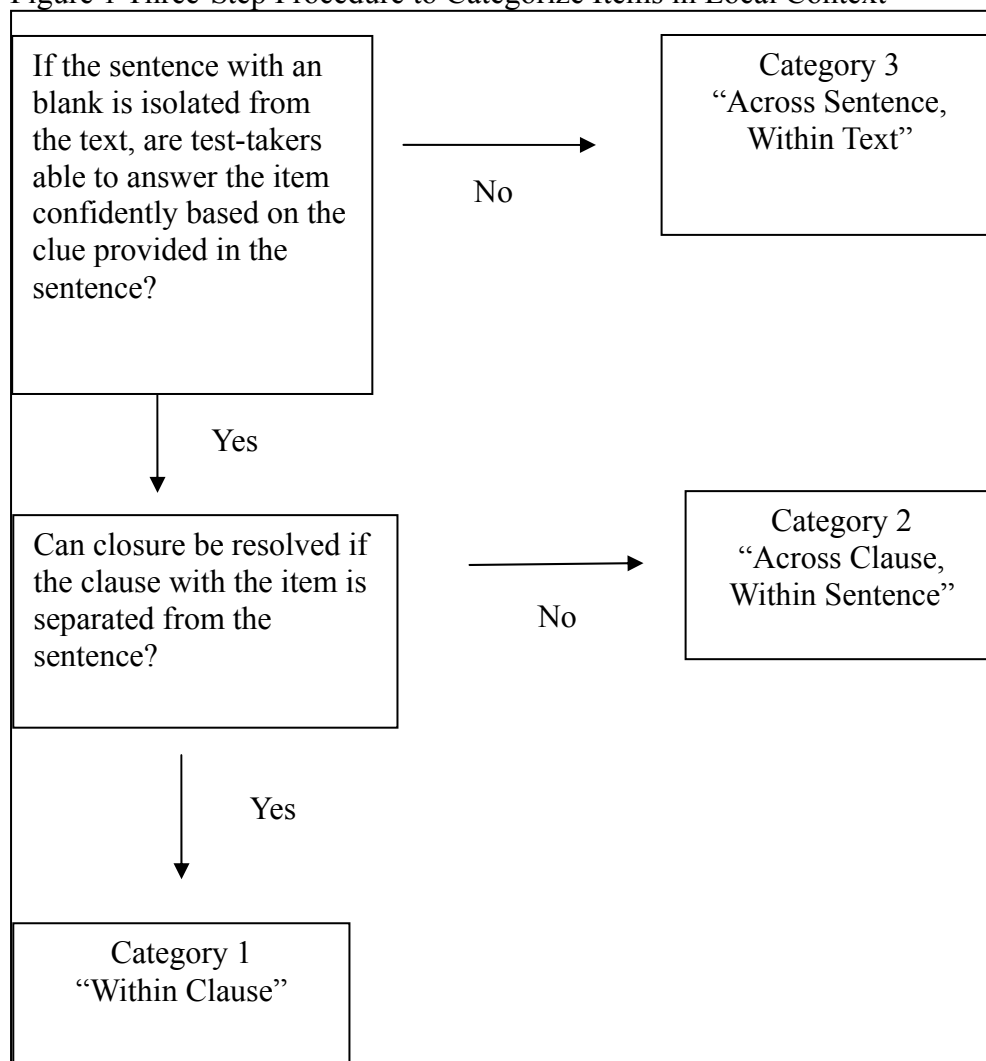
However, the application of the cloze test in local context is different from that in other contexts in two perspectives. First of all, in both SAET and DRET, cloze test is applied as multiple choice; four options are provided for test-takers to choose from. The standard cloze procedure requires the test-takers to fill in the cloze deletions by themselves. Secondly, in local context, more than one word may be deleted in one item, that is, a cloze item may contain a three-word phrase, but in standard cloze procedure, a blank is strictly for only a word. Because of the differences mentioned, Jonz' taxonomy does not completely fit in local context.

In local context, because cloze procedure is applied as multiple choices, that is, test-takers do not have to provide answers for each blank by themselves, what they have to do is to select the correct answer from the provided options. Under the circumstances, because one of the four options is correct, the options are served as clues for the test-takers. Therefore, category 4 is excluded in local context. Due to the multiple-choice format of the cloze test in SAET and DRET, a modified categorization procedure is proposed. Figure 1 reveals the three-step categorization procedure used in the current study (please refer to next page).

The Distinction between Global and Local Items

Yang (1996) points out that in Bachman's taxonomy, items categorized as type 1 and type 4 measure test-takers' bottom-up or discrete-point skills, while type 2 and type 3 require cross-sentential or global skills. In order to meet the purpose of the study, investigating whether the cloze test in local context functions as an integrative test or just as what Yang (1996) claims "a series of discrete-point completion items addressing the same topic" (p.62), an adjustment of Yang's proposition for discrete-point skill and global skill is made. Since test-takers are able to restore a blank within a sentence level, that is, if a sentence with a blank is isolated from the original passage, learners are able to answer the item correctly without referring to other sentences provided; then the item is not different from the sentence completion task. Therefore, in the present study, only items which require test-takers to refer to the clue beyond sentence level are considered as integrative items. In short, only items of category 3 are considered as global items, whereas, items classified into category 1 and 2 are regarded as local items.

Figure 1 Three-Step Procedure to Categorize Items in Local Context



Data Analysis

The coding and analysis procedure in the study include three steps: (1) have all the items in Zong-He-Ce-Yian section in both tests properly categorized into one of the three general categories, (2) find out the frequencies of each item type to know what kind of language skill, integrative or discrete skill, the modified cloze test tend to measure, and (3) analyze the passing rates on each type of item.

Three-Step Categorization Procedures

A three-step categorization procedure is proposed in the present study to classify items into three general categories (see Figure 1). The item categorization procedure starts from checking whether an item can be answered confidently and correctly without referring to clues beyond the sentence level. If the answer is negative, then the item is identified as Category 3, indicating that the item require learners to refer to a broad context that is beyond sentential level. On the other hand, if the answer is positive, then the researcher has to further explore whether the item require test takers to refer to clues at a sentential or clausal level. Therefore, Category 3 is the first item type being identified in current study. The following examples illustrate the three-step categorization procedure.

Category 3 — Across Sentences, Within Text

Based on Figure 1, to correctly categorize the items into three types, the researcher first isolates the sentences contains a blank from the original passage. Therefore, the original

passage was divided into around five or less than five sentences (in some cases, a sentence contains more than one blanks).

By doing so, the researcher is able to differentiate category 3 from the other two categories. If the researcher cannot confidently select the answer from the provided options, that is, the clue in the sentence is not sufficient; then the item is categorized as type 3.

Example 1

__(17)__ was the ability to identify the belongings of his predecessor, or rather his previous self.

- (A) It (B) All *(C) One (D) This

(SAET 2004)

The correct answer for Example 1 is (C). If we only look at the given sentence without referring to the whole text, we can find that both (A) and (D) are acceptable. When we refer to the original passage, we will select (C) as the correct answer. In the original passage, what follows the above example is, “Another requirement was that he should have large ears...” The word “another” specifies that there is a relationship between the two sentences, and “one” is the antecedent for “another.”

Category 2 – Across Clause, Within Sentences

If the clue provided in the sentence is sufficient, that is, based on the sentence provided, learners are able to answer the item confidently, then the researcher has to further segment the sentence into clauses to identify whether it is the clause that contain the item provide the clue or the clue is beyond the clause level.

Example 2

Certainly, children must start with enjoyment __22__ their interest in poetry dies.

- (A) or (B) and (C) so (D) then

(From DRET 2004)

Four conjunctions are provided as options in Example 2. To answer the item, test-takers have to refer to the two clauses around the blank to find out the relationship between the two clauses. Therefore, the clue of the item is beyond the clause level, and the item is classified as type 2, across clause within sentence.

Category 1 – Within Clause

If test-takers are able to answer the item that is within a clause, in other words, the clue for the item is in clause level; then the item is classified as category 1.

Example 3

The Chinese and Indians __16__ it for at least 4,000 years to treat everything from headache to depression.

- (A) would use (B) are using (C) had used (D) have been using

(From SAET 2008)

The correct answer for Example 3 is (D). There is only one clause in the example, and the clue for the clause is “for at least 4,000 years.” Since the test takers are able to answer the item correctly without referring to clues outside the clause, the item is classified as category 1.

In Jonz (1990), two subtypes are identified under category 1: clause-level syntax and clause-level lexis. His adjustment of Bachman’s taxonomy is under the assumption that between these two types, lexical choices represent higher-order knowledge than syntax. The present study also classifies items under category 1 into two subtypes, *Clause-level Syntax* and *Clause-level Lexis*, to investigate whether there is difference in test-takers’ performances

on these two subtypes.

The options provided are used as clues to classify items into subcategories. For example, in Example 3, the word “use” appears in the four options, and the differences among the four options are the syntactic elements. The type of item is categorized as *Clause-level Syntax*.

Example 4

Some environmental scientists suggest turning newspapers __16__ charcoal.

*(A) into (B) for (C) off (D) upon

(From DRET 2004)

In Example 4, options with four different prepositions are provided. Based on the clue provided in the sentence, test-takers' are required to select the best proposition for the clause. This type of item is categorized as *Clause-level Lexis*.

RESULTS

Distribution of Different Item Types in SAET and DRET

SAET

Table 1 presents the distribution of different item types in Zong-He-Ce-Yian section in SAET from 2004 to 2008.

Table 1. *Number of Different Types of Items in SAET from 2004 to 2008*

	Clause Level Syntax	Clause Level Lexis	Across Clause within Sentence	Across Sentence, within Text
2004	3	6	1	5
2005	3	4	5	3
2006	3	4	5	3
2007	1	7	4	3
2008	2	7	4	2
Total (%)	12 (16%)	28 (37%)	19 (25%)	16 (22%)

As shown in Table 1, in SAET from 2004 to 2008, among the four categories, 28 out of 75 items (37%) in Zong-He-Ce-Yian section were categorized as *Clause-level Lexis*, which was the most frequent item type. Category 2, *Across Clause within Sentence* items the second most (25%). *Across Sentence within Text*, the third (22%). The result of Chi-square test demonstrates that there is no significant difference among the four item types in SAET ($\chi^2 = 7.4$, $p = .06$), revealing that the test designers do not put emphasis on any of the four categories. In addition, the result of item frequency in SAET shows that around 75% of items are categorized as category 1 and 2, which are used to measure learners' discrete-point skill, and only 22% of the items are classified as global items.

DRET

Table 2 offers the distribution of the four item types in DRET from 2004 to 2008 (please refer to next page). As shown in Table 2, it is found that out of 95 items in DRET from 2004 to 2008, the proportion of the following three categories are similar – *Clause-level lexis*, *Across Clause within Sentence*, *Across Sentence within Text* – each category consists of around one third of the items. The result also demonstrates that only 4 items (4%) is classified as *Clause-level Lexis* items. Furthermore, if we further categorize items into two general types, global and discrete-point items, we find that up to 69% of items are discrete-point items, whereas only 31% are global items. The result is in accordance with what has been found in SAET, that is, the items in Zong-He-Ce-Yian section in both

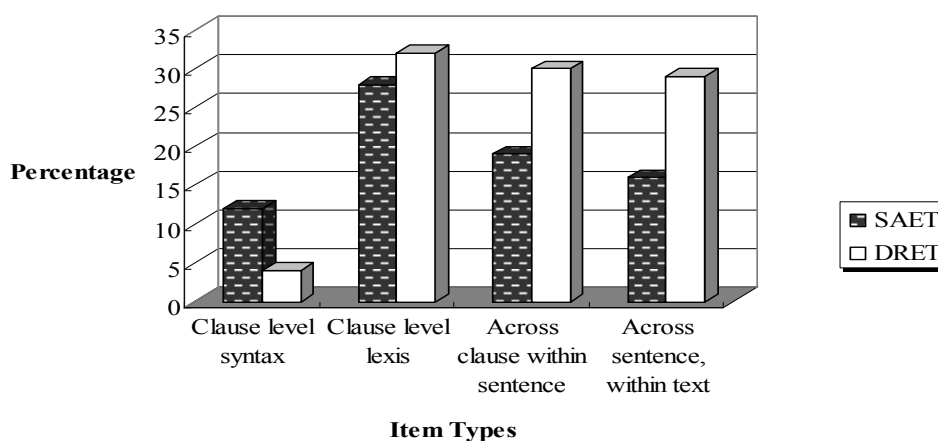
nation-wide examinations focus on assessing learners' discrete-point skills rather than integrative skills.

Table 2. Number of Different Types of Items in DRET from 2004 to 2008

	Clause Level Syntax	Clause Level Lexis	Across Clause within Sentence	Across Sentence, within Text
2004	1	4	7	3
2005	1	8	9	2
2006	0	6	5	9
2007	1	7	4	8
2008	1	7	5	7
Total (%)	4 (4%)	32 (33%)	30 (32%)	29 (31%)

Similarities and Differences of Item Types in SAET and DRET

Figure 2 Percentage of the Four Item Types in SAET and DRET from 2004 to 2008.



Similarities

As can be seen in Figure 2, the ranking order of frequency of the four types in SAET and DRET is identical. The most frequent item type is *Clause-level Lexis*, which is followed by *Across Clause within Sentence*, and *Across Sentence within Text*, and *Clause-level Syntax* accounts for the least proportion among the four categories. In the present study, only items that require test-takers to refer to clues that are beyond sentence level is defined as global items; as a result, only Category 3, *Across Sentence within Text* is defined as global items. 78% and 69% of the Zong-He-Ce-Yian items in SAET and DRET are classified as local items, and 22% and 31% in SAET and DRET respectively are global items. The result indicates that there is an approximately equal distribution of the four item types in local context. However, if we divide items into two general types, global and local items, we find that the modified cloze test puts emphasis on local items.

Differences

A close review of items year by year has found that there is a higher ratio of the global items in DRET from 2006 to 2008. Table 3 illustrates the proportion of the local and global items in SAET and DRET from 2004 to 2008.

Table 3. *Proportion of the Local and Global Items from 2004 to 2008*

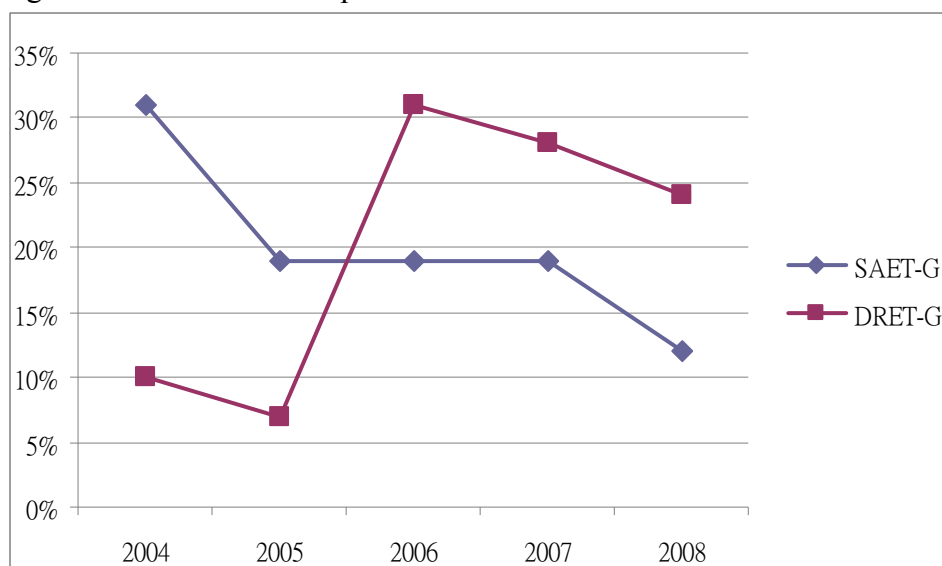
	2004	2005	2006	2007	2008
SAET-L	17%	20%	20%	20%	22%
DRET-L	18%	27%	17%	18%	20%
SAET-G	31%	19%	19%	19%	12%
DRET-G	10%	7%	31%	28%	24%

Note. SAET-L: the proportion of local items in SAET; DRET-L: the proportion of local items in DRET; SAET-G: the proportion of global items in SAET; DRET-G: the proportion of global items in DRET.

From 2004 to 2008, there is a total of 59 and 66 local items in SAET and DRET respectively. The average proportion of the local items in each year is 20%. There is no significant difference in the change of the proportion of the local items in both exams in these five years (SAET: $\chi^2 = 0.65$, $p = .95$; DRET: $\chi^2 = 3.3$, $p = .50$).

As for global items, there is a total of 16 and 29 global items in SAET and DRET respectively. If we look closely at the global items in both exams from 2004 to 2008, we find that the proportion of the global items in SAET decreases gradually, whereas, the proportion of the global items in DRET increases obviously. Table 3 reveals that in DRET, the proportion of the global items ranges from 7% to 31%. Figure 3 reveals the curves of the global items in SAET and DRET in the five years.

Figure 3 Curves of the Proportion of the Global Items from 2004 to 2008



As shown in Figure 3, we can see that there is an obvious change in the distribution of the global items in the five years. There is only a total of 5 out of 29 global items from 2004 and 2005, which only takes 17% of the global items; whereas from 2006 to 2008, there is an obvious increase in global items in these three years. If we further refer to Table 2, we can find that the proportion of Category 3 increases from 2004 which accounts only for 15% of the four categories to 2006, accounting for 45% of the items.

Test-takers' Performances on Each Question Type

The passing rates on each type of question were examined to see how the examinees in general performed on the four different types of questions.

SAET

Table 4. Average Passing Rates of the Four Categories in SAET from 2004 to 2008

	Clause Level Syntax	Clause Level Lexis	Across Clause within Sentence	Across Sentence, within Text
2004	42% (3)	47% (2)	52% (1)	38% (4)
2005	55% (3)	42% (4)	58% (2)	64% (1)
2006	53% (1)	50% (3)	52% (2)	47% (4)
2007	38% (3)	50% (1)	49% (2)	33% (4)
2008	69% (1)	63% (3)	68% (2)	51% (4)
Average	51% (2)	50% (3)	56% (1)	47% (4)

Note. Numbers in () represents the rank order of each category within a year.

The passing rate of the four categories in SAET is provided in Table 4. As shown in Table 4, the average passing rate of each category ranges from 47% to 56% from 2004 to 2008. Category 2 is the best performed item, with the highest passing rate (56%), followed by Clause-level Syntax items (51%), Clause-level Lexis (50%), and Category 3 (47%).

In general, in SAET from 2004 to 2008, the four types of items did not receive consistent ranks every year. For example, Across Clause within Sentence ranked the first in 2004, whereas in 2005, it is Across Sentence within Text that ranked first. However, if we vertically compare the rank order of each category year by year, we can find that in general, among the four categories, Category 3 had the lowest passing rate, except 2005. The result reveals that, among the four categories, Category 3, which requires learners to refer to clues that are beyond sentential level, is the most difficult item type in SAET.

DRET

Table 5. Average Passing Rates of the Four Categories in DRET from 2004 to 2008

	Clause Level Syntax	Clause Level Lexis	Across Clause within Sentence	Across Sentence, within Text
2004	60% (2)	51% (3)	44% (4)	64% (1)
2005	57% (1)	46% (3)	43% (4)	54% (2)
2006	0	51% (2)	44% (3)	56% (1)
2007	43% (1)	42% (2)	39% (3)	36% (4)
2008	67% (1)	42% (4)	52% (3)	56% (2)
Average	57% (1)	46% (3)	44% (4)	53% (2)

Table 5 reports the average passing rates of the four categories in DRET from 2004 to 2008. The average passing rate of each category ranges from 44% to 57%. Clause-level Syntax is the best performed item, with the highest passing rate (57%), followed by Category 3 (53%), Clause-level Lexis (46%), and Category 2 (44%). Like what has been observed in SAET, the four item types did not receive consistent rank order.

DISCUSSIONS

Item Types in SAET and DRET

The result reveals that the ranking order of the frequencies of the four categories is

identical. The most frequent item type is Clause-level Lexis, which is followed by Across Clause within Sentence, and Across Sentence within Text, and Clause-level syntax accounts for the least proportion among the four categories. Furthermore, the present study has found that in SAET and DRET, there is no significant difference in the proportion of the item types. Another noteworthy finding is that the ratio of the local items is consistent in SAET in the five years, whereas in DRET, the ratio of the global items increases from 2006.

Test-Takers' Performances on Each Question Type

In both exams, it was hard to see a general pattern on which types of questions were best or worst performed throughout the five years. Test takers performed quite differently on the same type of question every year, indicating that each type of question was at different difficulty level each year. In general, category 3 is the most difficult type in SAET, and the most difficult type in DRET is category 2. Besides, the result shows that the average passing rate of all of the Zong-He-Ce-Yian items in SAET (51%) is slightly higher than that in DRET (47%), indicating that SAET is slightly easier than DRET. The result of the study is in accordance with CEEC's proposition which suggests that SAET is supposed to be easier than DRET.

CONCLUSION

In general, up to 70% of the Zong-He-Ce-Yian items are categorized as local items. The result is in accordance with Yang (1996), Chang (1994), and Huang's (1994) findings, demonstrating that most of the deleted words in a cloze passage can be restored simply by using clausal or sentential clues. Zong-He-Ce-Yian items in SAET and DRET are designed to measure test-takers' discrete point skills. Chang (1994) suggests that to meet the function of the standard cloze procedure, which is used to measure learners' overall language ability, test writers should balance both local and global items.

To help test constructors develop items that are able to measure both learners' integrative and discrete-point skills, Wu (2002) proposes three item-generation approaches: the syntactic-, lexical, and discourse-oriented approaches. Items generated by either syntactic- or lexical-oriented approaches are considered as local items, and items generated by discourse-oriented approach is regarded as global items. Test-developers can take Wu's item generation approaches as references to balance the proportion of global and local items in local context.

As for test-takers' performances on both exams, it is found that test takers performed quite differently on the same type of question every year. In general, the best performed item type in SAET (category 2) is different from that in DRET (category 3). The finding suggests that item difficulty could not be controlled by question types. It seems that other factors such as topics, the difficulty level of vocabulary, and so forth. However, factor analysis on item difficulty is out of the scope of the present study; the issue can be further explored in future research.

REFERENCES

- Bachman, L. F. (1982). The trait structure of cloze test scores. *TESOL Quarterly*, 16(1), 61-70.
- Bachman, L. F. (1985). Performance on cloze with fixed-ratio and rational deletions. *TESOL Quarterly*, 19(3), 535-556.
- Chang, W. C. (1994). A brief analysis of the principles for constructing cloze items. *Newsletter for Teaching the Humanities and Social Studies*, 5(4), 69-80.
- Huang, T. S. (1994). *A qualitative analysis of the JCEE English tests*. Taipei: Crane.
- Jeng, H. H. (2001). A comparison of the English reading comprehension passages and items in the 1999 college entrance examinations of Hong Kong, Taiwan, and Mainland China.

- Concentric: Studies in English literature and linguistics, 27(2), 217-251.
- Jeng, H. H., Chen, L. H., Hadzima, A. M., Lin, P. Y., Martin, R., Yeh, H. N., Lin, C.C., Chang, B. Y., Tseng, T. Y., Sung, L. M., Ahrens, K., Chen, Y. S., Chen, C. R., Chang, H. H., Yu, H. Y., Liu, M. C., Alagata, A., & Wu H. C. (1999). *An experiment in designing English proficiency tests of two difficulty levels for the college entrance examination in Taiwan*. Taipei, CEEC.
- Jonz, J. (1990). Another turn in the conversation: What does cloze measure? *TESOL Quarterly*, 24(1), 61-81.
- Jonz, J. (1991). Cloze item types and second language comprehension. *Language Testing*, 8(1), 1-22.
- Taylor, W. L. (1953). Cloze procedure: A new tool for measuring readability. *Journalism Quarterly*, 30, 415-433.
- Wu, H. L. (2002). Rational cloze: Item-generation approaches and construct validity. *English Teaching and Learning*, 26(4), 85-106.
- Yang, T. H. (1996). Fundamental considerations in the test, with special reference to its use in EFL testing in Taiwan. *Sun Yat-sen Journal of Humanities*, 4, 57-77.

Holistic Language Learning through Genuine Learner Autonomy

Po-ying Chu (朱珀瑩)
Taipei National University of the Arts
pyrchu@gmail.com

This paper outlines an approach to learning that allows learners themselves to define their objectives, how these may be achieved, and the extent to which they have been successful. In this way, holistic learning is ensured because that is how learners naturally learn, when left to their own devices. It is common to hear teachers reflect on their teaching, but not much research has been done on the learning experience of students from a student's perspective. This paper looks at way of understanding students and helping them understand themselves through their own words and thinking.

Following a reflective worksheet based on the principles of Exploratory Practice (Allwright, 2005), English students write down their learning narrative before class, proceed to group discussion related to their individual learning, give feedback to each other, and clarify their learning goals. This combines actual language use (listening and speaking, reading and writing in English) with reflection, providing rich practice for language learning while also letting students' voices be heard and understood. The focus is not necessarily on problems and difficulties, but may be on students' puzzles about learning, or their often underestimated achievements.

I will consider examples of both the process and the results of student reflection using Exploratory Practice to show how exploratory practice can contribute simultaneously to language learning itself and to students' awareness, and even enjoyment, of the learning process.

INTRODUCTION

The main purpose of this paper is to use Exploratory Practice as an alternative approach to create learning opportunities for students inside and outside the classroom. The researcher and author has been offering a reflection-based course by following the principles of Exploratory Practice, EP (Allwright, 2005). Students in this one-semester elective course meet two hours a week. The course is taught in English and students are encouraged to use English too. In the first part of the lesson, students are requested to deal with some required reading materials in the course book, 'the study skills handbook'. In the second half of the class, students are asked to talk about their English learning experience, the discussion being based on the structure of an EP worksheet which guides students towards thinking and writing about some particular learning puzzle questions and narratives before class, and then retelling their learning stories in group during class. Students have many learning opportunities for reading, speaking, writing and listening before, during and after a lesson. Since 2007, the researcher has started using a course website to improve the quality of the course through giving sufficient information about each lesson, supplementary material for downloads, important links related to the course, group talk audio/video recordings, and students' electronic learning journals.

LITERATURE REVIEW

In its theoretical conception, this course, and therefore also this study of the course, follow the principles of Exploratory Practice (EP) (Allwright, 2005). EP helps both students

and teachers to understand their experience of learning and teaching languages, with the emphasis on understanding rather than on trying to solve problems. Understanding is prioritized because without adequate understanding, any action that is taken may be misguided, for example because the wrong question is being asked or because a false problem is being addressed. The focus of understanding can be positive as well as negative experiences, successes as well as failures. In fact, recalling success is just as important as looking at problems, if not more so. As Aoki and Yukio have said:

Teacher education should focus more on teachers' successes, small as they may be. A success, as well as a problem, can trigger reflection. Celebrating success could nurture realistic, constructive optimism among teachers. (Aoki & Yukio, 2003)

Likewise, in order to repeat successful experiences, students need to know why things went well in the past, and then they can try to repeat their individual success.

Learner autonomy has been broadly used in EFL, and some researchers (Benson, 2001, Wenden, 2002) promote strategy training as the main means of achieving learner autonomy as well as learner development. However, this paper takes a different approach, in which learner autonomy is seen as an ability students have already obtained. As Holiday (2003) suggests, autonomy is there to be discovered, rather than to be learnt. Aoki and Smith (1999) argue that autonomy should more properly be seen as being a willingness and/or ability to take charge of one's own learning, and ability which learners may already possess or not possess in varying degrees. Smith (2003) makes a distinction between 'weak' and 'strong' versions of pedagogy for learner autonomy. In his characterization of 'weak' versions of pedagogy for autonomy, students are seen as currently lacking a capacity for which they need to be trained. On the other hand, a 'strong' version of pedagogy for learner autonomy is based on the premise that students already are autonomous and capable of exercising this capacity, and simply need to be allowed to develop it.

As it shares the same assumption that students do not need training to be autonomous, this EP-based course actually matches Smith's 'strong' version of pedagogy for learner autonomy. Instead of training, what students need are learning opportunities. Learner opportunities are categorized by Allwright (1991:167) as either "encounter" opportunities or "practice" opportunities. "Encounter" opportunities are opportunities to meet whatever is to be learned, while "practice" opportunities are opportunities to do something with target material. Once learners themselves make good use of learner opportunities without being asked to, they are in the process of learner autonomy and learner development both in and out of class. In my opinion, learner development could be defined as the process of developing as learners. Learners willingly seek for or create learning opportunities for themselves and each other, consciously or unconsciously, whether inside the classroom or in real life situations. Learners' realization grows as their development increases and vice versa.

HOLISTIC LANGUAGE LEARNING THROUGH EP

How do I use EP to create learning opportunities?

In the whole-English course, students are given opportunities to rethink their goals in learning English, keep weekly learning journals, and regularly share their reflection with others in groups and also on the course website. Here is a sketch of a lesson in the course.

First half of a lesson	Second half of a lesson
Language warm-up activity: vocabulary dominoes, math game, jazz chants, treasure map etc.	EP Group Talk -Students are asked to form groups of 3-6 students and talk about their learning puzzles, narrative or ideas for understanding.
Course book context: looking at some practical ideas for general English or academic English	-They talk, listen and give feedback to each other. -Before the end of a lesson, students are asked to report to the class what puzzle questions they have discussed. This is done through either a brief oral report or a group poster.
Students usually have hands-on experience to have a taste of different ways of learning.	-Teacher gives comments on students' discussions.

STUDENTS' LEARNING OPPORTUNITIES FOR LANGUAGE & REFLECTION

The procedure and use of the EP worksheet

The EP worksheet is used to promote the following:

1. A puzzle question

Students are encouraged to think of an English learning puzzle question they are interested in understanding either from positive or negative learning experience.

Examples of Students' Puzzle Questions

Speaking	<ol style="list-style-type: none">1. Why have I become braver when I talk in front of people?2. Why is my pronunciation skill so bad?3. Why can't I show my opinion to my classmates?4. Why do I feel nervous when I talk to foreigners?5. Why can't I express my feelings in English?
Reading	<ol style="list-style-type: none">1. Why can't I read for a long time?2. Why do I feel confused when I want to choose a book to read?3. Why is my reading worse than others?4. Why can't I help myself reading something in a hurry?5. Why do I find magazines are more interesting and understandable than story books?
Writing	<ol style="list-style-type: none">1. Why can't I write a narrative well?2. Why can't I write an article clearly?3. Why can't I use good words for English writing?4. Why do I have no ideas when I write an article?5. Why do I always use simple vocabulary in my writing?
Listening	<ol style="list-style-type: none">1. Why is my listening ability bad?2. How can I increase my listening ability?3. Why can't I get general ideas during listening to the radio program?4. Why can't I understand movies or radio in English?5. Why can't I understand BBC news?
Vocabulary	<ol style="list-style-type: none">1. Why do I always forget English words and grammar?2. Why couldn't I use the new vocabulary I've memorized effectively when I was chatting in English?3. Why can't I learn vocabulary more efficiently?4. Why can't I have good spelling in English?5. How can I recite more vocabulary?

Grammar	<ol style="list-style-type: none">1. Why didn't my grammar progress?2. Why am I always confused about the English rule and also forget them easily?3. Why is my grammar worse than other skills?4. Why do I always make mistakes in speaking tests?5. Why can't I study grammar well?
Exams	<ol style="list-style-type: none">1. Why can't I pass GEPT?2. Why am I afraid of GEPT?3. Why do I have to take English proficiency test for improving my English?4. Why can't I get high grades in the reading section of TOFEL test?5. Why do I have to take exams to prove my English ability?
Other	<ol style="list-style-type: none">1. Why does my learning motivation usually depend on teachers?2. English could be improved greatly3. Why can't I spend much time studying English?4. Why do I always learn less things from English?5. Why can't I study effectively in English?6. Why can't I study English without an electronic dictionary?

2. A narrative

Students are asked to write a narrative to describe what and how their learning puzzle happens in real life situations. They write their narrative before class. Then, they are asked to tell other students their narrative in group. As Moon (2006:130) has stated:

“When a story is then told [i.e. after writing about it] there is a further opportunity for the teller to learn from the story. She may need to reformulate it if it is to be told orally (i.e. not directly from direct reading)”.

Therefore, students have very good learning opportunities to clarify what puzzles them in their head and in their narrative after they write about it and talk about it. Through this way of storytelling, students benefit not only from language practice but also from having good chances to clarify their puzzles. Here is an example from a student who appreciated writing a narrative.

Maybe sometime, I really don't know what I've learned or what I've thought about the topics, but when I started to move my pen and write something down, I just realized what kind of ideas I got. To tell the truth, I quite like this feeling, because it make me feel that I found myself!

In the following example of a student's puzzle and narrative, she felt happy when was clear about her learning goal.

Puzzle: Why do I need to study English now?

Narrative:

This question was brought by Ms. Chu's comment from my last exploratory practice worksheet. I think it's a good puzzle to think about. Everyone wants to have excellent and fluent English. BUT why do I need to learn English just now? Maybe the reason that I had no direction to improve my English is because I did not establish firmly or clear my goal. Now I am clarifying my goal, the reason I want to study English right now is because I want to study aboard. Yes, I want to study aboard after I graduate form TNUA.

3. Refining the puzzle

After students talk about their narrative in group, they then can decide whether they need to narrow down their puzzle question to make it clearer and close to what exactly puzzles

them. Many students benefit from this step without taking any action. For example, a student's original puzzle was 'Why is my reading so slow?', but after she discussed with her friends, she formulated a more specific question: "Why do I use an electronic dictionary so often when I read".

Examples of a puzzle

Why is my reading so slow?

Why can't I speak fluently to a foreigner?

Why is my pronunciation so bad?

Examples of a Refined Puzzle

Why do I use an electronic dictionary so often

Why don't I have interesting things to say?

Why don't I have confidence about my English?

4. Ideas for exploring the puzzle

Students are asked to think of ways of understanding why their puzzle question happens. Sometimes they have preliminary ideas from themselves, and they can also receive suggestions from their group partners or their teacher. Very often students talk about watching movies without subtitles to improve their listening.

How can I improve reading?

ideas for exploring the puzzle:
Go to a bookshop on Sundays

Why can't I practice English everyday?

ideas for exploring the puzzle:
Look for a language exchange partner
Listen to English songs

Why can't improve my English?

ideas for exploring the puzzle:
Talk to friends in English every Monday morning

5. Action ideas

This is an optional choice. In fact, this is not the aim of EP but it is students' choice if they decide to explore deeper in their learning by trying out some action ideas. As long as students have better understanding about their learning puzzle, they naturally move towards a different puzzle without using any particular action ideas to solve any particular problem.

Next is a complete example of the EP worksheet.

Exploratory Practice : elaborating , developing and exploring a puzzle .One way to do it

First idea :

A puzzle :

I would like to know how I can develop the confidence when I am in a group talk.

Exploring what the puzzle really is :

A narrative :

As we all know, we need enough confidence to learn well. Nevertheless, confidence is something I am really lack of. During group talk, I seldom express my opinions or give others my advice, for I do not really sure whether others would like to know. I do actually dislike my performance in both academics and oral. I do not know why but I feel bad about myself. I know I must pick up my confidence or I will never improve. And I will also be sorry to my group members.

Refining the puzzle :

My refined puzzle :

I want to pick up my confidence in a group talk.

Ideas for exploring the puzzle with the minimum of disruption and a maximum of understanding :

Ideas :

To develop confidence, I need to take chance creating a successful experience. Also, I need to see my performance in another aspect. Nothing is completely worthless, so do our any tiny advice.

What I will do first to explore my puzzle :

Plan :

I will try to speak out in group talks, being helpful to my friends. Therefore, I will stop dwelling on the negative criticism I have given to myself.

Clearly, students are given many opportunities for their language learning and language practice through an EP course. They constantly use English to think, rethink, write, talk, listen and read during the course. With a course website, students now can watch a lesson's video recording, listen to other students' group talk, read other friends' journals and give written feedback. They also receive immediate comments from teacher. I would like to use a student's journal to show her thinking and reflection in the following example.

"Learn from not only your failure but your success." This is the statement that the teacher has continuously reminded us. Since I have been instilled deeply that failure is the mother of success. It's quite striking and a brand new to me.

This new idea brought me to reconsider my attitude towards my "puzzle". I am always asking and worrying about why I "cannot". But the fact is, I never stop to think how I "can". Experiences talk. Successful experiences therefore tells you how you can do something. So, I tried to recall having been able to talk in English confidently. And I found that is the beginning of my English learning. Then the puzzle is not a puzzle anymore. Why? Because I know what's wrong with me.

Learning more and more in English, I also grow to know how to criticize. Overemphasizing on my worries, I spontaneously am only aware of my worries. For I tend to persuade myself that I am not oversensitive; to warn myself that the problem is actually there. So I care about grammar, about vocabulary, about accent, and so on. Nonetheless, when I learn to change the attitude, I see the difference. I don't blame myself when using wrong grammar. I told myself, "Angela, you'd been able to talk in English fluently and confidently before. So you can do now as well." I can do it I know, I had done it years ago, when I didn't know many rules about grammar; when I only wanted to express myself."

WHY ARE STUDENTS AUTONOMOUS?

Moon (2006) puts emphasis on understanding in learning. 'The task of the educational process is not just to promote awareness of prior conceptions, but to develop the understanding to form an appropriate basis for further learning' (Moon 2006:83). During this reflection-based course, students continuously have to decide what they want to investigate, why they want to learn, and how they feel comfortable with trying things out. For Little

(1997) this kind of conscious reflection on the learning process is a distinctive characteristic of autonomy. In addition, decision-making which is commonly mentioned in autonomy occurs all the time in the process of their learning. According to Boud (1988), the main characteristic of autonomy as an approach to learning is that students take some significant responsibility for their learning over and above responding to instruction. With regard to EP, students play an important role in their learning and investigation. Dickinson (1993) identifies five features associated with autonomous learners: (a) they can identify what has been taught; (b) they are able to formulate their own learning objectives; (c) they select and implement appropriate strategies; (d) they can monitor these for themselves; and (e) they know how to give up on strategies that are not working for them. Here is a table to show how EP presents students' autonomy.

EP	Dickinson's features about autonomous learners
-a puzzle	-formulating learning objectives
-a narrative	-identifying what has been taught and what has been puzzling them
-refining the puzzle	(-decision-making)
-exploring possible ideas	-select and implement appropriate strategies
-action ideas (if necessary)	-they monitor for themselves
	-deciding to give up on strategies that are not working for them

In short, it is the students who have to make up their mind about what learning strategies suit them and what do not. Moreover, students actually have *real* control over their own learning and freedom to establish their own creativity about learning. Miller's (1979) suggestion of generating creative ideas in writing resonates with the idea of students' reflection in EP. According to Miller, 'writing allows the mind to explore more freely since we do not run the risk of losing our previous thoughts.' With EP, it leads students to a better world of learning defined by themselves and discovered by themselves. This fits in with Naiman's view that 'language learning in general might be more effective if the students' ideas were used more frequently. Thus the learner would take an active part in deciding how he would like to learn.' (Naiman 1995:180).

CONCLUSION

This paper has discussed how EP can be used as an alternative way of helping students to be autonomous, taking an EP worksheet as example for demonstrating students' use of language to define their learning objectives, to describe what happens to their learning in real-life situations, to refine their learning target, to listen and give feedback to each other in group talk, to rethink and understand their learning. I also emphasized the importance of giving students learning opportunities instead of considering training as a way of *making* them autonomous. Students need to have opportunities of freedom to exercise and practice their autonomy. In Smith's (2003) 'weak' version of pedagogy for autonomy, learning arrangements tend to be determined by the teachers. Following such an approach, a learner might be told what to learn and how to learn. The 'strong' version of pedagogy for learner autonomy, on the other hand, holds that only if learners are given opportunities can they develop the ability to make sensible decisions about their own learning, and also take responsibility for the consequences of their decisions. EP matches with the 'strong' version of pedagogy for learner autonomy.

With EP, students are given both learning opportunities for language practice and understanding their learning and themselves at the same time. They feel less worried or frustrated and gradually they feel happier as language learners than before as they understand themselves more. Thus, they take English learning as a choice they make for themselves.

REFERENCES

- Allwright, D. (1991). The Characterization of Teaching and Learning Environment Problems and Perspectives. In Foreign Language Teaching Research in Cross-cultural Perspective. Amsterdam:John Benjamins, pp.161-174.
- Allwright, D. (2005). Developing principles for practitioner research: the case of Exploratory Practice. *Modern Language Journal* 89(3):353-366
- Aoki, N. with Hamakawa, Y. (2003) Asserting Our Culture: Teacher Autonomy from a Feminist Perspective. In David Palfreymen and Richard C. Smith, *Learner Autonomy across Cultures*.(pp240-253) New York :Palgrave Macillan.
- Aoki, N. and R.C. Smith (1999) Autonomy in cultural context: The case of Japan. In S. Cotterall and D. Crabbe (eds.) *Learner autonomy in language learning: Defining the field and effecting change*. Frankfurt am Main: Lang. 19-28.
- Benson, P. (2001). Teaching and Reseraching Autonomy in Language Learning. Essex: Pearson Education.
- Boud, D. (ed). (1988). Developing Student Autonomy in Learning (2nd edn). New York: Kogan Page.
- Dickinson, L. (1993). Talking shop: autonomous learning. *ELT Journal*47/4: 330-336.
- Holiday, A. (2003). Social Autonomy: Addressing the Dangers of Culturism in TESOL.In David Palfreymen and Richard C. Smith, *Learner Autonomy across Cultures*.(pp.110-128) New York :Palgrave Macillan.
- Little, D. (1997). Language awareness and the autonomous language learner. *Language Awareness*, 6(2/30:93-104.
- Miller, J. (1979). Keeping a psychological journal. *The Gifted Child Quarterly* 23:168-175.
- Moon, J. A. (2006). Learning journals. Oxford: Routledge.
- Naiman, G. and M. Frohlich, H,H. Stern, A. Todesco. (1995). The Good Language Learner. Clevedon: Multilingual Matters Ltd.
- Smith, C. R. (2003). Pedagogy for Autonomy as (Becoming-)Appropriate Methodology. In David Palfreymen and Richard C. Smith, *Learner Autonomy across Cultures*.(pp.129-146) New York :Palgrave Macillan.
- Wenden, A. (2002) Learner development in language learning. *Applied Linguistics*: 23/1:32-55.

Examining How Meaning, Form, and Context Interact in the Text of News Leads

Mei-hsia Dai (戴美霞)
Ta-Hwa Institute of Technology
daimeihsia@yahoo.com.tw

This study attempts to employ an integrative approach from a discourse perspective to explore the relationships between the syntactic structures of news leads and the cognitive level of foreground-background propositions. The foregrounding process involves information presented sequentially or the pragmatic notion of the main line of the information at a particular point in the *preceding* discourse. The backgrounding process involves the notions of presupposition or old information at a particular point in the *following* discourse in terms of thematic/topical importance. However, the dynamic nature of discourse grounding tends to shift from one particular frame to another. The data for this study were 50 news leads collected from two local English newspapers in Taiwan, with articles that were obtained from the Associated Press (20), Agence France-Presse (20), and Reuters (10). The structures of lead sentences provide evidence for the distinction between foregrounding and backgrounding operations. The study on the news leads shows that the following four types of clauses are more frequently used: relative clauses, subordinate clauses, *that*-clauses, and coordinate clauses. The characteristics of the use of the four types are: (1) Restrictive relative clauses will be used when the information of this lower clause is the main idea whereas nonrestrictive relative clauses will be used when the information of the higher main clause is the main idea. (2) The information of the main clause presents the most important theme when it is followed by a subordinate clause. (3) A *that*-clause carries the foregrounding information when it follows an initial authoritative main clause subject. (4) The information of the main clause takes the most important theme when it is followed by a participle derived from a coordinate clause.

INTRODUCTION

English newspapers usually consist of a wide range of up-to-date topical materials written in authentic English and provide EFL readers with a valuable resource for English teaching and learning. Some might consider newspapers to be more motivating and dynamic materials than published textbooks. For EFL learners, reading news articles is a challenge, and when EFL learners approach the task, they discover that news stories are usually composed of language particular to news reports which prevents them from comprehending the meaning as a whole, especially the headlines and the lead paragraphs. However, if they could receive instruction as to the structure of writing in this context, it may help with understanding and encourage students to read more news.

The study and analysis of actual language in use is the goal of discourse analysis. Halliday (1994: xiv) stated that the historical study of linguistics involved studying the morphology of language and was followed by the study of meaning of words at the sentence level. The ultimate goal of the analysis was to discover the meaning of the form of language. Such study is more linguistics oriented. According to Halliday's view, "a language is interpreted as a system of meanings, accompanied by forms through which the meanings can be expressed." His work then highlighted approaches to understanding the mechanism of how text is structured. The progression of text constructs a context for meaning. The relationship of such a linear progression between lines is particularly apparent in news articles. Stories start with a headline and a lead to introduce the framework of the story, subsequent details are

presented in descending order of importance from beginning to end.

Linguistic studies have shown that discourse features correlate with the Foreground (FG) - background (BG) structures, and therefore conflate the phenomenon of grounding with several other notions and levels of description. The notion of “grounding” consists of two separate psycho-communicative processes. The first process is anaphoric, which involves the grounding of a particular point in the discourse vis-a-vis the *preceding* discourse background in terms of shared knowledge. It traditionally correlates “background” with “presupposition”, “old information” or “topic”. The other process is cataphoric, which involves clues the speaker, gives the hearer at a particular point in the discourse to determine how to ground it vis-à-vis the *following* discourse in terms of thematic/topical importance. Previous studies induce syntactic structure on foreground and background clauses or sentences (Labov & Waletzky, 1967; Warvik, 1990), and cognitive level of foreground and background information (Grimes, 1975; Jones & Jones, 1979). However, of the news lead and its cognitive level there were few studies to combine the two, i.e. to investigate the syntactic structures on clauses and explore the cognitive level of foreground and background information. This paper therefore attempts to explore an integrative approach, i.e. to seek the relationships between the syntactic structures of news leads and its cognitive level of foreground-background proposition. More specifically, the paper will examine how meaning, form, and context interact in the text of a news lead.

LITERATURE REVIEW

The dynamic nature of discourse grounding

Traditional discourse studies view a text as foreground when the information presented is new and asserted, and background when the information is old, shared and presupposed. Such a view of discourse is static because it requires the researchers' external knowledge of discourse information to interpret the text. The text is treated as an object, whereas there is no relationship between the writer, text, and/or the receiver. However, the process of interactional discourse is dynamic. That there is a correlation between the notion of foreground and the thematic notion “information presented sequentially”, or the pragmatic notion “main line of the information” (Givon, 1987) has been proposed by several studies in the past. However, the notion of “background” and the notion of presupposition and old information often become problematic. The proposition of the backgrounding structure, such as a restrictive relative clause or adverbial clause, can be defined, but the problem lies in that it cannot correspond to the notion of “presupposition.” Givon (1987: 181) on the other hand suggests that based on constructed examples with minimal context, that adverbial clauses are “topical” or “pragmatically presupposed” only when they precede the main clause; but that they are “focus” or “asserted information” when they follow the main clause.

News-story Organization

According to van Dijk (1988: 15), typical news organization does not present topics in a continuous linear order, but rather according to a hierarchical ordination principle. The Summary (Headline and Lead) always comes in the first place and the Commentary is generally placed at the end of the news story. After the Summary come the Main Events and Background, which are related to the Summary. The summary, as a result, is foregrounded. The organization of the news discourse structure is not rigid, but follows a rather optional order.

Contrary to the expectation of linear time organization (chronological order of report) brought to the narrative, news events are organized in relation to their perceived importance to the theme, or ‘macroproposition’, of the story as represented in the Summary (van Dijk 1982: 180). Within the body paragraph of the news report, some information is foregrounded and some backgrounded. Main Event episodes, which are considered as primary importance

to the macroevent, are foregrounded and the following supporting details are backgrounded. Normally in many texts the writer places the prominence of a text in the foreground proposition, or use a foregrounded structure that expresses a lower grounding value. What is less prominent is usually a background proposition or a backgrounded structure that expresses a higher grounding value.

Studies on News

Discourse analysis of news articles has received some attention with regard to the function of pronouns in written newspaper articles, *that*-deletion, news commentary, advertisements, and the syntactic manifestation of affect in the lead sentence.

McCarthy and Carter (1994) used written newspaper and magazine data to show how *it*, *this* and *that* perform distinctly different function in the signaling of focus and topicality in the text. He argued that *it* signaled reference to continued, ongoing topics, *this* signaled the raising of a new or significant topical focus, and *that* had a variety of distancing or marginalizing functions.

Nagano (1983) investigated 38 different news commentary segments taken from Japanese television programs and found out that in orally presented television news commentary, the central message is announced at the beginning, presumably to make the audience understand what follows in the presentation. The announcement of the main message at the beginning serves two functions: One is to catch the audience's attention. The other is to reduce the listeners' cognitive overload. Khalil (2006) examines syntactic manifestation of affect in the lead sentence of Arabic news stories. It addresses the question of the pragmatic motivation for the occasional occurrence of spatiotemporal structures in text-initial position. However, Khalil does not explore the relationship of affective functions to syntactic structures which depend largely on interpretation – a process that requires knowledge of discourse structure as well as knowledge of various patterns of relationships, such as knowledge of text-type and specific foreground-background structure.

DATA TO BE STUDIED

Research has evidenced that textual structures express or signal various underlying meanings, opinions, and ideologies. The data for this study were 50 news leads collected from *Taipei Times* and *Taiwan News* in June, 2008, with articles that were obtained from AP(20), AFP(20), or Reuters(10). Most lead sentences consist of one or two clauses. The most frequently used types of clauses are the focus of this study: relative clauses, subordinate clauses, *that*-clauses and coordinate clauses. Adjective phrases that are derived from relative clauses such as “More than 1,000 rescue workers *including troops searched on yesterday for 13 people missing . . .*” or participles from coordinate clause such as “Kosovo's first constitution as an independent state came into force yesterday, *giving ethnic Albanians the right to executive power . . .*”, or adverbial phrases beginning “*because of . . .*” will be counted as one clause. They were considered a clause because each comprises a complete idea. All lead sentences start with a main clause. In addition to the main clause, relative clause occurs the most frequently (25), subordinate clause the second (21), followed by *that* clause (10) and coordinate clause (10). However, data also show that four lead sentences consist of only one single clause and 19 leads consist of two clauses. Most of their second clauses are source clause (6), subordinate clause (4), relative clause (5), and coordinate clause (4).

DATA ANALYSIS AND DISCUSSION

Foregrounding and backgrounding operations may occur in news lead sentences. The results of the 50 leads show that relative clauses were the most frequently used clauses as well as the most frequently appeared clause that follows the initial-main clause. Examples (1) illustrates the phenomenon of two typical relative clauses.

(1) US makers of steel pipe win against China in trade case

Washington--

(1a) US steel pipe manufacturers, (1b) who have been battling a surge in imports from China, won a major victory on Friday (1c) when the International Trade Commission cleared the way for the imposition of stiff penalty tariffs for the next five years.

(2) Japanese quake death toll rises to 10, but rescue search goes on

Tokyo--

(2a) The death toll in a powerful earthquake (2b) that hit northern Japan at the weekend rose to 10 yesterday (2c) as troops and rescue workers searched for survivors in the remote, mountainous area worst hit by the tremor.

The two typical relative clauses are: nonrestrictive and restrictive. The major distinction between the two depends on whether the antecedent is unique information or not. In lead sentence (1), the writer starts with the subject of the main clause (1a), embedded by a nonrestrictive relative clause, and followed by an adverbial clause. The headline and the lead sentence express typical foreground meaning about the most significant and long battling event that the manufacturers of US steel won a major victory. Contrary to this expectation, if the antecedent in the news lead is not unique information, the relative clause (2b) *that hit northern Japan at the weekend* is restrictive which serves to express the most significant meaning about the event to the antecedent.

It is well known that the complementizer *that* is omissible in English. The use of a *that* clause in the following example demonstrates different typical utilization in news leads.

(3) Organizers say Olympic torch to make stop in Tibet

(3a) Organizers of the Olympic torch relay said yesterday (3b) *that the flame would soon make a stop in Tibet*, (3c) but they declined to give an exact date for that leg of the trip, (3d) which has been shrouded in secrecy.

Unlike a final source clause, the grammatical subject of the initial main clause followed by a *that*- noun clause only functions to elicit the subject information for the subsequent clause, such as examples (3). The subject of the main clause usually exhibits specific or technical noun phrases or organizations which also demonstrate authoritative power, and to make readers to know *who* the real subject is. Such a main clause is usually followed by a *that*- noun clause which often carries the prominent message for the main subject.

Another typical syntactic structure in the lead is a main clause followed by a similar syntactic structure but it was changed into a participle form, such as the example (4).

(4) Kosovo moves to terminate U.N. rule as its first constitution comes into force

(4a) Kosovo's first constitution as an independent state came into force yesterday, (4b) giving ethnic Albanians the right to executive power held by the United Nations mission (4c) that has run the territory since 1999.

The writer uses a very emphatic syntactic structure: the lead sentence starts with a main clause. In the second part of the lead, though presenting a parallel structure with a coordinate conjunction, the writer uses a participle to present the sequential level of importance of information.

The results of the study indicate that subordinate clauses were the second highest used

clause in the lead sentence. Since subordinate conjunctions are various types and carry different functions and meanings. The following examples illustrate one common quality shared by subordinate clauses.

(5) Obama apologizes to Detroit women after snub at rally

(5a) US Democratic presidential candidate Barack Obama personally apologized on Thursday to two women (5b) after campaign volunteers barred them from appearing in a picture (5c) because of their head scarves.

Though subordinate clauses occupy the second biggest number of clauses following the main clause, they usually function and mean variously. However, they have a common role in news lead, that is, subordinate clauses follow a main clause and serve to give more detailed information to explain the main event.

CONCLUSION AND SUGGESTIONS

This article demonstrates the distinction between foreground-background structures and attempts to interpret the perceptual differences regarding prominence in the surface structural notion of proposition. The leads have a canonical grounding structure that takes propositions with descending importance. News discourse initiates with the Summary, foreground proposition, in the beginning of news reports, and the Main Event, History and Comments tend to support the foreground meaning. The evidence of the study show that lead structures correlate with the foregrounding information in the lead. Data analysis also shows that pragmatic constraints, such as prominence or relevance, may be affected by the way the foreground-background structure presented in the text. The interaction of grounding discourse is multilateral.

The structure of the news lead sentences exhibits the nature of discourse grounding. The meaning of foreground-background proposition derives not only from the syntactic surface structure but also from the cognitive structure as well as the knowledge and interpretation of the language users. The number of lead sentences analyzed in this study is limited to three news agents, more leads from more other newspapers may provide further evidence to verify the findings here or to modify or to expand the types of lead structures. Different editors may favor particular structures. However, from a pedagogical point of view, it is crucial to note that the findings of lead patterns are beneficial to teachers and learners. Teachers can become aware of the four types of the most frequently employed structural patterns and familiarize their learners with these patterns and the prominence of the grounding structures. The foreground-background propositions interplay and correlate with the patterns and seem to reveal a discourse pattern that should be taken into account in the way learners manage the topicality of news leads.

REFERENCES

- Givon, T. (1987). Beyond and Foreground and Background. In Tomlin (Ed.), *Coherence and Grounding in discourse* (pp. 41-58). Amsterdam: John Benjamins.
- Grimes, J. E. (1975). *The thread of discourse*. The Hague: Mouton.
- Halliday, M.A.K. (1994). *An Introduction to Functional Grammar*. London: Edward Arnold.
- Jones, L. B. & Jones, L. K. (1979). Multiple levels of information in discourse. In L. K. Jones (Ed.), *Discourse studies in Mesoamerican language, Vol. 1* (pp. 3—27). Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.
- Khalil, Esam, N. (2006). Communicating affect in news stories: The case of the lead sentence. *Text & Talk*, 26(3), 329-349.
- Labov, W. & Waletzky, J. (1967). Narrative analysis: Oral versions of personal experience. In

- J. Helm (Ed.), *Essays on the verbal and visual arts* (pp. 12-44). Seattle, WA: University of Washington Press.
- McCarthy, M. & Carter, R. (1994). Spoken Grammar: what is it and how can we teach it? *EL Journal*, 49 (3), 207-218.
- Nagano, M. (1983). Danwa ni okeru Jojutsu no Koozoo [The structure of predicate in discourse]. In *Danwa no Kenkyuu to Kyooiku* [Study of Discourse and Education], F. Minami and N. Tanaka (Eds.), vol, 63-89. Tokyo: Oookurashoo Insatsukyoku.
- van Dijk, Tuen A. (1982). Episodes as units of discourse analysis. In Deborah Tannen (Ed.), *Analyzing Discourse: Text and Talk*. pp. 177-196. Washington, DC: Georgetown University Press. *Georgetown University Round Table on Languages and Linguistics*, 1981.
- van Dijk, Tuen A. (1988). *News Analysis: Case Studies of International and National News in the Press*. Hillsdale, NJ: Lawrence Erlbaum.
- Warvik, B. (1990). On grounding in English narratives: A diachronic perspective. In S. Adamson, V. Law, N. Vincent, and S. Wright (Eds.), *Paper from the 5th international conference on English historical linguistics* (pp. 559-575). Amsterdam: John Benjamins.

新聞專業英語時體教學對新聞中譯英的影響

Da-hui Dong (董大暉)
Chang Jung Christian University
dongdahui@mail.cjcu.edu.tw

本研究運用前人研究英語時體分佈所提出之「話語假設」(Discourse Hypothesis)的理論，在大學部四年級新聞專業英語課中，教授學生用「前景」(foreground)、「背景」(background)二元法區分新聞段落，並總結出「前景」使用「過去時」，「背景」使用「現在時」之英語新聞「前景、背景時態框架」原則。本研究進而探討此新聞英語時體教學法，對學生中譯英新聞正確選擇時體的效果。我們對實驗組 27 位大學四年級新聞組學生和比較組 66 位大學四年級、9 位研究所一二年級學生進行一項教學實驗，結果證明在學習上述時體選擇法後，學習者可以成功地將主要依靠詞彙及語境表達時體之中文新聞，翻譯成以詞形變化為主表達時體的英文。

介紹

英語時態一直是英語習得上的一個難點。大量研究指出，母語非英語的人士(下稱 NNS)，特別是母語 L1 中沒有時體詞形變化的學習者(如母語中文的學習者)，常常出現時態的錯誤(Bardovi-Harlig, 1998; Bardovi-Harlig & Bergstrom, 1996; Coppieters, 1987; Guiora, 1983; Hinkel, 1992, 1997, 2004)。造成錯誤的原因有很多，母語的影響，對英語中時態並不是單純指示時間缺乏了解，對英語語篇中時體使用的約定缺乏了解等等，都會造成 NNSs 學習英語和使用英語上的困難。

在中譯英教學中，我們發現，學生在決定使用何種英語時態時，首先尋找原文中表示時間的標記(中文一般用時間副詞、上下文語境、體(aspect)等表示時間)。當原文沒有表示時間的標記時，學生就會不知所措，常選擇錯誤的時體。基於前人分析語篇中時間框架(temporal frame)的成果(Matthiessen, 1996; Smith, 2003)，本文提出，中文新聞語篇中段落的時間框架可以分為「前景」、「背景」兩類，中譯英過程中，在「前景」中應使用過去時(past)，「背景」中使用現在時(present)。透過教學實驗檢驗，這種方法能有效提升學生中譯英新聞過程中時態的正確率。

文獻探討

學者很早就注意到，要幫助 NNSs 克服英語時體上的錯誤，就必須從語篇著手(Dry, 1981, 1983, 1992; Hopper, 1979; Kumpf, 1984; Longacre, 1981)。他們觀察到，敘述文(narratives)基本上有兩個部分組成：「前景」(foreground)和「背景」(background)。「前景」部分所包含的事件，是敘述文的架構，推動語篇時間前行；而「背景」部分則並不敘述主要事件，而是提供支持性的素材，對「前景」部分的事件進行補充或評論(Dry, 1981, 1983, 1992; Hopper, 1979)。母語英語者在寫作敘述文時，時體使用遵守「前景」使用過去時，「背景」使用現在時的規則。不少研究發現，NNSs 寫作敘述文時，「前景」較多使用一般過去時，「背景」較多使用一般現在時，而不常使用現在時的其他體式(Andersen, 1991; Bardovi-Harlig, 1995, 1998, 2000; Bardovi-Harlig & Bergstrom, 1996)。針對母語中文 NNSs 的研究也有類似結果(Chen, 2005; 樊長榮和林海, 2002; 蔡韶亭,

2004)。然而，在其他語篇的寫作中，NNSs 使用時態雜亂無章的問題非常普遍，有必要針對性地施教(Hinkel, 1997)。Chen(2005)比較母語英語與母語中文的大學生英語作文的時態後發現，母語是中文的學生能夠掌握句子層面時態的用法，但沒有認識到語篇層面的時態用法上的約定要比句法上的用法優先，因而造成使用上的錯誤。

在研究如何詮釋英語不同類型語篇中時體時，Chafe (1972)發現英語語篇中存在「時間框架」(temporal frame)，有基礎時態，在同一個時間框架中時態應保持一致，除非出現新的時間標記。例如，在一個過去的框架中，一般過去時就是基礎時態，即使講到當前事件，也應該維持過去時。Dowty (1986)也持這種看法，他認為，語篇中一個句子的時間解讀，與前一句時間解讀應保持一致，除非出現確定的時間副詞，這也就是他所說的「語篇時間詮釋原則」(temporal discourse interpretation principle)。

對於敘述文，「時間框架」比較容易分辨，即是上述「前景」和「背景」中的過去時和現在時，相較之下，其他文類則稍微複雜些，這也可能就是為什麼一般研究都選擇敘述文的原因(Nakhimovsky, 1988)。按照Ferry (2006)對新聞報導的分析，硬性新聞語篇結構通常是「倒金字塔」式的，由「導言」和「正文」組成，在「導言」部分回答5W1H (when, what, who, where, why, how)問題，而「正文部分」除了需要補充導言部分沒有回答的問題之外，還需要對事件進行更細緻的描述，提供背景知識，而新聞報導中還有一塊稱為「核心內容」(nut graph)的部分，以提供背景訊息或說明事件影響點 (point of impact) 的方式，將新聞事件放入較大脈絡中，說明該新聞事件對讀者的重要性。

Matthiessen (1995)指出，新聞報導中這些區塊都有其基本時態，但沒有預設時態。我們在分析數篇英語硬性新聞報導之後發現，如果利用「前景」和「背景」的概念，也可以對硬性新聞時態進行分類。新聞導言部分通常是「前景」，其時態框架為「過去時」。如果一則新聞中的5W1H 在導言部分沒有交代完整，後續段落對5 W1H的補充也都是「前景」，其時態框架也是「過去時」。新聞中那些表達作者個人評論、看法或是對該則新聞背景介紹的段落，以及其「核心內容」可分類為「背景」，其時態框架為「現在時」。但如果有明顯的時間標記出現時，時態的選擇還是應該保持與時間標記一致。附錄 1 是一篇標準的「倒金字塔」式新聞報導，我們將其段落按照「前景」、「背景」進行了區分，其時態使用完全符合「前景」使用過去時，「背景」使用現在時的模式。

因此，本研究者在新聞中譯英教學中提出新聞語篇「前景、背景時態框架」的原則，在進行新聞中譯英翻譯時，應先決定中文新聞語篇中段落為「前景」或是「背景」，再按「前景」使用過去時，「背景」使用現在時這一原則使用時態。本研究旨在探討教授學生「前景、背景時態框架」原則，對於學生中譯英新聞時態正確率有何影響。

研究方法

受試者

本研究「實驗組」共 27 位長榮大學翻譯系四年級新聞學程學生，「比較組」共 66 位長榮大學翻譯系四年級非新聞學程學生及 9 位研究所二年級學生。該系目前共有新聞、科技、經貿三個學程，課程配檔都包括專業英語、口譯、筆譯、視譯、第二外語等語言、翻譯相關課程，區別在於每個學程所注重的專業領域不同，因而教學材料會有所不同。學生從大學三年級上學期開始選擇學程。新聞學程學生較其他學程學生接觸更多新聞方面的課程，如新聞英語，新聞編輯等。透過近兩年的學習，其對新聞的認識和掌握比其他學程的學生更加扎實。在試驗進行之前，新聞學程學生（實驗組）尚沒有學習過專業新聞中譯英。本研究亦對學生進行口頭調查，發現學生雖然有上過新聞英語和新聞翻譯等課程，但大多有關新聞學的一些基本理論、英語新聞閱讀、新聞英譯中等，

並沒有系統學習過新聞英語時態的選擇，其時態使用仍然嚴重依賴傳統英語文法教科書。研究所受試者經研究所考試錄取，來自不同學校，除兩位本校翻譯系畢業生外，其餘在大學並非翻譯專業，也無新聞專業。在研究所學習期間，並沒有像大學部一樣以學程分組，因此，他們翻譯新聞應與翻譯其他類型語篇相差不大。

實驗步驟具體說明

1. 篩選前測與後測之原文與原文時態框架分析；

我們過篩選，挑出兩篇中文新聞，皆為標準倒金字塔硬性新聞報道。接著我們確定了這兩篇新聞每段的時態框架及翻譯每個小句（clause）時正確的時態選擇。前測原文共產生 14 小句，後測共產生 26 小句。如下表所示：

表格 1：前測原文「山崩洪水重創北臺灣」

小句編號	段落	功能	時態框架	正確時態
1· 豪雨在二十四小時內累積了十六公分的雨量，	導言	描述事件	前景	過去時
2· 於臺北縣市部分地區造成淹水及山崩災情。	導言	描述事件	前景	過去時
3· 臺北縣以五股、三重地區積水將近一公尺最為嚴重。	正文	說明細節	前景	過去時
4· 山崩也迫使臺北市一隧道關閉。	正文	說明細節	前景	過去時
5· 目前無人員傷亡傳出，	正文	說明細節	前景	過去時
6· 損失情形尚無法估計。	正文	說明細節	前景	過去時
7· 縣委員會發言人 Harry Zhang 表示：	正文	引述評論	前景	過去時
8· 「我們今年的颱風防範系統已徹底瓦解，」	正文	引述評論	前景	現在時
9· 「上一次颱風災後重建工作才完成，就又再遭逢重創。」	正文	引述評論	前景	現在時
10· 這整個地區今夏度過數十年來最嚴重的颱風季節，	正文	說明脈絡	背景	現在時
11· 雖不如日本及韓國嚴重，	正文	說明脈絡	背景	現在時
12· 但是最近一次發生於兩週前的納莉颱風期間，	正文	說明脈絡	背景	現在時/ 過去時
13· 土石流淹沒了一整個村莊，	正文	說明脈絡	背景	現在時/ 過去時
14· 奪走近 20 條人命。	正文	說明脈絡	背景	現在時/ 過去時

表格 2：後測原文「美救 3 大車商 日批不公平競爭」

小句編號	段落	功能	時態框架	時態
1· 美國政府準備對三大汽車業者提供紓困，	導言	描述事件	前景	過去時
2· 日本汽車業表示對其他車廠並不公平，	導言	描述事件	前景	過去時

3·這種違反公平競爭原則的援手	導言	描述事件	前景	過去時
4·就是保護主義的抬頭。	導言	描述事件	前景	過去時
5·以經濟新聞報導為主的日本 12 號頻道「東京電視」，在其晚間 WBS 新聞節目中對美國政府準備救三大汽車廠表示憂心。	正文	說明細節	前景	過去時
6·東京電視台訪問在美國的豐田汽車與歐洲汽車界人士，	正文	說明細節	前景	過去時
7·他們都說	正文	引述評論	前景	過去時
8·「美國政府只救美國三大車廠，是不公平的競爭。」	正文	引述評論	前景	過去時
9·他們表示，	正文	引述評論	前景	過去時
10·現在汽車廠都已經跨越國界，	正文	引述評論	前景	過去時
11·外國車廠也都在美國設點生產汽車，	正文	引述評論	前景	過去時
12·如果美國政府要救，	正文	引述評論	前景	過去時
13·就該一視同仁，	正文	引述評論	前景	過去時
14 不能夠獨厚美國三大，	正文	引述評論	前景	過去時
15·這樣的話等於是保護主義的抬頭。	正文	引述評論	前景	過去時
16·1930 年代時的世界經濟恐慌，源於各國為了保護自我，	正文	說明脈絡	背景	現在時
17·而禁止他國貨品輸入，	正文	說明脈絡	背景	現在時
18·這個慘痛經驗，讓最近召開的 G20 與 APEC 會議中，都特別強調不能夠讓保護主義抬頭，	正文	說明脈絡	背景	現在時
19·否則將使全球陷於長期蕭條。	正文	說明脈絡	背景	現在時
20·日本對美國政府準備救三大的措施，首先表示異議，	正文	說明脈絡	背景	現在時
21·認為這就是變相的保護主義。	正文	說明脈絡	背景	現在時
22·日本認為	正文	說明脈絡	背景	現在時
23·美國汽車業早在 1970 年代的石油危機後，就已經經營困難。	正文	說明脈絡	背景	現在時
24·日本與歐洲的汽車業為因應石油危機，努力研究生產低消耗燃料的車種，	正文	說明脈絡	背景	現在時
25·但是美國的主力商品還是耗油的大型車，	正文	說明脈絡	背景	現在時
26·這種賣不出去的大車，當然是難以經營。	正文	說明脈絡	背景	現在時

2. 前測和後測

前測和後測之執行

前測答題要求受試者填寫個人資料、年級，學程，並明確要求受試者將原文翻譯成

英語新聞。時間為 100 分鐘，允許查字典，但不允許連接網路。後測也採用相同的程序。這兩組受試者都接受兩次評量，而且都間隔 10 週以上，兩組不同之處在於「實驗組」在 10 週中接受過「前景、背景時態框架」教學，「比較組」則否。我們記錄了所有受試者的前後二次評量結果。

3. 評量要點

在測驗前即告知受試者測驗時間是 100 分鐘，在規定時間內無法完成之測驗為無效。所有受試者均於規定時間內完成作答，無逾時之無效評量。評量試題不在教學內容中，以免受試者依賴背誦記憶方式學習。前、後兩次評量內容一樣，但間隔 10 週以上，並以無預警方式進行評量，目的在避免受試者因預習或複習可採用短期記憶答題。

4. 講授「前景、背景時態框架」

研究者對長榮大學大學部四年級新聞學程學生受試者，講授「前景、背景時態框架」。在講義中，研究者利用 Smith (2003) 所提出關於不同英語語篇時態的詮釋原則，「前景」、「背景」在新聞語篇中的應用，佐以 Matthiessen (1995) 有關文本區塊發展的基本時態選擇方式，以及 Ferry (2006) 對硬性新聞英文提出的段落分析方法，舉例說明英語新聞每個段落的時態選擇，同時證明在實際的閱讀和寫作應用中，「前景使用過去時」、「背景使用現在時」的原則可以準確解釋英語硬性新聞的時態使用規則。研究者在講授過程中使用一篇美聯社新聞“Japanese troops cross into Iraq from Kuwait”（附錄 1）作為例子，講述如何判定每段「前景」、「背景」分類，並比較其時態的使用規律。

5. 分析前測、後測譯文中時態的使用正確率

以表中所標註的小句正確時態為標準，對照受試者每小句所使用的時態，判定正誤，並計算出正確率。

6. 統計分析

統計分析工具：SPSS for Windows 軟體

- 獨立樣本 T 檢定
- 成對樣本 T 檢定
- ANOVA 檢定
- 平均數檢定

假設

本項研究透過教學前、後的評量，比較受試者的相關先備知識程度與實際學習成效所呈現之差異，並檢定經由教學活動，受試者是否能夠根據所學之「前景、背景時態框架」，正確判定中譯英新聞語篇不同段落時所應使用的時態。以下是相關的假設：
假設一、受試者未經過新聞英語「前景、背景時態框架」教學，在中譯英新聞語篇時態使用正確率評量上，雖然經過 10 週的在校翻譯學習，前、後兩次評量結果不會進步。
假設二、受試者英語程度、新聞專業學程與本研究之時態選擇正確率沒有關係。
假設三、受試者在新聞英語接受「前景、背景時態框架」教學後，無法學得相關時態知識，也無法在中譯英新聞語篇時，提升時態使用正確率。

研究結果與討論

假設一檢驗

我們使用「成對樣本 T 檢定」檢定由 66 位大學翻譯系四年級學生及 9 位研究所二年級學生所組成之「比較組」，在未經「前景、背景時態框架」教學情況下，相隔 10 週，其後測與前測的時態正確率是否有顯著進步。以下是該組學生兩次評量的統計結果。

成對樣本檢定								
	成對變數差異平均數	標準差	平均數標準誤差	差異的 99%信賴區間		t	自由度	Sig. (雙尾)
				下界	上界			
對照組前、後測差	0.477	4.648	0.701	-1.411	2.366	0.681	43	0.499

上表為無教學之「比較組」前、後測之「成對樣本 T 檢定」相關數據，其結果 P 值 $0.499 > 0.01$ 未達顯著標準，證明「比較組」前、後測之結果無差異。此結果符合本項假設，證明受試者經過 10 週的在校翻譯學習，在未經「前景、背景時態框架」教學情況下，前、後兩次評量結果沒有進步。這或許可以說明，傳統的新聞專業英語和翻譯教學方法對於學生中譯英新聞報導並沒有幫助。

假設二檢驗

本研究受試者包括三個類別：研究所、大學部翻譯系四年級新聞學程學生，以及大學部翻譯系四年級非新聞學程學生。本項假設是要檢驗這三類受試者在前測時，時態使用準確度是否存在統計意義上的差異。我們以「時態使用準確度」為依變量 (dependent variable)，以「受試者類別」為影響因素變量 (factor)，進行 ANOVA 檢驗。下表顯示，三類受試者之間，時態使用準確度雖然存在一些差異，但從統計學上看，這些差異並不顯著 (顯著性大於系統默認之 0.05 水準)。

「受試者類別」與「時態使用準確度」ANOVA(one-way)

Multiple Comparisons						
Dependent Variable: 時態使用準確度						
LSD						
(I) 受試者類別	(J) 「受試者類別	平均差 (I-J)	標準差	顯著性	差異的 95%信賴區間	
					下界	上界
新聞學程	研究所	-0.293	0.220	0.085	-1.74	1.14
	非新聞學程	-0.867	0.221	0.1916	-1.30	-0.43
研究所	新聞學程	0.293	0.220	0.085	-1.14	1.74
	非新聞學程	-0.574	0.239	0.3302	-1.05	-0.10
非新聞學程	新聞學程	0.867	0.221	0.1916	0.43	1.30
	研究所	0.574	0.239	0.3302	0.10	1.05

研究所與新聞學程受試者，P 值為 0.085，這說明，雖然研究所受試者之英文程度高於大學部受試者，但是在中譯英新聞時，其時態使用正確率上，並沒有比大學部新聞學程受試者高；同理，新聞學程受試者，雖然比非新聞學程受試者，接受更多有關新聞翻譯的教學內容，但在中譯英新聞時態使用正確率上，也並沒有比非新聞學程受試者高；因此，檢驗結果支持本假設。即，統計結果無法有力證實英文能力與是否新聞學程

與本研究之時態使用正確率有必然相關性。必須再檢驗教學後之後測結果才能確認。

假設三檢驗

本項假設目的在證明英語新聞「前景、背景時態框架」教學對受試者中譯英新聞語篇時，時態選擇正確率不會有明顯差異。檢定教學前、後其成效差異之統計結果如下：相依成對樣本檢定

成對樣本檢定								
	成對變數差異平均數	標準差	平均數標準誤差	差異的 99%信賴區間		t	自由度	Sig. (雙尾)
				下界	上界			
對照組前、後測差	-4.973	5.693	0.465	-6.186	-3.760	-10.698	149	0.000

27 位「教學組」受試者在參加教學前之前測後，經過英語新聞「前景、背景時態框架」等相關教學，相隔 10 週，再次接受測驗（後測）。上表顯示，教學前、後測之相依成對樣本 T 檢定的結果，P 值 $0.000 < 0.01$ 已達顯著標準，證明教學前、後測結果具顯著差異。此顯著差異的可能原因有兩種：時態正確率提升或正確率下降。因此再檢定教學前、後測之時態選擇正確數和正確率，以及相關統計數據如下：

平均數統計量

	正確數平均數	正確率平均數	標準差	平均數的標準誤差
教學前成績	6	42.8%	3.983	0.558
教學後成績	21	80.7%	4.338	0.436

受試者之時態選擇正確率，從教學前的 42.8%，進步為教學後的 80.7%，提升了一倍，證明教學結果是正成長，據此推翻本項假設。

結論

傳統英語時態的教學對於語篇中時態的時間框架缺乏重視，學生往往掌握了句子層面時態的用法，但卻未掌握語篇層面的用法，造成寫作或翻譯時時態的誤用。本研究透過實驗證明，傳統專業英語教學或者翻譯教學，對提高學生時態使用的正確率效果並不理想，學生習得英語時態的正確用法，勢必經過漫長的時間；本研究提出，在新聞英語的教學中，如果能夠用簡單明瞭的方式，教授學生判別語篇中每個段落的「前景」、「背景」功能以及所使用時態的約定，即可有效地提升學生中譯英新聞時時態使用的正確率。

參考文獻

- Andersen, R. W. (1991). Developmental sequences: The emergence of aspect marking in second language acquisition. In T. Huebner & C. A. Ferguson (Eds.), *Crosscurrents in Second Language Acquisition and Linguistic Theories* (pp. 305-324). Amsterdam: John Benjamins.
- Bardovi-Harlig, K. (1995). A Narrative Perspective on the Development of the Tense/Aspect System in Second Language Acquisition. *Studies in Second Language Acquisition*, 17(2), 263-291.
- Bardovi-Harlig, K. (1998). Narrative Structure and Lexical Aspect: Conspiring Factors in Second Language Acquisition of Tense-Aspect Morphology. *Studies in Second*

- Language Acquisition*, 20(4), 471-508.
- Bardovi-Harlig, K. (2000). *Tense and Aspect in Second Language Acquisition: Form, Meaning and Use*. Oxford: Blackwell.
- Bardovi-Harlig, K., & Bergstrom, A. (1996). Acquisition of Tense and Aspect in Second Language and Foreign Language Learning: Learner Narratives in ESL and FFL. *The Canadian Modern Language Review/La Revue canadienne des langues vivantes*, 52(2), 308-330.
- Chafe, W. L. (1972). Discourse structure and human knowledge. In J. B. Carroll & R. O. Freedle (Eds.), *Language Comprehension and the Acquisition of Knowledge* (pp. 41-69). Washington D. C.: V.H. Winston.
- Chen, Y.-c. (2005). Chinese Speakers' Use of Tense and Aspect in an English Past-time Context. *English Teaching and Learning*, 29(4), 87-105.
- Coppieters, R. (1987). Competence differences between native and near-native speakers. *Language* 63, 544-573.
- Dowty, D. R. (1986). The effects of aspectual class on the temporal structure of discourse: semantics or pragmatics? *Linguistics and philosophy*, 9(1), 37-61.
- Dry, H. (1981). Sentence aspect and the movement of narrative time. *Text*, 1(3), 233-240.
- Dry, H. (1983). The movement of narrative time. *The Journal of Literary Semantics*, 12(2), 19-54.
- Dry, H. (1992). Foregrounding: an assessment. In S. J. J. Hwang & W. R. Merrifield (Eds.), *Language in context: essays for Robert E. Longacre* (pp. 435-450). Arlington, TX: The Summer Institute of Linguistics and the University of Texas at Arlington.
- Ferry, T. (2006). *Your Guide to English News Language*. Taipei: Cosmos Culture.
- Guiora, A. Z. (1983). The dialectic of language acquisition. *Language Learning*, 33, 3-12.
- Hinkel, E. (1992). L2 Tense and Time Reference. *TESOL Quarterly*, 26(3), 557-572.
- Hinkel, E. (1997). The Past Tense and Temporal Verb Meanings in a Contextual Frame. *TESOL Quarterly*, 31(2), 289-313.
- Hinkel, E. (2004). Tense, Aspect and the Passive Voice in L1 and L2 Academic Texts. *Language Teaching Research*, 8(1), 5-29.
- Hopper, P. J. (1979). Aspect and Foreground in Discourse. In T. Givon (Ed.), *Syntax and Semantics: Discourse and Syntax* (pp. 213-241). New York: Academic Press.
- Kumpf, L. (1984). Temporal systems and universality in interlanguage: A case study. In F. Eckman, L. Bell & D. Nelson (Eds.), *Universals of second language acquisition* (pp. 132-143). Rowley, MA: Newbury House.
- Longacre, R. E. (1981). A spectrum and profile approach to discourse analysis. *Text*, 1(4), 337-359.
- Matthiessen, C. (1995). *Lexicogrammatical Cartography: English Systems*. Tokyo: International Language Science Publishers.
- Matthiessen, C. (1996). Tense in English seen through Systemic-Functional Theory. In M. Berry, C. Butler, R. Fawcett & H. G. W. (Eds.), *Meaning and Form: Systemic Functional Interpretation* (pp. 431-498). Norwood: Ablex.
- Nakhimovsky, A. (1988). Aspect, Aspectual Class, and the Temporal Structure of Narrative. *Computational Linguistics*, 14(2), 29-43.
- Smith, C. S. (2003). *Modes of discourse*. Cambridge: Cambridge University Press Cambridge.
- 樊長榮和林海。(2002)。中國學生英語時體習得中的兩大「誤區」。34(6)，頁414-420。
- 蔡金亭。(2004)。情狀體和語篇結構對英語過渡語中一般過去時標記的影響。36(1)，頁49-55。

附錄

附錄 1：教學範例

Japanese troops cross into Iraq from Kuwait

Updated Sun. Feb. 8 2004 12:05 PM ET (Associated Press)

1. KUWAIT CITY -- A contingent of Japanese troops crossed Sunday into Iraq from Kuwait as part of the country's first deployment to a combat zone since 1945. 「前景」
2. The troops, most of them engineers, are part of a deployment slated to total about 800 troops in a humanitarian mission in southern Iraq. Another 200 are to remain in Kuwait. 「背景」
3. The troops who crossed the border Sunday traveled in a 25-vehicle convoy from the U.S. Camp Virginia in the Kuwaiti. The convoy included armored vehicles, trucks, personnel carriers, and an ambulance. 「前景」
4. The number of the troops in the convoy was not disclosed. 「前景」
5. Japan has already dispatched advance teams and three C-130 cargo planes to the region. The troops will purify water and carry out other humanitarian tasks in Iraq. 「背景」
6. The deployment to Iraq is Japan's largest and riskiest since World War II. It has faced strong opposition at home amid fears for their safety with the insurgency in Iraq. 「背景」
7. In Tokyo on Sunday, Japan's defense chief, Shigeru Ishiba, said the government wouldn't order its forces out of Iraq until they have completed their humanitarian mission, even if Japanese soldiers are killed in attacks. 「背景」
8. The mission is the first time Japanese troops have gone to a combat zone since 1945. While Tokyo has sent soldiers on peacekeeping missions in the past, they have only gone to areas where fighting has subsided. 「背景」
9. Japan's constitution, adopted in 1947 during the U.S. postwar occupation of the country, renounces the use of military force. 「背景」
10. The Japanese troops in Iraq will be armed with pistols, rifles, machine guns and anti-tank guns to use in self defense only. 「背景」
11. The decision by the government of Prime Minister Junichiro Koizumi to deploy the troops despite widespread opposition reflects a shift in the government's attitude since the 1991 Gulf War, when Tokyo shouldered a portion of the financial burden but sent no soldiers. 「背景」
12. Koizumi and his allies in the ruling Liberal Democratic Party have been eager to contribute Japanese troops to back the United States, its most important ally. 「背景」

A Needs Analysis of Culinary Arts Majors' EFL Learning in Taiwan's Context

Hsiao-i Hou (侯曉憶), Min-yu Li (李敏瑜)

National Kaohsiung Hospitality College
Chang Jung University

hsiao@mail.nkh.edu.tw
michelle@mail.nkhc.edu.tw

This study reports on an investigation into Taiwan's culinary arts majors' English needs for their advanced ESP courses. None of the recent ESP/EAP research has focused on this group of students and their needs in ESL/EFL learning. The design of this study has been largely influenced by the works of Hutchinson and Waters (1987), Bhatia (1993), and Swales (2004, 1990) commenting that English for specific purposes (ESP) curriculum development is guided by learner's target needs in the target situation. Through questionnaire surveys with culinary arts majors including sophomores, and seniors with intern experience, and interviews with instructor and chef, the multi-faceted investigation has led to a much deeper understanding of the English communication needs placed on Taiwan's culinary arts majors and the industry. It is believed that the results help to reconsider current ESP communication course content offered at Taiwan's vocational colleges and universities. Recommendations for ESP curriculum designed and teaching methodology in Taiwan's vocational programs are discussed.

INTRODUCTION

There is no denying the dominance of English in today's world and as a lingua franca of the international business scene. In recent years, employers in Taiwan also have advocated that college graduates should be proficient in English especially in their professional fields, so that they can function effectively in the workplace. A recent nationwide survey conducted in Taiwan regarding college student English learning performance shows that 73% of undergraduates lack confidence in their English ability (MOE, 2004). There has been an impetus to integrate English for Specific Purposes (ESP) program with undergraduate curriculum course design, and movement in this direction has occurred. However, designers and teachers of courses for ESP are not particularly familiar with the communications needs of certain industries (Long, 2005). Students or adults learners are increasingly dissatisfied with those teaching materials or methodology developed or adopted in the classroom. To solve those problems, a global level needs analysis was carried out in this study. The focus of this study was one of Taiwan's unique industries, the culinary arts group. To understand the English communication needs for this particular group, a triangulation procedure by comparing of results from student questionnaires, and faculty and employer interviews was conducted.

RATIONALE FOR THE STUDY

This study focuses on the culinary industry in Taiwan because it is a major promotional industry and has been a major contributor to its tourism industry development. According to the 2007 World Economic Forum (WEF), Taiwan's Travel and Tourism Competitiveness

Index (TTCI) was ranked as the 4th in Asian countries (in Ministry of Economic Affairs, 2007). In 2006, Taiwan government has implemented several policies and strategies to promote Taiwan's tourism to all over the world, and it attracted over 1,510,000 international tourists. Comparing with 2005, the increasing rate of tourists was over 9%. In addition, among those tourists, 62.21% of them came to Taiwan for the Taiwanese cuisine. The culinary industry in Taiwan, therefore, has become a key element in promoting Taiwan's tourism and international fame.

Another important reason to focus the study subjects as the culinary arts majors is because of the increasing enrollment numbers in vocational institutes in recent years. According to MOE's data (2008), there are about five higher institutes enrolling students in culinary arts or in baking and pastry arts majors. Among those institutes, one is in national sector and four are in private sector. There are about 600 freshmen enrolling in the field of culinary arts every year in Taiwan (Table 1). Until now, none of the ESP research has focused on this particular group of students, especially their needs assessment. The demand for English teaching in culinary field is an example of English for specific purposes (Hutchinson & Waters, 1987; Robinson, 1991). By examining this group of students' needs in order to specify language learning objectives, this study sought to develop an English program that would facilitate language learning in general and satisfy the specific needs of the culinary profession.

Table 1. *Number of Students Enrolled in Culinary Arts Departments at Vocational Institutions in Year 2009*

	Public Sector	Private Sector	Total
Chinese Culinary Arts	41	100	141
Western Culinary Arts	41	100	141
Baking and Pastry Arts	41	55	96
Culinary Arts	0	234	234
Total			612

Another significant impetus for this study is from previous research results (Bhatia, 1993; Charles, 1996; Forey, 2004; Long, 2005; Swales, 2004; 1990; Swales & Feak, 2004) showing that there is a gap existing between the workplace and in-class resources. It seemed that business people, who are familiar with workplace materials in their professional lives, should have the chance to report what they need. In the context of an EFL classroom, some teachers are often involved in using business texts as exemplar texts and possibly many teachers themselves may have only limited exposure to the business world. There is very little research available focusing on ESP curriculum development (Swales, 2004). None can provide a precise understanding of the nature of workplace English communication for graduate employees, especially in Taiwan's settings. Thus it is important to obtain a far deeper understanding of the day-to-day activities of the real business world. Such a need is commented on by Hutchinson and Waters (1987) when they state that the "great majority of ESP teachers have not been trained in ESP teaching. They need, therefore, to orientate themselves to a new environment" (p.157).

This study is designed based on Swales (1990) and Bhatia's (1993) ideas of genre analyses. According to Swales (1990; 2004) and Bhatia (1993), in order to understand the nature of texts, it is important to study their use as instruments of communication, the roles texts play, and the way texts are related to their uses and users in special interest settings. As Swales stated (1990), textual knowledge is insufficient for a full account of genre; to find the rationale it is necessary to go beyond the text by methods such as interviewing, participation and protocol analysis. This is exactly what this study aimed at, examining the

features of the actual environment where business messages are exchanged. The concept of genre, as defined by Swales (1990; 2004), is certainly very useful in the study of business discourse. According to him, the communicative purpose is the principal property of a genre and “other properties, such as form, structure and audience expectations operate to identify the extent to which an exemplar is prototypical of a particular genre” (Swales, 1990, p.52).

The English language needs of culinary arts students in Taiwan were assessed in order to increase teaching and learning effectiveness. An initial global level needs analysis was conducted in order to obtain up-to-date information on the types of English communication required in this industry. It is hoped that this study can also provide useful information to help ESP teachers and program designers in Taiwan.

DESIGN OF STUDY

The design of this needs analysis has been largely influenced by the work of Hutchinson and Waters (1987). English for specific purposes (ESP) curriculum development is guided by learner’s target needs, defined by Hutchinson and Waters (1987) as what the learner needs to do in the target situation. The target needs are further broken down into the necessities, wants and lacks of the learners and these needs were identified at the global level (Tarone & Yule, 1989). Questionnaires were developed and used to seek information on students' professional ability, their perception of English in their future career, perceived language skill needs and problems, the activities needed in their English course, suggestions for development of course content, their view of the usefulness of the culinary English programs they took in their previous semesters, and their demographic information. The questionnaires were written in Chinese, piloted, and modified according to feedback from 6 respondents: 4 students and 2 faculty members from a private university and a national college.

For the faculty and chef participant group, with similar content as the student group, a semi-structured interview instrument was developed and conducted. The interview was used to supplement the information obtained via the questionnaire in order to achieve greater reliability and validity. The use of these multiple sources helped to corroborate the results of the data collected.

Subjects of the Study

The subjects involved in this study were 225 college students who majored in western culinary arts, Chinese culinary arts, and baking and pastry arts from a national institute in southern Taiwan. Among them, 126 (56%) were freshmen and 99(44%) were seniors. In addition, 44% were male and 56% were female.

The faculty member participated in this study was forty-four years old and has taught at this college for eight years. He obtained his PhD degree from a university in Australia and had his master’s degree in F & B Management from a state university in US. He had worked as a chef and captain at different types of restaurants in US over 3 years. An interview with the instructor was conducted in Mandarin to gather details of the communication profiles and to obtain his views about the students particular English communication needs. For the chef participant, he was fifty-one years old and has been working in the industry over 10 years. He has worked at his current position as a sous chef at a four-star hotel in Southern Taiwan for six years. An interview with the chef was conducted in Mandarin to gather details of the communication profiles at this industry and to obtain his views about the culinary arts majored students’ particular English communication needs.

Data Analysis

The data were computer-analyzed using an SPSS 12.0 program. Student responses were divided into two groups based on their year: sophomores (SO) in one group and seniors (SE) in the other. Based on the school curriculum design, all the senior participants in this study have finished their one-year internship training in the industry during their junior year (July 2007 – June 2008). To compare the differences between one group of students with internship and the other group without internship, it's believed that the results of the study could obtain up-to-date and broader information on the types of English communication required in this industry. Descriptive analyses, Chi-square, t-tests, and ANOVA analyses were conducted.

RESULTS

The students' demographic data showed that the average age for the participants was 21 years old. 41% of them have obtained at least one type of English proficiency licenses, including TOEIC, TOEFL, or GEPT. Among them, 37% have been to a foreign country and the average grade for their previous English classes was 74.4.

Importance of English for Studies and Future Careers

Most students believed English is important for their current studies, future career development and advanced studies. 93.2% of the respondents said English was very important or important (SO: 95.2%; SE: 90.7%) for their current studies at school. There was no significant difference between these two groups regarding this viewpoint. In terms of perceptions of the importance of English for their future careers, 97.3% (SO: 97.5% ; SE: 96.9%) said English was very important or important with no significant difference between these two groups. In terms of perceptions of the importance of English for their future advanced studies, 94.2% (SO: 92.7% ; SE: 95.9%) said English was very important or important with no significant difference between these two groups.

Participants' Professional Ability

When the participants were asked to evaluate their ability in the four basic English skills, both groups showed less confidence in their writing and speaking skills. In addition, they also showed much more confidence in their reading skill than in listening skill. This fact can be noticed in Table 2. There is a significant difference between these two groups of students regarding their self-evaluation of the reading ability. The senior students have less confidence in their reading skills than the sophomore group. The overall patterns were consistent with the current MOE's (2004) national survey showing that college students in Taiwan perform better in English reading skill, than in speaking and listening skills.

Table 2. Participants' Self Evaluation of Their English Ability

	Poor		Fair		Good		Excellent		χ^2
	SO	SE	SO	SE	SO	SE	SO	SE	
Reading	42.4%	63.6%	40.0%	30.3%	15.2%	5.1%	1.6%	1.0%	*
Writing	64.0	76.8	27.2	20.2	7.2	2.0	0.8	1.0	
Listening	49.6	59.6	36.8	34.3	10.4	5.1	2.4	1.0	
Speaking	56.0	65.7	34.4	28.3	7.2	5.1	1.6	1.0	

English Language Curriculum, Course Content, and Materials

In response to the topics which should be included in the ESP course, the sophomore reported that English related to food ingredients and foods should be the most needed topics to be included in ESP course. For the senior group, they indicated that topics related to food ingredients and culinary arts should be the most needed ones to include in the course content. Between these two groups, there was no significant difference regarding their viewpoint on

this issue. (Table 3).

Table 3. *Topics to be Included in the ESP Courses*

	SO		SE		P
	M	SD	M	SD	
Daily life	2.15	0.459	2.07	0.479	
Cultural knowledge	2.00	0.622	1.99	0.505	
Management	2.03	0.567	2.01	0.544	
Food and Beverage	2.16	0.482	2.09	0.497	
Food ingredients	2.24	0.465	2.20	0.515	
Foods	2.27	0.464	2.19	0.488	
Food cultures	2.14	0.546	2.07	0.558	
International etiquettes	2.03	0.538	2.11	0.513	
Beverages	2.14	0.534	2.10	0.544	
Culinary arts	2.21	0.480	2.20	0.515	
Travel	2.11	0.585	2.12	0.500	
Hotel	2.07	0.570	2.13	0.508	
Tourism	2.11	0.540	2.12	0.540	

*For each topic, in ranking, 1 is strongly disagree, and 4 is strongly agree. So that the higher the mean, the higher of the agreement.

When captured students' perceptions about which of the skills should be included in the ESP class, there is a significant difference observed between these two groups of students. This difference can be found in skill related to business writing (Table 4). Senior students produce a higher mean than the sophomores in showing that business writing should be included in their ESP class. For both groups, they all show different patterns of viewpoints on this issue. Sophomore students indicated that public speaking and language test taking skills were rated highest. For senior students, they indicated that interview skills, study skills and language test taking skills were rated highest.

Table 4. *Skills to be Included in the ESP Courses*

	SO		SE		P
	M	SD	M	SD	
Public speaking	2.15	0.459	2.07	0.479	
Study skills	2.05	0.566	2.10	0.547	
Reading professional articles	2.07	0.526	2.08	0.531	
Telephoning skills	1.93	0.611	2.01	0.489	
Business writing	1.76	0.616	1.95	0.525	*
Filling out forms	1.90	0.607	2.02	0.574	
Language test taking skills	2.13	0.595	2.10	0.681	
Interview skills	2.08	0.517	2.20	0.591	

*For each skill, in ranking, 1 is strongly disagree, and 4 is strongly agree. So that the higher the mean, the higher of the agreement.

Regarding the participants' view of the usefulness of the ESP programs they took in their previous semesters, there were significant differences observed in instructors' teaching

methods, the materials, and the improvement of listening and speaking skills (Table 5). For the senior group, they indicated lower means regarding the usefulness of the teaching materials and the improvement of their listening and speaking skills in the previous ESP course than the sophomore group. For teacher's instructional methods, the senior group indicated a higher mean in showing that their teachers used traditional lecturing methods in their ESP courses.

Table 5. *Usefulness of the Previous Culinary Arts English Programs*

	SO		SE		P
	M	SD	M	SD	
My teacher used traditional lecturing methods.	1.77	0.556	1.92	0.550	*
Pairing or grouping activity is useful for my English learning.	1.91	0.568	1.77	0.655	
My teacher encouraged active learning.	2.06	0.453	1.95	0.598	
The credits for the ESP courses are enough.	1.51	0.738	1.41	0.655	
The materials are very practical.	1.70	0.686	1.51	0.721	*
The ESP course improved my listening skill.	1.80	0.650	1.58	0.687	*
The ESP course improved my speaking skill.	1.77	0.582	1.57	0.658	*
The ESP course improved my reading skill.	1.73	0.613	1.61	0.698	
The ESP course improved my writing skill.	1.62	0.658	1.53	0.721	

*For each statement, in ranking, 1 is strongly disagree, and 4 is strongly agree. So that the higher the mean, the higher of the agreement.

Interview with the Instructor

Findings from the interview with the instructor provided valuable input to the needs analysis. On the whole, the instructor indicated that the culinary arts majors' overall English abilities are not as good as other peers because they only focused on their culinary skills training during their high school years. English was not a major focus when they were in high school. In addition, in their college entrance exam, English is one of the required exam subjects, but it is not the key subject influencing their enrollment. He also indicated that he was not satisfied with current materials and textbook he used in the ESP classes for the culinary arts majors. He mentioned that the textbooks in the market are not very practical and authentic. Therefore, he had to spend a lot of time providing supplementary materials for the students. Another major problem for the current ESP courses is the size of the class. In his classes, there were over 50 students in the language class. When asked about the professional English skills the culinary arts majors needed to improve, his view was the same as most of the students. He thought that oral communication skills should be enhanced because this is most Taiwanese students' weakness. In the future, if the students work in those chain hotels in northern Taiwan, they will probably have to work with foreign chefs. Under this circumstance, oral skill will be very important. He did not express concern about grammatical errors as long as they did not seriously affect the learner's comprehension.

Interview with the Chef

Findings from the interview with the sous chef provided another perspective to this study. On the whole, the chef indicated that he was not satisfied with his employees' English communication skills, especially when require to perform more demanding tasks which involved a better command of English. When asked about the professional English skills the employees needed to improve, his view was the same as most of the study subjects. He suggested that the culinary arts ESP courses should emphasize on students' oral and listening communication skills training. In addition, for their reading practice, the students should learn how to read recipes in English. He also expressed concern about professional genre that students should acquire as long as they want to work in the field in the future.

CONCLUSION

The multi-faceted investigation has led to a much deeper awareness of the communication demands placed on Taiwan's culinary arts majors. It is believed that the results will help to reconsider current ESP course content offered at Taiwan's undergraduate level.

Based on the study findings, the culinary arts students have lower confidence in their English writing and speaking skills. This groups of students need to be encouraged to learn extensive use of English. According to the instructor and chef interviews, students should be alerted to the very definite likelihood that they will need to communicate in English with co-workers, many of whom will not, like themselves, be native speakers of the language. In addition, they should learn how to read or write recipes. Therefore, writing and speaking practices should be enhanced in their ESP courses.

The seniors with one-year internship experience in industry provided valuable input for this study. From the seniors' perceptions, this study showed that topics and vocabulary related to the culinary professions, including various food topic and ingredients, should be taught in context. Certain skills should also be taught including interview skills, language test taking skills and study skills. It is recommended that techniques such as pairing or group work be used in the ESP course to increase students' proficiency in the oral/aural skills. In addition, authentic materials such as magazine, advertisements or recipes should be used to encourage reading and writing practices.

The detailed investigation into the needs of English learning for the culinary arts group can be enable the course designers and developers of teaching and learning materials to provide more specifically focused English courses. While this study has concentrated on communication demands it is believed that students in many other profession-related majors, such as hair dresser or fashion designer would have similar needs, and these are areas that would benefit from further study. It is hoped that the findings described above will be value to course designers and teachers working with personnel who need to communicate in a similar environment.

REFERENCES

- Bacha, N. & Bahous, R. (2008). Contrasting views of business students' writing needs in an EFL environment. *The ESP Journal*, 27, 74-93.
- Bhatia, V. (1993). *Analysing genre: language use in professional settings*. NY: Longman.
- Bilbow, G. (1991). *Business writing for Hong Kong*. Hong Kong: Longman.
- Charles, M. (1996). Business communication: Interdependence between discourse and the business relationship. *The ESP Journal*, 15, 19-36.
- Department of Technology and Vocational Education, Ministry of Education. (2008). The Online Expo of Vocational Education. From: <http://techexpo.moe.edu.tw/>
- Department of Investment Service, Ministry of Economic Affairs. (2007). Taiwan's TCCI. From: <http://investintaiwan.nat.gov.tw/zh-tw/news/200705/2007052101.html>

- Department of Statistics, Ministry of Economic Affairs. (2003). The 2003 Industry and Commerce Statistic Reports. From: http://2k3dmz2.moea.gov.tw/GNWEB/news/news21.aspx?menu=2&id=1&item=02_01
- Forey, G. (2004). Workplace texts: do they mean the same for teachers and business people? *The ESP Journal*, 23, 447-469.
- Hoggatt, J. (2003). Communication needs of business. *Business Education Forum*, 58, 26-33.
- Hutchinson, T. & Waters, A. (1987). *English for specific purposes*. Cambridge: Cambridge University Press.
- Long, M. (2005). *Second Language Needs Analysis*. Cambridge: Cambridge University Press.
- Ministry of Education in Taiwan. (September, 2004). *The evaluation report of implementing English education in universities in Taiwan*. (Issue No. 362). Taipei, Taiwan.
- Robinson, P. (1991). *ESP Today: A practitioner's guide*. New York: Prentice Hall.
- Schutz, N. W. & Derwing, B. (1987). Needs assessment in ESP. In R. Mackay & J. Palmer (Ed.), *Languages for specific purposes program design and evaluation*. MA: Newbury House Publishers. Pp. 40-45.
- Swales, J. (2004). *Research genres: explorations and applications*. UK: Cambridge University Press.
- Swales, J. & Feak, C. (2004). *Academic writing for graduate students : essential tasks and skills*. Ann Arbor : University of Michigan Press.
- Swales, J. (1990). *Genre analysis*. Cambridge: Cambridge University Press.
- Tarone, E. & Yule, G. (1989). *Focus on the language learner: Approached to identifying and meeting the needs of second language learners*. Oxford: Oxford University Press.

中英網路溝通對比分析及其教學意義

Chia-ling Hsieh (謝佳玲), Rong-yu Yu (余榕毓)

國立台灣師範大學

國立台北科技大學

clhsieh@ntnu.edu.tw

s3540303@ntut.edu.tw

在資訊科技的蓬勃發展下，網際網路已成為搜尋與交流知識的新式平台。過去文獻從社會、心理、媒體等角度探討網路溝通的本質，至今仍缺乏跨語言的對比研究。本文以台灣與美國「奇摩知識家」電腦網路版的資訊交流現象為例，分析中文與英文網路溝通策略的異同。結果顯示中英網路媒體皆發展出不同於口語對話的言語特徵，形式簡潔且功能透明，但使用者尋求訊息的方式仍反映中英個別的禮貌文化。中文傾向軟硬兼施，以感謝、讚美、自謙等軟策略拉抬對方地位；以命令、求救、抱怨等硬策略激勵對方行動。英文偏好言簡意賅，較少加綴其他具補強效果的言語行為，且常以穢語表達遭遇困難的不滿或挫敗情緒，達到動之以情的目的。結論顯示現代英語教學者需正視網路的力量，引導學生瞭解中英電子語言與文化的共性與特性，方能在跨文化的資訊交流中應對得宜，順利達成溝通目標。

簡介

隨著資訊科技的快速發展，電腦多媒體應用與網路通訊技術全球普及，全世界的網路使用者已突破十億人口 (World Internet Usage Stats 2008)，網際網路儼然成為現代社會新興且盛行的溝通媒介，以電子郵件、電子佈告欄、電子報、討論區、留言版、即時通訊等各種不同的形式讓人類進行同步或非同步的交流。¹⁸ 網際網路除了建構出一個比傳統溝通媒介容許更多交際機會的虛擬社會之外，日常生活中常見的人際互動行為也普遍出現於網路溝通中，如資訊交換、教育學習、商業交易等。透過網路平台尋求知識或技能上的協助也是網路溝通中最常見的語言行為之一，也是語言學習者在教室外可能經常接觸甚或嘗試的溝通行為。對於面對面提出請求協助的語言行為，過去許多中外文獻曾針對其策略類型與禮貌功能提出詳盡的分析與見解，然而語言學界尚未探究請求協助的語言行為在數位資訊交流環境下所展現的風貌，以及不同語言是否偏好不同的線上禮貌策略。在現代網路媒介普遍流行的情況下，這些問題對於成日沈浸於網路世界的語言學習者更為重要，唯有瞭解不同語言在網路界面上請求協助的慣常模式，學習者才能在跨文化的網路溝通中應對得宜，順利獲得協助。

因此，本研究的目的在於觀察中文與英文電腦使用者透過網路請求協助的語言慣例與行為特色，一方面探討網路溝通媒體對人際互動方式的影響，另一方面瞭解中文與英文網路溝通文化的異同。為使分析結果更具代表性，研究的中英文語料皆取自在台灣與美國皆擁有眾多訪客的著名網站，即台灣的「YAHOO! 奇摩知識+」(網址 <http://tw.knowledge.yahoo.com/>) 與美國的「YAHOO! Answers」(網址 <http://answers.yahoo.com/>)。透過分析這兩個網站討論區中網友彼此請求協助的真實語料，本研究比較中英網路語言偏好的線上求助策略，希冀研究結果能讓中英文語言教學者與學習者認識網路溝通的風格，並意識到網路文化對當代語言使用的影響。

¹⁸ 全世界的網路使用者人口請參考 <http://www.internetworldstats.com/stats.htm>。

文獻回顧 求助行為

本研究擬探討的線上請求協助屬於「請求」(request)行為的一種,亦即請求者要求被請求者執行某項行為以滿足請求者的欲求(Austin 1962),因此可視為一種侵犯他人自由與展現請求者權力的行為,屬於「威脅顏面的行為」(face-threatening act)之一(Brown & Levinson 1987)。由於「禮貌」(politeness)的產生是為了降低人際互動過程可能產生的摩擦(Fraser 1975:64),因此在正常情況下的人際互動中,請求者會避免請求行為對他人顏面所造成的危害,因而會運用「正向禮貌策略」(positive politeness strategy)或「負向禮貌策略」(negative politeness strategy)降低對被請求者顏面的傷害,前者用來積極地維護聽話者的「正向顏面」(positive face),也就是個人希望保有正面形象的需求;後者則用來消極地滿足聽話者的「負向顏面」(negative face),也就是個人希望保有自由意志的需求(Brown & Levinson 1987:65-68)。

西方文獻在英文的口語請求模式上已累積相當豐碩的研究結果。早期的 Clark 與 Schunk (1980) 揭示請求的禮貌程度與「損益理論」(cost and benefit theory)有關,若能讓被請求者感覺受益越多則顯得越禮貌;而情境與字面意義的交互作用是決定損益的關鍵。Blum-Kulka & Olshtain (1984) 則指出幾乎所有的語言都會採用慣例型的間接策略 (conventionally indirect strategies)。Craig 等 (1986) 開始注意到請求行為使用大量且多樣的禮貌策略,最常見的是 Brown 與 Levinson (1987) 的正向禮貌與負向禮貌策略,呈現複雜的混用狀況,必須仰賴情境才能分辨其適用性,而禮貌策略的解讀也會遇到不確定的因素,例如,間接式的請求可能被視為約定俗成的表達方式,但也可能流於拐彎抹角,徒增認知的負擔。Blum-Kulka (1987) 繼而研究請求行為中「間接」(indirectness) 與禮貌的關係,指出過去文獻認為間接可增加選擇性,減少強迫的意味,所以代表禮貌,然而事實上間接請求需要較多推理過程,可能解讀為較不禮貌,因此禮貌是基於語用功能清晰以及未作強制要求這兩者之間的平衡。而且不同文化容許的語用清晰程度不同,所以對禮貌的界定也相異;不過人類普遍傾向盡量縮小推論過程,減少認知負擔,藉此達到禮貌的目的。Blum-Kulka (1989:18) 進一步將請求策略區分為直接策略 (direct strategies)、慣例型間接策略 (conventionally indirect strategies) 與非慣例型間接策略 (non-conventionally indirect strategies),而英文最常使用的是慣例型間接策略 (conventionally indirect strategies)。¹⁹ Fukushima (1996) 也強調禮貌在不同文化中有不同的體現,例如英國人較重視人際距離,日本人較重視團結認同,因此日文比英文更常使用直接請求,且較常以直述句表明願望;在支援對策的類型上兩個語言也偏好不同類型,英文最常採用提出「根據」(grounder),日文最常採用「減少負擔」(Imposition Minimizer)。Rinnert 與 Kobayashi (1999) 則發現美國人對禮貌的感知與請求的正式級別相關,但不如日本人明顯,而請求的語意內容對美國人的影響比日本人來得大。無論是在下位對上位或上位對下位的請求中,日本人的請求方式都比美國人正式。

中文亦不乏請求行為與禮貌本質方面的研究成果。中國人的禮貌行為建構在能分辨與察覺在社交場合中展現合乎禮儀的行為,而中文的「禮」(ceremony; courtesy;

¹⁹ 此三種策略涵蓋不同次類。直接策略包括語氣導出 (mood derivable)、行動動詞 (performative)、遮言式行動動詞 (hedged performative)、義務陳述 (obligation statement) 以及需求陳述 (want statement); 慣例型間接策略 (conventionally indirect strategies) 則可細分為建議式套語 (suggestory formulae) 與探詢式預備 (query preparatory); 而非慣例型間接策略 (non-conventionally indirect strategies) 包括強暗示 (strong hint) 與弱暗示 (mild hint), 詳參 Blum-Kulka (1989) 的討論。

etiquette)與「貌」(appearance)明確要求個人在公眾與私下場合的行為合宜(Lee-Wong 1999:24-25)。在請求行為的文獻中，Lee-Wong (1994) 歸納出中文的請求都包含「主體行為」(head act)，有時附帶「支援對策」(supportive move)，最常見的中心行為是祈使句、疑問句、「需求陳述句」(want/need statement) 以及「推論陳述句」(presumptive statement) 等直接的請求策略。其中祈使句的使用率最高，而且經常包含「動作動詞」(action verb)，據此清楚呈現命題內容，表達說話者的態度真誠，並常附加許多禮貌標語。在支援對策方面，中文偏好以稱呼或禮貌用語的「內部修飾」(internal modification) 來降低請求的直接程度。Lee-Wong (1994) 也指出請求的策略類型受到年齡與地位的社會因素以及要求的合理程度影響，中華文化特別著重社會地位與社會距離的價值觀會規範中國人的言行舉止，而對於真誠度與團結性的重視也會影響祈使句的運用。Zhang (1995) 也研究中文的間接請求與禮貌關係，發現間接請求是透過外部修飾 (external modification) 表示。Hong (1996) 贊同請求策略深受談話雙方社會權力、社會距離以及請求內容的影響，這一點反映在支援對策的使用、請求的長度以及稱呼的選擇上。Ma (1996) 研究中國和美國學生的網路聊天行為則發現中國學生的語言使用比美國學生間接。Yeung (1997) 檢視社會距離、相對地位與請求帶來的負擔此三項變因對中文與英文商用書信語言的影響，發現商用語體會影響上述三項變數的預測能力，顯示 Brown 與 Levinson (1987) 的理論無法充分反映中文的情況。Hong (1998) 則對比中文與德文的請求，結果顯示中文的請求禮貌度比德文更容易受地位與距離變數的影響。

綜觀以上關於請求行為的中外文獻，目前語言學界的研究成果以面對面口說請求為核心，也有少數針對書面請求所進行的分析，而網路的求助行為則尚待探討。針對口說與書面的請求，文獻的共識包括以下幾點。首先，不同語言對禮貌概念的詮釋不盡相同，因而也接受或容許不同方式的行為體現。其次，在請求的表達上，文獻普遍認為西方語言（如英文、德文等）比東方語言（如中文、日文等）直接，但部分文獻指出中文也使用如祈使句式這種相當直接的策略，而且中文的請求更容易受到地位與距離因素的牽制。在以上這些文獻的觀點基礎下，本研究考察中文與英文的網路求助行為是否展現不同於口語請求的特點，希望藉由中英對比的結果讓語言學習者瞭解兩個網路文化的差異。

網路文化

「電腦中介溝通」(Computer-mediated Communication; CMC) 是指以電腦網路為傳播媒介的溝通行為，溝通的主角是人類，而電腦網路僅為媒介，也就是經由電腦設備為媒介在人際間所發生的互動行為 (Herring 1996)。電腦中介溝通原是一種電子書面溝通的形式 (Barnes 2003)，具備同步與非同步溝通的特性。而透過線上討論區所形成的「電子社群」(electronic community) 或「虛擬社群」(virtual community) 可將相同嗜好的網路使用者連結在一起 (Furlong 1989; Van Gelder 1990)。先前關於電腦中介溝通研究多集中在電子郵件、部落格、聊天室、論壇等不同溝通平台的語言特色 (如 Baron 1984; Turkle 1995; Collot & Belmore 1996; Werry 1996; Yates 1996; Parks & Roberts 1998; Wallace 1999; Beibwenger 2000; Crystal 2001; Ziegler 2002; Baron 2003; Lewis 2003; Herring 2004; Marcoccia 2004)，許多文獻認同電腦中介溝通融合口說與書寫的語言特質，且使用者傾向減少打字數，注重增進溝通的速度與效率，並運用富含創意與意涵的文字傳達意義 (Herring 2001)。

盧諭緯 (1997) 以台灣學術網路中最具社群性的電子佈告欄 (Bulletin Board System; BBS) 為觀察對象，歸納出幾點重要的語言現象。其中之一是網路溝通以文字

為基礎，但參與者是根據語音優先的原則進行溝通，社群成員基於共知此原則而得以順利交流。此外，語氣詞（如“嗯嗯”、“嗚嗚”、“哈哈”等）與表情符號（如動作表情類如 *^_*、聲音語氣類如 :0、心理狀態類如 >_<、個人形象類如 :-8）的使用是增進溝通的線索，藉以建立良好的社會互動，進而促成社群的形成。稱呼的選擇也反映參與者尚未建立清楚的角色關係，且仍無法脫離既有的社會地位，呈現虛擬社會與真實社會的糾葛。這些語言現象顯示，一個社群的形成，其成員不僅在語言使用上具有相近的風格，同時溝通的本質也需側重成員態度與經驗的分享。

Baron (1998) 針對電子郵件的溝通歸納四點特色。首先，電子郵件比傳統寫作不正式，裡面常出現隨興用語，排版較不謹慎，使用者也常會擴大彼此的親密程度。其次，在缺乏面對面聽覺或視覺接觸的情況下，使用者感覺較不受限制，可發展出平等的對話關係。此外，用電腦比用傳統的紙筆更容易流露私人資料，促進私密性的對談。最後，電子郵件使用者常會加上情緒符號，所以也允許情緒的表現。因此，Baron (1998) 的研究顯示電子郵件呈現許多口語語言的特性。Flanagin 與 Metzger (2001) 透過問卷調查電腦中介溝通如何滿足個人需求，發現受訪者普遍使用傳統媒介（如面對面溝通、電話、書籍、雜誌、報紙、電視）與網路媒介完成「資訊擷取」(information retrieval)、「資訊提供」(information giving)、「交談潛力」(conversation capability) 三種目標。網路能像電話與電子郵件等「間接人際科技」(mediated interpersonal technology) 一樣具有對話的特色，也可像報紙、電視、書籍雜誌等「大眾傳播媒體」(mass media channel) 般具有資訊擷取與資訊提供的功能。可見網路與傳統溝通媒介並無太大差異，反而因其「內容聚合性」(content convergence)，亦即傳統由報紙或電視播報的資訊內容現在亦可由網路呈現，導致媒介的分類不再只依照媒介固有的屬性，而應衡量使用者在選擇某一媒介時所考慮的媒介功能。Flanagin 與 Metzger (2001) 也指出，不論使用何種媒介，人類最大的需求皆為資訊擷取，而網路就是現代人達成此目的之主要工具。

總結來說，文獻已歸納出網路溝通比口語溝通更著重傳訊效率，形式也更自由，且中文與英文的網路語言皆有其特色，可視為形成獨特的體裁。但學者也強調，網路媒介除了應根據其原有的特性在屬性上進行分類之外，更應該兼顧社會功能的屬性。本研究擬觀察上述網路媒介的特性如何影響中英文網路使用者在討論區中請求協助時採用的語言形式。

研究方法

語料來源

本研究的中英文語料分別取自台灣的「YAHOO! 奇摩知識+」(亦稱為台灣「YAHOO! 奇摩知識 plus」)(網址 <http://tw.knowledge.yahoo.com>) 與美國的「YAHOO! Answers」(網址 <http://answers.yahoo.com/>) 網站中關於電腦網路（美國稱為 Computer & Internet）知識領域的討論區。這兩個網站都是 YAHOO! 網際網路服務公司所成立的開放性數位資訊交流平台，提供免費且社群導向的資訊服務。使用者必須先利用該網站設計的介面申請帳號，註冊成為正式會員，而後才能依照自己有興趣的知識領域在不同的平台上提出問題或給予解答。根據美國 YAHOO! Answers 於首頁上提供的相關連結，目前全世界至少有二十六個地區已經建立了 YAHOO! 知識+，如韓國、日本、台灣、中國、香港、美國、法國、德國、西班牙、義大利、印度、巴西等地。其中最早的 YAHOO! 知識+ 是韓國於 2003 年 6 月架設，繼日本於 2004 年 5 月設立之後，台灣與美國也分別於 2004 年 11 月 18 日與 2005 年 12 月 7 日成立。各地區使用的網站名稱略有差異，如台灣的「YAHOO! 奇摩知識+」、美國的「YAHOO! Answers」、中國大陸的「雅虎知識堂」、法

國的「YAHOO! Questions réponses」、德國的「YAHOO! Clever」等。²⁰除了知識+之外，YAHOO! 公司也提供一系列的資訊、消費、娛樂、社群等各類網際網路服務，包括搜索引擎、電子信箱、即時通訊、購物拍賣、新聞報導、部落格、家族、交友等。²¹

參考國際知名網路流量分析公司（如 Alexa、Comscore、Netcraft 等）調查統計的數據，YAHOO! 是全球網際網路訪客最多的網站之一，擁有 4 億以上獨立 IP 用戶的訪客，平均每日接受瀏覽的網頁高達 34 億，成為台灣與美國最受歡迎的網站之一。²² 根據台灣「創市際市場研究顧問公司」於 2003 年 6 月進行的「ARO 網路測量研究」數據，在台灣地區搜尋引擎類型的網站之服務整體表現中，YAHOO! 奇摩搜尋服務的使用者有 686 萬人，遠高於 Google、MSN、PChome、蕃薯藤、Openfind 等。²³ 而崛起於 2004 年的台灣 YAHOO! 奇摩知識+ 也是台灣網路使用者最常使用的資訊搜集平台之一，根據該網站於 2009 年 3 月在首頁末尾顯示的資料，知識平台上已解決的問題數已超過一千三百萬筆，問題數則超過四十五萬筆，表示平均而言每個問題能獲得 2.8 個解答，且發問完成率高達 97%。在美國方面，ZDNet 公司於 2008 年 10 月針對美國搜尋引擎所進行的市佔率排名結果顯示，YAHOO! 的市佔率達 17.74%，排名第二。²⁴ Hitwise 網路測量公司也指出，截至 2008 年 3 月為止，網路使用者利用知識網站發問的比例較 2007 年同期成長 118%，其中美國 YAHOO! Answers 佔所有的知識搜尋平台的第一名，高達 74.05%。²⁵

根據 Pew Internet & American Life Project 公司於 2008 年 10 月份進行的調查，全球男性電腦網路使用者的人數由 2002 年的 33% 增加至 2008 年的 53%，女性電腦網路使用者的人數亦由 2002 年的 25% 增加至 2008 年的 45%；其中 2002 年約有 33% 的使用者運用搜尋引擎找尋資料，到 2008 年其比例已增加至 49%。²⁶ 台灣創市際公司於 2007 年 4 月調查指出，八成四的台灣網路使用者會利用網路蒐集資訊，對於資訊的追求已展現化被動為主動的趨勢。在搜尋工具方面，網友常使用的平台依序是搜尋引擎、資訊收集平台（如 YAHOO! 奇摩知識+、維基百科等）、專業內容網站（如手機王、UrCosme 等）、BBS 或論壇以及主題部落格。在搜尋資訊的類別方面，以娛樂類（46.42%）、電腦及通訊類（43.26%）與消費資訊類（36.04%）為大宗。²⁷

在本研究觀察的中英兩個網站中，台灣 YAHOO! 奇摩知識+ 區分十一個知識類別，分別為電腦網路、生活資訊、手機通訊、休閒嗜好、視聽娛樂、運動體育、社會人文、商業金融、教育學習、科學醫療以及煩惱心事。截至 2009 年 3 月，YAHOO! 奇摩知識+ 的首頁顯示其電腦網路版的問答筆數超過兩百三十萬件，在所有知識類別中排名高居第一。²⁸ 而美國的 YAHOO! Answers 則涵蓋更多知識領域的大類，包括藝術與人文（Arts &

²⁰ 詳參維基百科 (<http://zh.wikipedia.org/wiki/>) 中關於「Yahoo!奇摩知識+」的介紹。

²¹ 台灣地區詳見 <http://tw.yahoo.com/>；美國地區詳見 <http://www.yahoo.com/>。

²² 詳參維基百科 (<http://zh.wikipedia.org/wiki/>) 中關於「Yahoo!」或「雅虎」的介紹，Alexa 的調查結果參見 <http://www.alexa.com/>，Comscore 的調查結果參見 <http://www.comscore.com/>，Netcraft 的調查結果參見 <http://news.netcraft.com/>。

²³ ARO 指標（Access Rating Online）（ARO Index = Reach x Web Session Reach x Duration per Visit）係指評估網站總體表現的指標，類似傳統媒體的收視率調查，整體表現考慮因素包含使用者到達率（Reach）、上網次數到達率（Web Session Reach）以及單次造訪停留時間（Duration per Visit）在所有網站中的相對值。詳參創市際市場研究顧問公司網站的說明 http://www.insightxplorer.com/news/news_08_11.html。

²⁴ 詳參 ZDNet 公司網站 <http://blogs.zdnet.com/ITFacts/?cat=27&paged=2>。

²⁵ 詳參 Hitwise 公司網站 <http://www.hitwise.com/>。

²⁶ 資料來源 <http://blogs.zdnet.com/ITFacts/?cat=27&paged=2>，詳見 Pew Internet & American Life Project 公司網站 <http://www.pewinternet.org/>。

²⁷ 詳參創市際市場研究顧問公司的調查報告 http://www.insightxplorer.com/news/news_04_04_07.html。

²⁸ 資料來源請參見台灣 Yahoo!奇摩知識+ 網站 <http://tw.knowledge.yahoo.com/dir/dir>。

Humanities)、美容與流行 (Beauty & Style)、商業與財政 (Business & Finance)、汽車與交通運輸 (Cars & Transportation)、電腦與網路 (Computers & Internet)、消費性電器產品 (Consumer Electronics)、外食 (Dining Out)、教育與參考文獻 (Education & Reference)、娛樂與音樂 (Entertainment & Music)、環境 (Environment)、家庭與人際關係 (Family & Relationships)、飲食 (Food & Drink)、遊戲與消遣 (Games & Recreation)、保健 (Health)、居家與園藝 (Home & Garden)、地方商業 (Local Businesses)、新聞與事件 (News & Events)、寵物 (Pets)、政治與政府 (Politics & Government)、懷孕與養育 (Pregnancy & Parenting)、科學與數學 (Science & Mathematics)、社會科學 (Social Science)、社會與文化 (Society & Culture)、運動 (Sports)、旅遊 (Travel)、YAHOO! 產品 (Yahoo! Products) 幾項，每項大類再區分許多次類。其中電腦與網路也是使用率較高的討論區之一。因此，本研究選擇較具代表性的電腦網路討論區作為中英對比的對象。

以上提供之數據資料顯示，在網路上主動搜尋資訊已是台灣和美國網路使用者的普遍行為。由於台灣 YAHOO! 奇摩知識+ 與美國 YAHOO! Answers 兩個網站皆移植同樣的設計概念，提問與解答的運作模式也相同，再加上兩個網站都設有電腦網路討論區，使用者人數眾多，資料更新速度快。基於以上極高的相似度，本研究以這兩個網站的電腦網路討論區為研究對象，比較中英文在網路溝通中採用的求助策略，以瞭解中西方語言與文化的異同。

蒐集方法

本研究於 2008 年先後從台灣 YAHOO! 奇摩知識+ 與美國 YAHOO! Answers 取得使用者在電腦網路討論區求助或提出問題的語料，台灣 YAHOO! 奇摩知識+ 的中文語料是從 2008 年 1 月 26 日起往前回溯至 2008 年 1 月 23 日，共計 807 筆合格語料，美國 YAHOO! Answers 的英文語料是從 2008 年 10 月 31 日起往前回溯至 2008 年 8 月 15 日，共計 807 筆合格語料。以上即為本研究中英對比的語料來源。所謂合格語料，是指使用者透過這兩個網站平台對電腦網路相關知識尋求幫助或解答的語料，討論區中摻雜少數主題不符的語料不納入本研究的範圍。以下舉兩例，中文如**錯誤！找不到參照來源。**) 所示，英文如**錯誤！找不到參照來源。**) 所示。²⁹

- (1) 不知道日本 TOKUYO 的按摩椅在日本當地的風評好嗎，因為最近開了一家三洋健康生活館，裡面賣很多 TOKUYO 的按摩椅，功能相當讚，店員說日本熱賣喔，不知哪一位大大可以提供這一個品牌的按摩椅好嗎，哪裡有在做優惠活動
- (2) What song is in the very beginning of this clip?
I know, its used in a lot of slow motion things, but I don't know the song name
<http://www.youtube.com/watch?v=jgTkXAvtn...>

對比結果

完成語料蒐集後，本研究針對台灣 YAHOO! 奇摩知識+ 與美國 YAHOO! Answers 兩個網站的使用者在討論區中求助時所採用的語言行為進行策略分類與分析，並統計各個策略的出現頻率，進行中英對比。結果顯示中文與英文偏好不同的求助策略，以下分四個小節說明對比結果，前兩小節分述中英文的網路求助特色，後兩小節呈現中文與英文的共同點與相異點。

中文網路求助特色

台灣 YAHOO! 奇摩知識+ 的語料顯示，由於求助的行為會干預被求助者的行為，威脅其顏面，所以中文網路使用者傾向採用軟硬兼施的方式求助，一方面以軟策略拉抬可

²⁹ 本文保留語料之原始形式，未修改文字、標點符號與句型之誤用。

能回應的網友們的地位，另一方面以硬策略激勵網友們付出行動。軟策略包括在求助的主體行為中使用禮貌用語，以及在求助主體行為之外使用其他的輔助或支援行為，依照使用率高低排列涵蓋感謝、讚美、自謙、回報、自清、道歉等次類；硬策略則包括以祈使句式所形成的求助主體行為，或是以求救、抱怨等方式所出現的輔助或支援行為。以下按照使用頻率由高至低說明這些策略的特色。

(一) 禮貌用語：中文網路求助者常藉由具備禮貌功能的用語修補顏面，加強求助主體行為的禮貌程度，常見的禮貌用語為“請”、“請問”、“請教”、“麻煩”以及“拜託”等。其中“請”字單獨使用的比例最高，佔中文所有求助語料總筆數的 50.1% (404/807)，如**錯誤！找不到參照來源。**)；“請問”與“請教”合計佔所有求助語料總筆數的 60.0% (484/807)，如**錯誤！找不到參照來源。**)；而“拜託”與“麻煩”的出現比例分別為 9.0% (73/807) 與 22.9% (185/807)，如**錯誤！找不到參照來源。**) 與**錯誤！找不到參照來源。**) 所示。

- (3) 有甚麼好玩的單機遊戲? 類似 暗黑破壞神. 世紀帝國. 新仙劍奇俠傳. cs. 俠盜獵車手. 末日危城 或可以自由改裝配備的賽車遊戲 只要好玩都請告訴我 謝謝
- (4) 請問有沒有哪位大大有 RPG" 尋秦記" 的遊戲載點!! 或是光碟片!! 最近有點想玩!!
- (5) 我用最新的硬碟下載但他說容量不夠 我看有足夠的容量但不能 下載 我試了很多次但不能 (V062-V065) 希望能給我直接下載ㄉ 我要 065 ㄉ 我不要官網ㄉ拜託ㄉ
- (6) 3.5代 太舊 也不能帶人練 5代 需要卡片才能玩其他歌 或 功能 4代我找不到 / \ 迪若 也沒有 麻煩一下~ ps:4代中間有滑球

(二) 祈使：求助是講求禮貌的行為，然而本研究發現與面對面溝通相同的是，中文網路使用者也經常運用直接的手段請求他人幫忙，而以祈使句式形成的求助主體行為即為常見的硬策略之一。與禮貌用語以及以下將介紹的軟策略相較之下，這種直接的手段似乎禮貌度較低，語氣較為強硬，因此，為了彌補硬策略對禮貌的負面影響，並凸顯得到協助的必要性，網路使用者常在祈使句式中加入禮貌用語如“請”與“拜託”，並以疊字加強懇求的語氣，或透過表必要性的情態動詞或副詞如“要”與“一定”等強調網友們提供協助的必要性，常見的慣用語為“請”、“拜託”、“要”或“一定”與可重疊式的動詞“幫”或“教”的組合，如“請幫我”、“拜託幫幫忙”、“一定要教教我”等。這種硬策略的出現比例佔中文語料總筆數的 39% (275/807)，如例句**錯誤！找不到參照來源。**) 與**錯誤！找不到參照來源。**)。

- (7) 請問一下：關於機動戰士鋼彈 SEED DESTINY GENERATION of C.E. 的第 6 關[7 關]要怎麼破 那 1 關是沙漠的地形我不懂那一關阿斯蘭作戰的意思我也有聽說過是要毀滅砲台 可是我不知道要怎麼會滅砲台 還有最後真好像會從後面攻擊敵人的樣[脈衝鋼彈沒有搭配羽翼形 巨砲型和巨劍型] 拜託請各位幫幫忙 小弟我已經打了很多次都沒破 請各位幫忙
- (8) 各位大大~[如標題]-我用鍵盤玩而且我是用 [2p] 的 我想要把滑壘按鍵的 [ctrl] 改成空白鍵~ 也就是把 [2p] 原本 [ctrl] 的功能改成按 [空白鍵] 就有原來只有按 [ctrl] 的功能~{應該了解我問的意思吧~如果了解我可再說的更清楚} 以上這個問題就請各位大大一定要為我解答~~ 也請大家踴躍回答{ ~ ~謝謝~ ~}

(三) 感謝：中文求助者常在求助或提出問題後運用感謝行為，常見的用語包括“謝謝”、“感謝”、“感激”與“感恩”，以及加上程度副詞而更為禮貌的“非常感謝”與“感激不盡”，另外還有些附帶口語語尾助詞、程度副詞與網路火星文的感謝語，如“謝啦”、“感恩哈”、“超級感謝”與“3Q”。上述感謝語的出現比例佔中文所有求助語料筆數的 37.2% (300/807)，如例句**錯誤！找不到參照來源。**) 到**錯誤！找不到參照來源。**)。

- (9) 我已經下載完了!是 8.0 的可以交我該如何將資料夾鎖密碼嗎?我不太會用，希望說的詳細一點的。謝謝! ^_^
- (10) 請問有玩過極速快感 11 的大大們~有人知道該怎麼樣調校車輛嗎?如果可以希望敘說完整一點~調校的功能何在及用途~還有什麼樣的調校適合什麼樣的賽道?請大大敘說的明白一點~感恩~我的即時通 ikmogto@yahoo.com.tw

- (11) 誰可給我凱大.R版的完整資料夾 請寄到 hahoo8747@com.tw 謝啦 我不要再有毒的喔
(12) 洛汗剛剛免費!! 我想練智精靈 請教我配點好口?? 順便告訴我 為什麼在 NPC 商店買的裝備素質都不一樣?? 3Q ^^

(四) 讚美：在請求協助時，中文求助者常伴隨讚美的行為，藉此推崇或稱讚可能願意或有能力提供協助的網友，對方的地位因而提升，可能因此更樂意提供協助。常見的讚美是透過“專家”、“高手”、“達人”、“大大”等稱謂來表達，出現比例佔中文所有求助語料筆數的 28.1%(227/807)，如例句**錯誤! 找不到參照來源。**)與**錯誤! 找不到參照來源。**)。

- (13) 一些網頁的 flash 都無法顯示 只顯示一個 f 網路上得影片也無法看 youtube 天空 等等 影片也變成 f 點進去變空白 沒辦法看 flash 也更新到 9 了 .. 求求電腦高手幫幫忙解決吧@@
(14) 最近一些活動逼的我必須買吳克群的「老子說 EP」，問題來了，活動現場有提供 CD player，可是老子說這張 EP 很特殊是少見的 AVCD 格式，我只需要裡面的 CD 格式部分，請一些多媒體的達人告訴我 如果只要播 AVCD 裡的 CD 格式那可以在一般的 CD player 播嗎？拜託拜託現在時間超趕的！！

(五) 求救：電腦網路討論區的資料量龐大，為了能夠在眾多的求助訊息中獲取注意力，中文求助者也常在主體策略中使用求救策略以吸引網友的目光。求救語給予被求助者壓力，屬於硬策略之一，但在浩瀚的網路中卻是可接受的表達方式。求救策略還具備幾個附加效果。一方面，求救能將求助者定位為亟需獲得幫助的弱者，藉此增加得到救援的機會；另一方面，求救也能傳達求助事件的急迫性，有助於問題儘速獲得解決。求救策略可透過祈使句式如“救救我”表達，也能透過疑問句式如“誰能救救我”、“有什麼方式能救救我”傳遞。語料中常見的求救語包括“急”、“求救”、“救命”、“救救我”等，出現比例佔所有求助語料筆數的 17% (118/807)，如例句**錯誤! 找不到參照來源。**)與**錯誤! 找不到參照來源。**)。

- (15) 想自己設一個伺服器，但安裝 Appserv 後，密碼帳號竟然進不去，也不小的問題是出在哪，誰能教教我解密，還有如果要設定 100 組的帳密要如何設定阿，救救我吧!
(16) 誰可以給我製作樂果和音樂盒的軟體順便教我怎麼製作可拿到 20 點 拜託 給我製作樂果和音樂盒的軟體 ~ 順便教我怎麼製作 !! 急 (!) > < '' 拜託 • v •

(六) 回報：網路使用者在 YAHOO! 奇摩知識+ 發問或解答，可根據在此平台受益與貢獻的程度而扣減或新增個人的點數。依點數的不同，此平台將使用者區分成八個等級，從低至高為知識貧民、初學者、實習生、研究生、專家、大師、知識長與知識名人；點數愈多愈有助於使用者升級，並享有更多的權利與尊重。網路求助者深暗點數升級的道理，因此常會以高額點數作為回報，利誘與激勵網友們踴躍幫忙。而且，由於求助者與提供協助者雙方都可從點數贈予機制得利，因此這項輔助策略有助於提高問題的回答率，出現比例佔所有語料筆數的 6.3% (51/807)，如例句**錯誤! 找不到參照來源。**)與**錯誤! 找不到參照來源。**)。

- (17) 請大家幫我尋找此問題網站，本人懸賞 20 點，請大家多多幫忙，本人因為跟露天拍賣賣家發生買賣起紛，有沒有要幫我和解請傳手機電話到我信箱:awz8619@yahoo.com.tw
(18) 我是說如果在 10 內的話有獎品 沒在 10 內也沒關西我會在開別之贊助 20 點 40 點夠力ㄇ? 配點方面 4 力 1 準

(七) 自謙：除了讚美被求助者之外，中文求助者也同時採取自謙或自貶的輔助策略以突顯自己獲得幫助的必要性，藉由讚美與自謙兩者塑造的強烈對比，網友們可能更樂意積極地幫助弱者，如此便達到此策略的目的。常見的自謙或自貶用語包括自稱語“小弟”、“白痴”等或形容詞“很蠢”、“很菜”等，出現比例佔中文所有求助語料筆數的 2.9%(23/807)，如例句**錯誤! 找不到參照來源。**)與**錯誤! 找不到參照來源。**)。

- (19) 朝鮮訓練場裡那個"火箭術"技術是作啥用的 請知道的大大告訴小弟
(20) 還有我想請問 新手訓練中有一把很像弓的東西但不是弓類似十字槍吧 我都沒子彈根本無法發射阿 我真的很蠢吧 不好意思 各位大大幫幫我吧 可以的話 即時方便請教嘛謝謝囉各位大大

(八) 道歉：因有求於人，求助者也常以道歉為輔助行為，放低姿態以顯示自己對

於麻煩網友幫忙的愧疚。此策略平均出現在求助主體行為的前面或後面。常見的道歉用語為“不好意思”、“抱歉”與口語的“拍謝”，但不曾使用“對不起”。此策略的出現比例為 2.4%(19/807)如例句**錯誤! 找不到參照來源。**)與**錯誤! 找不到參照來源。**)。

(21) 不好意思你們好呀！想請教各位之是大老們 我呀！玩了 Xbox360 的遊戲 Forza 2 (極限競速 2) 可是我要用他兩人玩呀！就是不太會用耶！

(22) 這是原圖→http://www.pandafile.com/uploads/55299438_41fcc0c3.png 修改完後→
http://www.pandafile.com/uploads/40322706_2-11.png 這是用 ps 還是 P I ? 是如何改色的? 能給我做法嗎? 很像要嘗試這類改色。希望可以教教學√用成圖比較麻煩點√抱歉噢

(九) 自清：網路求助者常在發問前，先澄清自己曾努力找尋相關答案但未果，因此才有尋求幫忙的需要，藉此將求助的行為合理化，甚或取得網友們的理解與同情。此輔助策略的出現比例佔中文請求語料總筆數的 1.7% (14/807)，如例句**錯誤! 找不到參照來源。**)與**錯誤! 找不到參照來源。**)。

(23) 我一直找我一直找我一直找我一直找我一直找都找不到 因為我家電腦不能用 options language sytem 來用就是了

(24) 為什麼安裝 QuickTime 以後要看網路上的影片或玩小遊戲都會出現"藍色的 Q，中間再一個白色問號"? 我查過知識+，可是都沒用 誰能幫幫我?!

(十) 抱怨：中文網路使用者偶爾會以抱怨的硬策略來表現求助行為的動機，亦即對造成個人遇到問題或困難的相關人事物提出抱怨，有時也涉及評論或指責網友們回覆答案的好壞。這種輔助策略禮貌度雖低，但卻可能激發網友們的同理心，只是在中文的使用率並不高，出現比例僅佔中文語料總筆數的 0.2% (2/807)，如例句**錯誤! 找不到參照來源。**)與**錯誤! 找不到參照來源。**)。

(25) 台中通訊鑽石通訊 為何取消我勿出價記錄 又將我列入黑名單 因為想要衛星導航手機 所以出價勿 只想知道原因~~這樣時在太過份勿 欺騙大眾什麼都不說~~逃避請問我該如何去問 我只想知道為何還是不巧沒人出價所以~~~ 不想賣也能說 Y 幹麻這樣

(26) 我電腦是新的! 我也才剛灌好也才剛掃完毒 這是甚麼爛回答

英文網路溝通特色

美國 YAHOO! Answers 的語料則顯示，雖然求助屬於威脅顏面的行為，但是英文網路使用者卻偏好直接求助，以傳遞訊息的簡潔與清晰為目的，鮮少如中文般運用其他修補顏面的輔助或支援策略。此外，英文網路求助者常使用穢語表達對求助事件的不滿情緒，以引起被求助者的共鳴與回覆。以上兩者構成英文網路溝通的主要特色，以下分別舉例說明。

(一) 言簡意賅：英文網路溝通簡單明瞭的特性，常見於以下兩種情形。一種是求助者在網站平台所提供的問題標題欄位上直接以疑問句式詢問求助標的，並無內文說明求助的原因或相關細節，這種情形的出現比例佔英文語料的 13.0% (105/807)，如例句**錯誤! 找不到參照來源。**)與**錯誤! 找不到參照來源。**)，兩句皆取自標題列。另一種情形是求助者雖然在標題欄載明問題並提供內文，但內文僅扼要說明與求助事件相關的背景訊息，並未解釋求助的原因等相關訊息，這種情形的出現比例佔英文語料筆數的 5.0% (40/807)，如例句**錯誤! 找不到參照來源。**)與**錯誤! 找不到參照來源。**)。由此可見英文網路溝通傾向目標導向，求助者的重點在於如何讓問題癥結一目瞭然，如此有助於溝通的效率。

(27) 標題 Is there a tutorial that shows how to format only c:\ the drive with windows?

(28) 標題 Where can i download a free RM-Mp3 converter?

(29) 標題 What should i do to heal virus infection if my pc effected with virus?

內文 i'm using AVG FREE anti virus software.

(30) 標題 Anyone know how to automatically play videos on your youtube channel?

內文 Especially automatically playing your favorites.

(二) 動之以情：英文網路使用者進行求助行為時，偶爾會以穢語表達受挫無奈和

不滿的強烈負面情緒，穢語所指示的對象為求助事件本身，可讓網友們感同身受，達到動之以情的目的。常見的穢語如“damn”、“heck”等，出現比例佔英文所有語料的1.0% (8/807)，如例句**錯誤！找不到參照來源。**)與**錯誤！找不到參照來源。**)。

(31) Why the heck does Google Chrome keep crashing? I just downloaded it. It keeps on crashing every few seconds, when I try to search in the bar.

(32) Obviously it's something to do with the ink cartridges, but I have no idea why it's doing it. I underline and bullet point stuff on MS Word in red, but when I print out the document it comes out yellow. Can I fix it or do I need to buy some new ones? Damn these expensive things.

中英網路溝通共同特色

除了上述的個別特色之外，中文與英文的求助語料也具備共同的特色，包括以下的解釋原因與網路用語。

(一) **解釋原因**：中英網路求助者都常解釋求助的原因，以便讓網友們了解問題的來龍去脈，作為求助的輔助策略之一。解釋原因出現的位置無固定模式，可放求助句的前、中、後，或與請求反覆交替出現。但基於前因後果的邏輯關係，中英文的解釋原因都較常出現在求助句的前面。這個策略在中英文語料中的出現比例分別佔總語料筆數的56.1% (453/807) 和 53.0% (428/807)，可見在兩個語言的使用頻率相當，如例句**錯誤！找不到參照來源。**)與**錯誤！找不到參照來源。**)。

(33) 我把照片檔放在D巢但是不知情況下把它刪掉,又把資源回收筒清除了 我要怎麼樣才能救回已經被清除的檔案? 麻煩大家!

(34) my stupid brother and was having fight. during our fighting he throw pen against me but pen get strike to my dearer computer's screen. it get cracked. previously he broke my keyboard' space bar. please tell how much cost will it take to get it repair. and also tell what punishment should i give to my stupid bro.

(二) **網路用語**：中文流行的網路用語包括注音文與相似音兩種形式。注音文係指網路使用者為講求時效與便利，以注音符號取代原本要表達的詞語，常見的注音文如“ㄉ”（不）、“ㄇ”（嗎）、“ㄉ”（的）、“ㄊ”（他）、“ㄋ”（呢）、“ㄟ”（吃/痴）、“ㄑ”（去）、“ㄌ”（欸）、“ㄝ”（也/耶）。這種用法使得中文原本同音不同形的詞語（如“喝”、“呵”等）在電腦螢幕上以同一個形體出現（取其音“ㄉ”）。相似音則指網路使用者以中文相近發音的詞語代表外來詞語，如“猴塞雷”（粵語表厲害）、“卡哇伊”（日文表可愛）、“甘八爹”（日文表加油）等。英文的網路用語則以首字縮寫與諧音字為主，常見的首字縮寫例子如“BTW”（表 by the way）、“ASAP”（表 as soon as possible）；常見的諧音字例子如“plz”（表 please）、“u”（表 you）、“n”（表 and）、“dunno”（表 don't know）、“thx”（表 thanks）等。還有一種網路用語是跨越語言界線的，中文與英文皆普遍使用，即表情符號。表情符號指網路使用者以鍵盤上的符號組成臉部表情或動作以傳達感受或情緒，有助於活化溝通或拉近距離。常見的表情符號如“;-)”（微笑），“T_T”（哭得很傷心），“Orz”（佩服得五體投地）。中文與英文求助語料中使用上述網路用語的頻率相當，分別佔中英所有求助語料總筆數的37.2% (300/807) 與 32.0% (258/807)，如例句**錯誤！找不到參照來源。**)至**錯誤！找不到參照來源。**)所示。

(35) 請問誰有棒棒堂ㄉcss 因為我ㄉ知道ㄑ哪裡找棒棒堂ㄉcss 請位大大幫忙 謝謝

(36) 經過網X亂改後=有狠多技能都跟以前不一樣 所以我想問各位大大們我練ㄉ羊逝水雷羊+案係女巫我想問 我ㄉ技能藥學啥? 因為既能有歌改過所以很多人ㄉ回答都ㄉ型ㄉ 所以我從新發問 請大大們告訴我 水雷暗羊藥學ㄉ既能有哪些

(37) 從7-11買完整全新的希望光碟包，回家灌好時，遊戲跑出更新的畫面，也更新好了，結果要玩的時候就跑出3個亂碼的視窗，然後就不能完，網路下載我也試過了，就是不行，這是第5次灌了，為啥不能玩阿，之前灌過就可以玩，現在玩就不行 太奇怪了吧 是哪裡發生問題阿，都可以玩完美世界了，希望竟然不能跑 請各位電腦高手幫忙小弟我一下吧 我由衷感謝你們 ORZ

(38) doing a report n i need it double spaced help p|zzzzz!!!! thnak u!!!

(39) Has anyone been using this new browser for some time now and has any opinions? Thanks!

8)

(40) Ok, so my parents finally trust me enough to have our old family computer fixed, and have more memory added to it. It's going to be in my room from now on, and I just wanted to know if anybody knew of any good anti-virus and anti-spyware software I can download off the internet? I mean yea I could go buy it, but I had to pay half the bill to fix the dang thing, so for now I'm resorting to downloading one! Any help would be appreciated! BTW I'm getting it back today 9/6/08 so ASAP would be awesome!!

結論

本研究的中英對比分析顯示，網路語言因其媒介特性形成特定的溝通文化，且因語言不同而受制於不同的文化框架，展現不同的策略運用，中文與英文即具備不同的網路溝通文化，反映在網路使用者的求助策略上。一方面，中國人在網路上最常使用的求助主體行為之一是語用功能清晰的祈使句，並偏好以禮貌用語的內部修飾手段或其他各種軟硬兼施的外部修飾手段作為增加禮貌程度的輔助或支援策略，這一點與面對面溝通的文獻結論是相同的。然而，美國人在網路上的求助行為卻比面對面溝通更為直接，所以面對面時偏好的慣例型間接策略在網路求助中並不多見，反而是以直接提出問題為典型的求助方式。因此，英文在面對面溝通的情況下運用間接的表達方式雖可達到禮貌，但在網路溝通的情況下則更重視語用功能的清晰，盡可能減少網友認知的負擔；再加上考量到網友並無提供協助的義務，強制要求反而較能凸顯需求，進而獲得注意。此外，在不同文化中網路使用者慣常容許的語用清晰手段也有所差異，中文網路求助者傾向以祈使句的動作動詞述明網友們被期待協助完成的事件；而英文網路求助者傾向以疑問句詢問完成事件的方法。而且，不同文化也要求不同程度的顏面修補，中文網路求助者習慣在求助前後加上較多輔助或支援策略，英文網路求助者則喜好直接呈現問題、切入重點，而這些差異也會決定中英文網路使用者如何判斷其他求助者的行為是否合乎禮貌。

本研究蒐集的語言材料與歸納的對比結果可做為中文與英文語言教學者與學習者的參考資料，藉此瞭解中文與英文在禮貌認知上的差異，以及認識網路溝通媒介對兩個語言產生的影響。本研究的結論顯示，現代語言教學者應注重網路世界的語言發展模式，幫助學習者瞭解不同語言在網路溝通上的差異，以便在線上進行合宜的跨文化溝通。本研究未來可朝向其他溝通平台（如電子郵件、部落格、即時通訊等）以及其他言語行為（如交易拍賣、資訊分享、心情抒發等）進行中英對比研究，以進一步驗證本研究的結論，並有助於更完整地勾勒出中英網路溝通特色的全貌。

參考書目

- Austin, J. L. (1962). *How to do things with words*. Oxford: Clarendon Press.
- Barnes, S. B. (2003). *Computer-mediated communication: human to human communication across the Internet*. Boston: Pearson Education.
- Baron, N. S. (1984). Computer mediated communication as a force in language change. *Visible language*, 18.2, 118-141.
- Baron, N. S. (1998). Letters by phone or speech by other means: the linguistics of email. *Language and Communication*, 18, 133-170.
- Baron, N. S. (2003). Why email looks like speech. In J. Aitchison & D. Lewis (Eds.) *New media language* (pp. 85-94). London: Routledge.
- Beibwenger, M. (2000). *Kommunikation in virtuellen Welten: Sprache, Text und Wirklichkeit* [Communication in virtual worlds: Language, text, and reality]. Stuttgart: ibidem.
- Blum-Kulka, S. (1987). Indirectness and politeness in requests: Same or different? *Journal of*

- Pragmatics*, 11.2, 131-146.
- Blum-Kulka, S. & Olshtain, E. (1984). Requests and apologies: a cross-cultural study of speech act realization patterns (CCSARP). *Applied Linguistics*, 5.3, 196-213.
- Blum-Kulka, S., House, J. & Kasper, G. (Eds.) (1989) *Cross cultural pragmatics: Requests and apologies*. Norwood, N.J.: Ablex.
- Brown, P. & Levinson, S. C. (1987). *Politeness: some universals in language usages*. Cambridge: Cambridge University Press.
- Clark, H. H. & Schunk, D. H. (1980). Polite responses to polite requests. *Cognition*, 8, 111-143.
- Collot, M. & Belmore, N. (1996). Electronic language: A new variety of English. In S. C. Herring (Ed.) *Computer-mediated communication: linguistic, social and cross-cultural perspectives* (pp. 13-28). Amsterdam: John Benjamins.
- Craig, R. T., Tracy, K. & Spisak, F. (1986). The discourse of requests: assessment of a politeness approach. *Human Communication Research*, 12.4, 437-468.
- Crystal, D. (2001). *Language and the Internet*. Cambridge: Cambridge University Press.
- Flanagin, A. J. & Metzger, M. J. (2001). Internet use in the contemporary media environment. *Human Communication Research*, 27.1, 153-181.
- Fraser, B. (1975). Hedged performatives. In Cole, P. and Morgan, J. (Eds.) *Syntax and semantics 3: speech acts* (pp. 187-210). New York: Academic Press.
- Fukushima, S. (1996). Request strategies in British English and Japanese. *Language Sciences*, 18, 671-688.
- Furlong, M. S. (1989). An electronic community for older adults: the seniornet network. *Journal of Communication*, 39.3, 145-53.
- Herring, S. C. (1996). *Computer-mediated communication: linguistic, social, and cross-cultural perspectives*. Amsterdam: John Benjamins.
- Herring, S. C. (2001). Computer-mediated discourse. In D. Schiffrin, D. Tannen, & H. Hamilton (Eds.) *The handbook of discourse analysis* (pp. 612-634). Oxford: Blackwell Publishers.
- Herring, S. C. (2004). Slouching toward the ordinary: Current trends in computer-mediated communication. *New Media & Society*, 6, 26-36.
- Hong, W. (1996). An empirical study of Chinese request strategies. *International Journal of the Sociology of Language*, 122, 127-138.
- Hong, W. (1998). *Request patterns in Chinese and German: A cross-cultural study*. Munchen: LINCOM EUROPA.
- Lee-Wong, S.-M. (1994). Imperatives in requests: direct or impolite—observations from Chinese. *Pragmatics*, 4.4, 491-515.
- Lee-Wong, S.-M. (1999). *Politeness and face in Chinese culture*. Frankfurt: Peter Lang.
- Lewis, D. (2003). Online news: A new genre? In J. Aitchison & D. Lewis (Eds.) *New media language* (pp. 95-104). London: Routledge.
- Ma, R. (1996). Computer-mediated conversation as a new dimension of intercultural communication between East Asian and North American college students. In S. C. Herring (Ed.) *Computer-mediated communication: linguistic, social and cross-cultural perspectives* (pp. 173-185). Amsterdam: John Benjamins.
- Marcoccia, M. (2004). On-line polylogues: conversation structure and participation framework in internet newsgroups. *Journal of Pragmatics*, 36, 115-145.
- Parks, M. R. & Roberts, L. D. (1998). “Making MOOsic”: the development of personal relationships on line and a comparison to their off-line counterparts. *Journal of Social and Personal Relationships*, 15, 517-537.
- Rinnert, C. & Kobayashi, H. (1999). Requestive hints in Japanese and English. *Journal of Pragmatics*, 31.9, 1173-1201.

- Turkle, S. (1995). *Life on the screen: Identity in the age of the Internet*. New York: Simon & Schuster.
- Van Gelder, L. (1990). The strange case of the electronic lover. In G. Gumpert & S. L. Fish (Eds.) *Talking to strangers: mediated therapeutic communication* (pp. 128-142). Norwood, NJ: Ablex.
- Wallace, P. M. (1999). *The psychology of the Internet*. New York: Cambridge University Press.
- Werry, C. C. (1996). Linguistic and interactional features of Internet relay chat. In S. C. Herring (Ed.), *Computer-mediated communication: linguistic, social and cross-cultural perspectives* (pp. 47-63). Amsterdam: John Benjamins.
- Yates, S. J. (1996). Oral and written linguistic aspects of computer-conferencing. In S. C. Herring (Ed.) *Computer-mediated communication: linguistic, social and cross-cultural perspectives* (pp. 29-46). Amsterdam: John Benjamins.
- Yeung, L. N. T. (1997). Polite requests in English and Chinese business correspondence in Hong Kong. *Journal of Pragmatics*, 27.4, 505-522.
- Zhang, Y. (1995). Strategies in Chinese requesting. In G. Kasper (Ed.) *Pragmatics of Chinese as native and target language* (pp. 25-68). Honolulu: University of Hawaii Press.
- Ziegler, A. (2002). Email – Textsorte oder Kommunikationsform? Eine textlinguistische annäherung. In A. Ziegler & C. Durscheid (Eds.) *Kommunikationsform Email* [Communication on form email] (pp. 9-32). Tübingen: Stauffenburg.
- 盧諭緯 (1997)。說文解字：初探網路語言現象及其社會意義。《第二屆資訊科技與社會轉型研討會》，12月18至20日，台北：中央研究院社會學研究所。

A Study on the Acquisition of English Expletives by EFL Learners in Taiwan

Aili Hsin (忻愛莉), Yi-ru Lin (林怡如)

National Kaohsiung Normal University

gealhsin@nknuc.nknu.edu.tw

j010629@yahoo.com.tw

This study aims to investigate the influence of Chinese parameter settings on the supply of subject expletives in English and to explore the acquisition development of EFL learners' choice between expletives of *there* and *it*. The aspects we address to are (a) the influence of whether Chinese *you* (有) and *sihu* (似乎) appear in sentences on the supply of subject expletives, (b) the triggering effect of expletives on [-null subject] parameter resetting, (c) the acquisition differences of expletives *it* and *there*, and (d) the correlation between the acquisition of the syntactic features in expletive constructions and the accurate choice of expletives. Four groups of participants we invited in the study were 30 native English speakers, 43 English major graduates, 45 English major seniors in the university and 51 senior high students. The instruments consisted of a Chinese-to-English translation, a blank-filling task and a test of grammaticality judgment. The data collected were then analyzed with SPSS software. The findings are briefed as follows. First, the translation items with Chinese *you* and *sihu* were supplied with significantly more expletives than those without, and the association of *you-there be* and *sihu-it seems* even further influence the senior high participants to generate the most expletive supply in items without Chinese counterparts. Second, *there* in *there be*, instead of *it*, might be a potential trigger of [-null subject] parameter change. Third, *it* was acquired better and improved faster than *there*. Fourth, the syntactic features of expletive constructions were acquired separately for the EFL groups and thus did not facilitate expletive acquisition much. These findings suggest explicit classroom instructions and adequate input on expletives are essential and may accelerate the acquisition.

INTRODUCTION

English expletives do not have semantic meaning and serve to satisfy the syntactic requirement that a sentence must have an overt subject, so English is named a non-null subject language with the feature [-null subject] (Rizzi, 1986). The feature means a subject position must be filled with an overt subject which cannot be dropped. But this requirement does not exist in Chinese (Huang, 1989). Chinese subjects known in contexts are allowed to be dropped, so even though the subject positions are empty, we can know the semantic contents there. This is why Chinese is termed a null subject language with the feature [+null subject]. The feature means Chinese subject positions are allowed to be empty.

Based on the Chinese-English differences of (a) subjects can be dropped or not, and (b) subjects are semantic or not, the researcher designed a small-scale, pilot study for three groups of participants: the third-year junior high students, the second-year English majors and the English majored graduates. The instrument was Translation Task from Chinese to English. The Chinese sentences were the structures with subjects dropped. We attempted to examine whether the participants could provide subject expletives or referential subjects which have semantic contents in English sentences. We found that in Chinese *you* sentences as in (1), the rate that the participants provided subject expletive *there* was high (over 84 % for each group). However, the rate of supplying *there* in the sentences without *you* as in (2)

decreased a lot. Only 57.9% of the English graduates used *there*. The English sophomores even decreased to 39.1% and the third-year junior high students got the worst, 38.4%. Besides, for those who did not use *there*, they supplied referential subjects as in (2b) and (2c).

- (1) 院子裡有一隻狗，很兇。
 a. There is a dog in the yard. It is fierce.
 (2) 餐廳來了個新廚師。
 a. There comes a new cook in the restaurant.
 b. The restaurant comes a new cook.
 c. A new cook comes to the restaurant.

From the results, we observe three phenomena: (a) when the existential marker *you* appears in a sentence, expletive *there* was highly possible to be triggered, (b) when *you* did not appear, the participants preferred referential subjects to subject expletive *there*, and (c) the participants knew that English subjects were obligatory, so they provided a subject to the empty subject position of the second clause as in (1a). Therefore, based on the observations from the translation task above, expletive *there* is related to the English setting, [-null subject], the obligatoriness of lexical subjects, congruent with Hyams' (1986) observation. But we question whether the participants have truly acquired expletive *there* on the ground of the fact that the use of expletive *there* tended to rely on whether the Chinese *you* appeared or not. The participants may regard *there be* as a chunk with the Chinese meaning of *you*, so when *you* appears, *there be* structure is often triggered. Similarly, this chunk learning might extend to expletive *it*, which frequently co-occurs with the verb, *seem*. Learners may see *it seems* as a chunk with the Chinese equivalent of *sihu*. We intend to explore whether the supply of expletives stems from the chunk learning of the Chinese equivalents or from the acquisition of expletives, and whether the supply of both expletives could trigger [-null subject].

In addition to expletive supply, there seem to have some problems with learners' knowledge of what expletive they should fill in the blank in (3). Even some advanced learners were not sure of the answer to the blank either. This indicates they were still unclear with what complement could fit in the expletive construction in (3). It makes us wonder why these advanced learners' expletive accuracy is not sound yet after more than ten years of learning.

- (3) _____ seems to be nothing interesting. [it/there]

From both the translation task for expletive supply and the blank-filling item for expletive accuracy, we can understand expletive acquisition is not easy for EFL learners. Hence, this study mainly centers on the factors that affect expletive acquisition in attempt to provide pedagogical implications for efficient expletive learning and teaching.

PROPERTIES OF EXPLETIVE CONSTRUCTIONS

Zwart (1991) stated that English existential *there* constructions have four properties exemplified in Table 1. First, the postverbal DP shows definiteness effect. The DP (i.e. "a girl" in (4)) has to be indefinite. Second, the verb has to be unaccusatives, inclusive of unaccusatives, raising verbs, copula (*be*) and unergative verbs with locative PPs. Each kind of verbs is shown in examples (5) to (8). Third, these constructions trigger an existential reading. Zwart compared (9a) and (9b). (9a) is ambiguous in that it has "presuppositional" and "existential" readings, shown in (11a & b) respectively. On the contrary, (9b) is disambiguous with "existential" reading only.

- (11) a. many men are such that they are in the garden.
 b. a situation exists such that many men are in the garden. (Zwart, 1991)

Table 1. *The four properties of there constructions in Zwart (1991)*

Properties	Examples
Definiteness effect	(4) There is <i>a girl</i> at the door.
Unaccusative verbs, raising verbs, copula and unergative verbs with locative PPs	(5) There arrived a package. (unaccusative verb) (6) There seems to be a dog in the yard. (raising verb) (7) There was happiness. (copula) (8) There hung a picture on the wall. (unergative verb with a locative PP)
Existential reading	(9) a. Many men are in the garden. b. There are many men in the garden.
Stage-level predicates	(10) a. There are many people <u>sick</u> . b. *There are many people <u>intelligent</u> .

The fourth property mentioned by Zwart is that stage-level predicates, not individual-level predicates, are allowable in existential constructions, as the examples in (10) show. The final property this study examines is the necessity of “to be” in *there*-sentences with raising verbs as in (12). According to Lasnik (1992), “be” is a case assigner, so “to be” is necessary.

- (12) a. [TP/AgrsP there_i seems [TP/AgrsP t_i to be [PP a man in the room]]]
b. *[TP/AgrsP there_i seems [TP/AgrsP t_i [PP a man in the room]]]

For *it* constructions, according to Rothstein (1995), expletive *it* occurs with raising verbs, passive verbs and adjectives “thematically selecting only an internal CP,” as the following examples.

- (13) It seemed/appeared/turned out that they were late.
(14) It was widely believed that the world is flat.
(15) a. It is likely that I’ll be on time.
b. It is possible that I’ll pass the course.
c. It is certain / obvious that he will win. (Rothstein, 1995)

Agreement in Expletive Constructions

McCloskey (1991) observed that a main verb agrees with the postverbal DP in *there*-sentences as in (16), but a main verb does not agree with the postverbal clause in *it* sentences as in (17).

- (16) a. There exist no good *solutions* to this problem.
b. *There exists no good *solutions* to this problem.
c. There exists no *solution* to this problem. (McCloskey, 1991)
(17) a. It seems equally likely at this point *that* the president will be reelected *and that* he will be impeached.
b. * It seem equally likely at this point *that* the president will be reelected *and that* he will be impeached. (McCloskey, 1991)

It is obvious that (16b) is incorrect because the main verb (i.e. exists) does not agree with the postverbal plural DP (i.e. no good solutions). That is, the structural subject, *there*, must syntactically associate with a postverbal DP, like the way (14a & c) do.

The link between *there* and a postverbal DP cannot apply to *it* and its associated clause, as exemplified in (17a & b). The matrix verb, *seems*, must be singular even though two complement clauses conjoin. The subject expletive *it* is directly in agreement with the main verb, *seem*.

METHOD

Three tasks were designed to test three EFL groups of learners in Taiwan with different proficiency and a group of native English speakers. The three EFL groups consisted of 51 third-year senior high students from Kaohsiung Municipal Sin Sing Senior High School, 45 English major seniors from National Kaohsiung Normal University, and 43 English major graduates at National Kaohsiung Normal University. Thirty native speakers were recruited and their scores were the norm of analyses. The following introduces the three tasks and their functions.

Chinese to English Translation Task

This task was set up to test the influence of the appearance of *you* and *sihu* in sentences on the supply of expletives (see Table 2 for the test items) and to examine the triggering effect of expletives on [-null subject] (see Table 3 for the test items). There are a total of 15 items. In Table 2, the participants were required to translate the parts underlined. *There*-sentences include the ones with and without *you*. The verb in the sentences with *you* is “be,” and the verb in the ones without *you* is unaccusative. *It*-sentences contain the ones with and without *sihu*. The verbs in the sentences with *sihu* is “seem,” and the verbs in the ones without *sihu* are passive verbs and raising adjectives (e.g. is possible). The participants were demanded not to use adverbs or adverb phrases in sentences. We intend to know whether they would use subject expletive *it*.

Table 2. *The items to test the influence of you and sihu on supply of expletives*

<i>There</i>	<i>It</i>
1. 你看， <u>有一隻狗在那裡</u> ，是隻狼狗。	6. 房子的燈亮了， <u>似乎 Bob 已經到家了</u> 。
2. 兩年前， <u>有一家郵局在公園的轉角</u> 。	7. 經過一陣子的討論， <u>似乎那些人的想法是對的</u> 。
3. 三小時前的餐廳大火， <u>逃出來了三個</u> 人。	8. 天空烏雲密布， <u>似乎要下雨了</u> 。
4. 哇！ <u>進來了一個美女</u> 。	9. 我估算了一下時間， <u>可能我們會遲到</u> 。
5. 剛剛發生了一件有趣的事。	10. <u>媽媽很明顯是在生氣</u> ，你還這樣不聽話。
	11. <u>預料 Bob 會準時到</u> 。

The scores of the sentences with *you* were correlated with those of the ones without *you* to inspect the correlation between *there be* and the supply of *there* in *there*-unaccusative verb sentences (Items 3~5). The postverbal DPs in Items 3, 4 and 5 are nonspecific indefinites and they stay in the VP domain. According to Diesing's (1992) assumption on indefiniteness, a nonspecific indefinite inside VP existential closure is given existential reading. So the precise translations for Items 3 to 5 are the *there*-unaccusative verb pattern. Hence, if the scores of the sentences with and without *you* are not positively, significantly correlated, then the supply of *there* in *there be* might be more possible to be a chunk learning. The acquisition of *there* is not stable enough to achieve positive correlation. The score analyses of *it*-items are much the same as those of *there*-items. The scores of the sentences with and without *sihu* were correlated to test the supply of *it*.

The items in Table 3 were designed to make the subjects of the second clauses empty. The empty subject positions in English should be filled with referential DPs. The purpose of the design is to examine the triggering effect of expletives on referential subjects. So the scores of these items were correlated with those of the translation items with *you* since Yang (2001) reported that the *there* in *there be* is the trigger of [-null subject], and were correlated with the scores of the items with *sihu* and raising adjectives because they appear earlier in textbooks than passive verbs.

Table 3. *The items for the triggering effect of expletives on [-null subject]*

12. 我上星期去坐高鐵，真的很快。
13. 爸爸剛把新的沙發搬回來，看起來好像很累。
14. 要是我妹妹沒睡覺，會生氣的。
15. 每次姐姐看到老鼠都會尖叫。

Blank-Filling Task

This task focuses on expletive accuracy as well as [-null subject]. If the participants recognized the necessity of an overt subject, they next had to decide to use *it* or *there*. The key to the choice was the postverbal element. When a DP (*a cat*) or an IP (*to be a cat*) was in the complement position, *there* was triggered. When a CP (*that a cat is in the yard*) occupied there, *it* was legitimate. In order to control the choices of subjects in sentences, the answer to the blanks has the options of *it*, *there* or ϕ (which means nothing needs to be filled in). The total number of the items is 17. Nine of them were for *there*, and the others were for *it*. Some of the items are listed in Table 4.

Table 4. *The items for Blank-filling task*

- There-DP /IP (to be+DP) complements:** [it/there/ ϕ]
1. The weather reporter said that _____ is likely to be an earthquake in Japan today.
 2. _____ were too many people in the room at that time.
 3. _____ still remained a little work to be done.
-
- It-CP complements:** [it/there/ ϕ]
1. Tom looked sleepy. _____ seemed that he was very tired.
 2. Did you know in the 16th century, _____ was widely believed at that time that the earth was flat?
 3. _____ is expected that Wang will win the championship this season.
 4. It's ten. _____ appears that Mary will not be coming.

The data from this task were analyzed in SPSS by one-way ANOVA. The purpose is to investigate the differences on expletive accuracy across groups, and the acquisition differences between *there* and *it* in each group.

Grammaticality Judgment

This task served to explore the correlation of expletive accuracy with the syntactic properties of expletives constructions. The expletive accuracy for *there* constructions means the relationship between *there* and postverbal DPs. The syntactic properties refer to “definiteness effect,” “to be,” “agreement” and “stage-level predicate.” The test items were designed based on these properties. The scores of each property were correlated with those of *there*-DP to test whether the acquisition of syntactic properties would promote the accurate rate of *there*. The total number is 23. Some of them are listed in Table 5.

Table 5. *The items for there in Grammaticality Judgment*

<p>There (True/False)</p> <p>To be</p> <p>1. That girl appears to be a princess. 2. *The story turns out a happy ending.</p> <p>Definiteness effect</p> <p>*There is the man in the room. There arrives a crazy singer.</p> <p>Agreement</p> <p>There stands a tree in front of the house. *There exists no good solutions to this problem.</p>	<p>Stage-level predicates</p> <p>There remain two computers available for us. *There appeared to be some students honest.</p> <p>There-DP/IP</p> <p>*There comes that this event won't be forgotten. 2. There rose the sun above the horizon.</p>
---	---

The expletive accuracy for *it* constructions signifies the relationship between *it* and postverbal CP complements. The syntactic properties refer to “agreement” and “raising verbs, raising adjectives and passive verbs.” The rationale and the purpose of the design are the same as those of *there* constructions. *It* constructions contain a total of 15 items. The following in Table 6 are some of them.

Table 6. *The items for it in Grammaticality Judgment*

<p>It (True/False)</p> <p>Agreement</p> <p>1. *It seem that the girls are sad and that they are going to cry. 2. It is known that Sam and Paul are Chinese. 3. It is clear that Tom left early and that he went home late.</p> <p>Raising verbs/adjectives/passives</p> <p>1. *It occurs a crash in the air. 2. It is believed that Jolin is famous.</p> <p>It-CP complements</p> <p>1. *It seems to be a baby in the room. 2. It continued to upset Sherry that she had lost her mom's necklace.</p>

RESULTS AND DISCUSSION

The following presents the findings in four aspects: (a) the influence of *you* and *sihu* on the supply of expletives, (b) the triggering effect of expletives on [-null subject], (c) the acquisition differences of *there* and *it* within each group and across groups and (d) the correlation between the syntactic properties of expletive constructions and expletive accuracy. The findings are briefly discussed in the following.

The Influence of *You* and *Sihu* on the Supply of Expletives

To know the influence, we correlated the scores with and without *you*, and those with and without *sihu* from Translation Task. The results are shown in Table 7; the percentages of expletive supply are provided in Table 8. The percentages of expletive supply in the items with *you* and *sihu* are much higher than those without, so significant correlations are hard to achieve. The corresponding English of the test items without *you* and *sihu* are legitimate sentences, but expletives were not often activated as the items with *you* and *sihu*. Hence, the participants possibly established the relationships between *you* and *there be*, and between *sihu* and *it seems*, so when *you* and *sihu* appear, expletives were easier to be triggered. *There be* and *it seems* were regarded as chunks, corresponding to *you* and *sihu* respectively.

Table 7. The correlation coefficients between the scores with and without *you* and *sihu*

Correlation Groups	With and without <i>you</i>	With and without <i>sihu</i>
G(graduate) (n=43)	.187	.143
U(ndergraduate) (n=45)	.266*	.168
S(enior) (n=51)	.039	.345*

*p<.05

Table 8. The percentages of participants supplying expletives

Percentage Groups	With <i>you</i>	Without <i>you</i>	With <i>sihu</i>	Without <i>sihu</i>
	There be	There-unaccusatives (e.g. There occurred...)	It seems	It is adj./P.P.
G	62.8	3.9	82.2	32.6
U	80.0	5.9	79.3	28.1
S	89.2	12.4	90.2	37.3

In Table 8, we found the seniors provided the most expletives in the items without *you* and *sihu*. Following a scrutiny of the raw data of the items without *you*, most of the EFL participants who supplied *there* adopted *there be* to translate the sentences without *you*, which resulted in ungrammatical run-on English sentences. Take “發生了一件有趣的事” as an example. Among the seniors who provided *there*, most of them wrote the sentence like “There was something interesting happened,” instead of “There happened/occurred something interesting.” If we translate “There was something interesting happened” into Chinese, we have “有一件有趣的事發生了.” The seniors reinterpreted the sentence and translated it in the way they were familiar with. This could be a recency effect which explains the phenomenon that the newly learned messages would be memorized better because the learners are concentrating on them (Slavin, 2005; Greene, 1986).

The low percentages of there-unaccusatives and it is adj./P.P. in production indicate that expletives *there* and *it* are not productive in these constructions and hence expletives are learnt by L2 learners as a frozen phrasal chunk of *there be* corresponding to *you* and *it seems* to *sihu*, instead of a syntactic rule of filling in a semantic dummy element in an empty subject position. With such a tremendous discrepancy between the with vs. without constructions, we observe no activating effect of *you* and *sihu* on the production of English expletives on the more marked structures of there-unaccusatives and it is adj./P.P.

The Triggering Effect of Expletives on [-null subject]

[-Null subject] means the obligatoriness of overt subjects. English is this kind of language. On the other hand, Chinese is [+null subject] because a subject can be dropped when the subject is known in context. When learning English, learners in Taiwan need to switch the parameter setting from [+null subject] to [-null subject]. Hyams (1986) maintained that expletives could trigger [-null subject]. To examine the triggering effect of expletives, the scores of the translation items in Table 3 were correlated with those of the items with *you*, *sihu* and raising adjectives respectively. Table 9 shows the correlation coefficients. The graduate and senior groups made significant and positive correlation between [-null subject] and *there*, but the undergraduate group did not. Besides, no group gets significant correlation between *it* and [-null subject].

Table 10 indicates that the participants have acquired the setting of [-null subject] in general. If expletives are the trigger, the scores of expletives and [-null subject] should be in positive and significant correlation. According to the correlation coefficients, *there* is a

potential trigger instead of *it*. We cannot affirm *there* is the trigger since the undergraduates did not reach significant correlation. This insignificance may result from the fact that they did not choose *there* to translate the items with *you*. Moreover, the significant correlation from the graduates and the seniors indicates that even though the participants regarded *there be* as a chunk, this learning might be able to promote the parameter change from [+null subject] to [-null subject] too.

Table 9. *The correlation of [-null subject] with there and it*

Correlation Group	[-null subject] and <i>you-there be</i>	[-null subject] and <i>sihu-it seems</i>	[-null subject] and <i>it is adj.</i>
G	.336*	-.114	.195
U	-.093	-.122	.198
S	.316*	.034	-.020

*p<.05

Table 10. *The percentages of supplying overt subjects*

Percentages Groups	[-null subject]
G	100
U	97.8
S	94.1

Acquisition differences of *there* and *it*

There and *it* are different in feature specifications (Groat, 1995; Hornstein, Nunes & Grohmann, 2005), syntactic associations with the postverbal elements (Chomsky, 2000; Park, 2005), co-occurring verbs on the part of *there*-unaccusatives and *it*-raising adjectives and past participles (Zwart, 1991, Han, 2001; Nakajima, 2001; Rothstein, 1995), and occurring frequency. Our goal is to examine whether these differences affect the acquisition pace and the choice of *there* and *it*. The data from the Blank-filling Task were analyzed in SPSS by one-way ANOVA. In the responses of *there* and *it* items, the statistical results show that the groups from the most accurate responses to the least accurate are in the order, NS>G>U>S. That is, the higher the proficiency is, the more accurate the responses are.

In addition to the comparison of each expletive across groups, we also intend to know which expletive performed better in each group. Table 11 shows the differences in correct responses between *there* and *it* in each group by one-way ANOVA in SPSS. Apparently, the correct rates of *there* and *it* in each group are significantly different. The percentages indicate that *it* is remarkably more accurate than *there*.

Table 11. *The accuracy differences between there and it in each group*

Differences Group	There in %	It in %	Sig.
N(ative) S(peaker)	77.4	99.6	.000**
G	69.5	99.4	.000**
U	61.7	98.3	.000**
S	61.7	91.2	.000**

*p<.05 **p<.01

The percentages of *there* in Table 11 surprises us in that even the NS group did not perform as well as we predicted (77.4% correct responses), and the accuracy of *there* improves very slowly for the EFL groups. The 22.6% incorrect responses (100%-77.4%=22.6%) from the NS group mainly originated from the structure of *there*-unaccusatives. Take “*There* occurred a crash in the air” as an example. They felt neither *there* nor *it* was the right answer. Although they thought *there* is the nearest answer, they did not use *there* and *occur* this way frequently, so they did not feel the combination was right¹. That is, the co-occurring frequency of *there* and *occur* controlled their performance on this Blank-filling task.

For the EFL participants, another source of incorrect responses is the pattern of *there*-raising verbs/adjectives such as “*There* is likely to be an earthquake in Japan.” This situation resulted from the fact that the connections of *is likely* and *seems* with *it* were so strong that once the EFL participants saw *is likely* and *seems*, they supplied *it* without thinking². That is, *it is likely* and *it seems* are high-frequency to the EFL groups. The source of the input is from textbooks and class instruction. A senior high teacher told the researcher that the co-occurrence of *there* and *is likely* and *seems* was rare in textbooks and class instruction, if there was any. Based on the performance of the native speakers and the EFL participants, we know that co-occurring frequency affects a person’s judgment on grammaticality of expletive usages and choices of expletives much.

Correlation Between Syntactic Properties and Expletive Accuracy

We intend to explore whether the acquisition of the syntactic properties of *there* and *it* constructions could facilitate learners’ association of *there* and postverbal DPs, and *it* and CP complements. In the following, *there* and *it* are divided into two sections.

Syntactic Properties of *there* constructions. The syntactic properties we targeted include (a) to be, (b) S(*there*)-V agreement, (c) definiteness effect and (d) stage-level predicates. The task adopted was Grammaticality Judgment, which required the participants to underline the part(s) they considered incorrect (for the samples of the test items, see Table 5). The scores of each property were correlated with that of *there*-DP. Table 12 shows the correct percentages and the correlation of *there*-DP with each property. The NS group got positive, significant correlations on the properties of *to be*, *definiteness effect* and *stage-level predicates*, whereas none of the EFL groups had positive, significant correlation between any feature and *there*-DP.

Table 12. The correlation between each property and *there*-DP

Features <i>There</i> -DP Group	To be		Definiteness effect		Agreement		Stage-level predicates		<i>There</i> - DP
	r	%	r	%	r	%	r	%	%
NS	.371*	95.0	.404*	61.7	.226	77.1	.386*	82.3	100
G	.098	70.9	.016	86.6	-.074	77.4	-.193	62.2	77.9
U	-.209	74.4	.037	85.6	-.232	69.5	-.176	60.0	75.6
S	-.110	61.8	-.033	73.5	-.304*	58.8	-.011	45.1	36.3

* $p < .05$

Note. There are four items for *there*-DP, two are grammatical and the others are ungrammatical. The grammatical ones are in the structure of *there*-unaccusative-DP. The grammaticality judgment of this structure

¹ This was some of the NS participants told me during and after the questionnaire.

² This is the reason that some of the participants I interviewed told me why they chose *it*.

was affected by the co-occurring frequency of *there* and unaccusative verbs a lot, so the scores of the two grammatical items were excluded from data analyses.

The NS group did not get significant correlation between *agreement* and *there-DP*. The main problem is the agreement of *there* and plural postverbal DPs. Take “*There seem to be George and Mary*” as an example. This is a grammatical sentence, but 70% of the NS group marked it incorrect since they thought the verb *seem* should be singular. Some of them added “s” to *seem* on the questionnaire and the others underlined *seem* to express the ungrammaticality of the verb. However, according to Chomsky’s (2000) Agree Theory and Park’s (2005) adapted Agree Theory, the agreement in *there* constructions is related to the number feature of the postverbal DPs. The inconsistency between theory and the responses of the participants is worth studying.

Based on the results, the relationships between the syntactic properties of expletive constructions and expletive accuracy of *there* are distinct in the grammar of the NS group and the interlanguage grammar of the EFL groups. *To be, definiteness effect* and *stage-level predicates* are in a clustered relationship with *there-DP* for the NS group. But these properties are separate for the EFL groups. In other words, the EFL participants learned these properties separately, which might be responsible for the slow acquiring process of *there-DP*. Furthermore, the use of *there-unaccusatives/raising* elements does not co-occur with the acquisition of *there-DP* for the EFL learners; that is, even though they adopted *there-unaccusatives/raising* elements, they have not necessarily acquired the association of *there-DP*. But the acquisition of *there-DP* shows developmental improvement as proficiency increases.

Syntactic Properties of *it* constructions. The properties we explored contain (a) S(it)-V agreement, and (b) raising verbs, adjectives and passive verbs. The task is Grammaticality Judgment too (for the samples of the test items, see Table 6). The scores of each feature were correlated with that of *it-CP*. Table 13 presents the correct percentages and the correlation of *there-DP* with each property. Only the undergraduates have significant and positive correlation between *it-CP* and raising elements, but the other groups do not. Besides, all the groups have no significant correlation between agreement and *it-CP*.

Table 13. *The correlation of it-CP and each feature, and the percentage of each feature*

<i>It-CP</i> Group	Features		Agreement		Raising verbs, adj. and P.P.	<i>It-CP</i>
	r	%	r	%	r	%
NS	-.061	93.4	-.022	94.7		90.8
G	.005	90.0	-.001	91.3		80.8
U	-.071	91.9	.339*	87.1		73.4
S	.171	80.1	.150	72.54		61.8

*p<.05

On the ground of the results, agreement does not cluster with *it-CP* or enhance the connection of *it-CP* much since subject (*it*)-verb agreement was affected by the numbers of the postverbal clausal complements, which is true across groups. Take “**It seem that the girls are sad and they are going to cry*” for example. This sentence contains two clausal complements. The correct rate of this item for the NS group is 80%; that is, 20% (100%-80%=20%) of them marked this incorrect sentence grammatical. On the other hand, the sentence, “**It turn out that Tom is the killer,*” has only one clausal complement; the correct rate is 96.7%. Almost every one of them recognized its ungrammaticality by underlining the main verb, *turn*, or adding “s” to the verb. Accordingly, the relationship of agreement between subject-*it* and verbs in *it* constructions was not completely fixed.

Moreover, the category of *it*-raising elements does not cluster with *it*-CP either. Table 13 shows that both the graduate and NS groups got over 90% correct rate on *it*-raising elements and the correct percentages are better than those of *it*-CP. The insignificant correlation might derive from *it*-CP. Following the scrutiny of the raw data, the problematic source lies in the two ungrammatical items for *it*-CP. The correct percentage in each group is listed in Table 14.

Table 14. *The correct percentages of the ungrammatical items for it-CP*

Groups	Percentages	*It seems to be a baby in the room.	*It is certain to be John here.
NS		80	83.3
G		53.5	72.1
U		37.8	60
S		17.6	39.2

One of the sources of the incorrect rate might stem from a mixture of referential and expletive *it*. A NS participant told the researcher that the sentence “*It swam little boys and girls to the shore*” could be grammatical if the *it* meant “a turtle.” Obviously, the NS group may judge grammaticality of a sentence by imagining a context for it.

Referential and expletive *it* both co-occur with raising verbs, such as *it seems* and *it is likely*. One way to distinguish them is the postverbal element. Referential *it* could appear with an IP such as *It seems to be a problem*. The *it* may mean a harsh situation, which seems to be a problem. On the other hand, expletive *it* co-occurs with a CP such as *it seems that John is a troublemaker*. The *it* does not refer to anything; it is semantically empty.

However, the distinction between referential and expletive *it* is difficult for the EFL participants since the relevant input and context are not sufficient for them to disambiguate the differences. The obscure distinction between referential and expletive *it* may contribute to why expletive *it*-DP is still existent. This may explain why 20.9% of the graduates marked the incorrect sentence, “**It occurs a crash in the air*” grammatical.

CONCLUSION

This study centers on four aspects: (a) the influence of *you* and *sihu* on the supply of expletives, (b) the triggering effect of expletives on [-null subject], (c) the acquisition differences of *there* and *it* and (d) the correlation between the syntactic properties of expletive constructions and expletive accuracy. The findings indicate that *there be* and *it seems* were highly possibly regarded as chunks corresponding to Chinese *you* and *sihu*. These chunk learning does not promote the supply of expletives in the pattern of *there*-unaccusative verbs and that of *it*-raising adjectives and passive verbs. But the chunk learning of *you-there be* could make the *there be* a potential trigger of [-null subject], supporting Yang’s (2001) conclusion.

However, when examining the expletive accuracy of the EFL participants, we found that they have not completely acquired expletive usages such as the relationships between *there* and postverbal DPs. Furthermore, the ways of the EFL and NS groups acquiring the relationship of *there*-DP is different on the basis of the findings of the correlation between the syntactic properties of *there* constructions and expletive accuracy. When the NS group acquired the relationship of *there*-DP, most of the syntactic properties were acquired too, but *there*-DP and the properties were acquired separately for EFL participants.

In this study, we found that co-occurring frequency of expletives and predicates and the adequacy of input affected the participants’ supply of expletives and grammaticality judgment. Pedagogically, we suggest that explicit classroom instructions and adequate input on

expletives are essential and may accelerate the acquisition.

REFERENCES

- Chomsky, N. (2000). Minimalist Inquires: The framework. In Martin, R., Michaels, D. & Uriagereka, J. (Eds.) *Step by Step: Essays on Minimalism in honor of Howard Lasnik* (pp. 89-155). Cambridge: MIT Press.
- Diesing, M. (1992). *Indefinites*. Cambridge, Mass.: MIT Press.
- Greene, R. L. (1986). Sources of recency effects in free recall, *Psychology Bulletin*, 99, 221-228.
- Groat, E. M. (1995). English expletives: A minimalist approach, *Linguistic Inquiry*, 26, 354-65.
- Han, J.-Q. (2001). Existential sentences in English and Chinese: towards a generative analysis, *Modern Foreign Languages* 24 (2), 143-158.
- Hornstein, Nunes & Grohmann. (2005). *Understanding Minimalism*. Cambridge: New York.
- Huang, C. T. J. (1989). Pro-drop in Chinese: A generalized control theory. In Jaeggli, O and K. Safir (Eds.) *The null subject parameter* (pp. 185-214). Dordrecht: Kluwer.
- Hyams, N. (1986). *Language acquisition and the theory of parameters*. Dordrecht: Reidel.
- Lasnik, H. (1992). Case and expletives: Notes toward a parametric account, *Linguistic Inquiry*, 23, 381-405.
- McCloskey, J. (1991). There, it and agreement, *Linguistic Inquiry*, 22, 563-67.
- Nakajima, H. (2001). Verbs in locative constructions and the generative lexicon, *Linguistic Review*, 18, 43-67.
- Park, I. (2005). A study on structures and agreement of infinitives, *Studies in Generative Grammar*, 15 (2), 209-230.
- Rizzi, L. (1986). Null objects in Italian and the theory of pro. *Linguistic Inquiry*, 17, 501-57.
- Rothstein, S. D. (1995). Pleonastics and the interpretation of pronouns, *Linguistic Inquiry*, 26, 499-529.
- Yang, Xiaolu. (2001). Triggering data in L2 acquisition of English by Chinese learners, *Foreign Language Teaching and Research*, 33 (1), 48-56.
- Zwart, C. J.-W. (1991). Expletive raising and expletive replacement in Dutch. *Proceedings of the Eastern States Conference On Linguistics*, Baltimore, the U.S.A, 8, 393-404.
- 張文哲譯 (2005)。教育心理學:理論與實務。台北:臺灣培生教育出版:學富文化發行。(Slavin, R. E., 2003)

A Holistic English Classroom through Drama: Applied Drama Techniques, Process Drama, and Readers Theatre

Liang-fong Hsu (徐良鳳), Shin-mei Kao (高實玫), Gary Carkin

Tainan U. of Technology
National Cheng Kung University
Southern New Hampshire University

td0011@mail.tut.edu.tw

This study explored the effect of using drama as a learning tool on the improvement of college students' English ability after a three-week intensive course. Applied drama techniques were used in the first week as orientation activities to familiarize students who had not been exposed to this kind of learning method, such as verbal games, body movements, mimes, strategic interaction and role-play. The second week was devoted to learning through three process dramas. Process drama refers to educational drama that is created by the participants. It may make use of a story's beginning, middle, and end to explore, develop, and express ideas and feelings through dramatic enactment. It is, however, always improvised drama, participant centered, and not intended for an audience. In the third week, the students used Readers Theatre format to perform a script created from their learning and improvisation content in the second week.

The participants were evaluated by GEPT Intermediate level test before and after the course as the pre- and post-study evaluation and the results of these underwent statistical analysis. The results showed that the participants had made significant progress only in speaking ability, with trivial difference in listening, reading and writing abilities.

Below are some key comments by the participants:

1. Students believed that learners with low-proficiency and low-confidence could benefit greatly from drama;
2. They thought drama provided them with a broad range of opportunity in learning English;
3. Overall, they thought drama could motivate them to, and give them confidence in, learning English.

INTRODUCTION

The traditional way of separating the four skills in English learning curriculum is considered dogmatic and difficult to implement in actual practice either in the classroom or in real life. Only a holistic teaching and learning method can reflect the actual language practice. Because drama fosters a virtual environment which gives meaning to language practice, it provides a holistic instruction and learning methodology.

New techniques of teaching English through drama emerge constantly and adaptations of older methods evolve. In this study, the authors combined a variety of applied drama techniques, process drama and readers theatre to explore the impacts on the participants when such innovative approaches were integrated in an intensive course. In general, Taiwanese college students, especially technological college students, are not familiar with drama; they tend to shun it. They think they might be required to "act." Therefore, it is crucial to orient the students to drama through various dramatic activities in the beginning, such as verbal games, body movements, mimes, story-building, strategic interaction, and role-play. The applied drama techniques were structured by levels of sophistication and difficulty and later incorporated in the dramatic structures of the process drama and readers theater.

Process drama offers many opportunities for negotiation and input of the participants. Essentially, it is the product of a collaborative meaning-making process by all the participants through the medium of role. In the procedure, the participants are constantly required to listen, speak, read, and write attentively in order to respond to each other and contribute ideas to the story or task on hand. One of the three process dramas derived in this study is developed into a performing script for the Readers Theatre format in the third week. Readers Theatre was included in the study because it can enhance the participants' reading fluency through repeated reading of the script for an anticipated performance. From repeated rehearsing, the participants learn to enunciate and articulate their lines.

This study explored the effects of using drama as a learning process for the development of English reading, writing, listening, and speaking skills. The subjects were Taiwanese university students of low to intermediate proficiency who were inexperienced in the use of drama to learn language. Applied drama techniques were used in the first week to familiarize students with drama exercises and methods. The second week was devoted to learning English through three process dramas. In the third week, the students used readers theatre format to perform a script developed through the use of the process drama of the second week.

What Is Process Drama?

Process drama, a type of educational drama, differs from drama methodologies in which the performance of a play is the focus. Heathcote (1984, p.49) defines the educational drama thus: "A broad definition of educational drama is "role-taking", either to understand a social situation more thoroughly or to experience imaginatively via identification in social situations. Dramatic activity is the direct result of the ability to role-play – to want to know how it feels to be in someone else's shoes."

The term process drama (Haseman, 1991; O'Tool, 1992; O'Neill, 1995) is a relatively new term in drama education. O'Neill (1995) used the term to describe the "pre-text" and the development of a drama created by the participants. The teacher starts the session with a situation but leads to a discussion about what had led up to the situation. The participants gradually build up a background. Taking time to work this way, the group of participants expands a body of information, deepens understanding, and develops a richer and more interesting interpretation and enactment. Kao & O'Neill (1998) further explain that process drama is extended over time and built up from the ideas, negotiations, and responses of all the participants in order to cultivate social, intellectual and linguistic development.

Stinson and Freebody (2006, p.29) believe that, while each process drama is carefully pre-planned, the structure offers many opportunities for negotiation and input by the participants. Process dramas are designed to offer chances for individuals and groups to contribute to the dramatic action, to solve problems and to employ higher-order thinking processes. They do not lead to performances of the devised materials to audiences who are outside of the drama.

What Is Readers Theatre?

Readers theatre developed after World War II from the speech and drama fields of oral interpretation and conventional theatre (Adams, 2003) and differs from traditional plays in that the readers (or performers) do not memorize lines but read directly from the script. It was originally developed by and for adults but was quickly adapted for use with older children. Most readers theatre performances are arranged to have the actors stand or sit on a stage or other performance area (Kerry Moran, 2006, p.317). No attempt is made to hide the scripts that the performers hold in folders or place on lecterns in front of them. Lines are distributed among individuals, pairs, small groups, and the whole group (Flynn, 2004, p. 360).

Readers theatre is a minimal theatrical performance requiring students to express meaning through fluent and prosodic readings of scripted stories, poems, chants, and rhymes. It does not require the use of props, costumes, or scenery. Staging and movement is kept to a minimum. Students' voices are the only tool used to communicate meaning or to bring character to life. There is no attempt to create a sense of reality on stage because it is co-constructed between readers and the audience. The readers participate in this joint project by suggesting interpretations through fluent or expressive reading while the audience uses his or her imagination to complete the meaning making.

Fluency gains may be associated with readers theatre participation (Millin & Rinehar, 1999). One possible explanation for the improvement is that readers theatre demands repeated reading during the rehearsal stage for an anticipated performance. The desire to put on a performance is more motivating than simply requiring that a story be read a prescribed number of times. Most importantly, the increased reading fluency does not end with texts that have been practiced but carries over to new and unpracticed texts (Tyler & Chard, 2000).

RESEARCH METHOD

Participants

The participating students were 32 female university students who volunteered to take a special intensive summer course entitled Holistic English at Tainan University of Technology. They came from the four year, two year, five year, seven year and graduate programs of various departments, including music, accounting, childcare, finance, international trades, foreign languages, fine arts, product design, fashion design, and styling & cosmetology. The student's level of English varied from high beginner to intermediate.

Course Design

The course was a three week intensive course which met for 3 hours a day, five days a week, for a total of 45 hours of class time. The first week was devoted to the use of applied dramatic techniques, the second week involved three process dramas, and the third week was dedicated to readers theatre. The work of each week built upon that of the previous; thus, the applied drama techniques were structured by levels of sophistication and difficulty and later incorporated in the dramatic structures of the process drama. Likewise, the readers theatre incorporated a story and some of the work developed during the second week of process drama methodology. The course was co-conducted by Dr. Shin-Mei Kao, a process drama expert, and Dr. Gary Carkin, who specialized in readers theatre and directing of theatre productions, with the researcher-observer, Dr. Liangfong Hsu, in the classroom.

The First Week: Applied Drama Techniques. The various drama techniques include games for concentration, pronunciation, image building, body movement, verbal exercises, story building, dramatic reading of short plays, strategic interaction, structured improvisation, and process drama episodes.

The Second Week: Process Drama. There are three carefully pre-planned processes. The first one is *The Seal Wife*, an Irish folk tale, taken from the book entitled *Pretext and Storydrama: The Artistry of O'Neill and Booth* by Philip Taylor (1992). The picture book used is *The Seal Mother* by Mordecai Gerstein (1986). The instructor has designed 9 steps in the structure:

- Step 1. The pre-text is launched;
- Step 2. Transforming the pre-text;
- Step 3. Forming the community;
- Step 4. The family;
- Step 5. The private dream worlds;
- Step 6. Ten years after the wife has left;

- Step 7. Patrick's song;
- Step 8. A thousand years later.

The second process drama is based on the picture book entitled *Tuesday* by David Weisner (1994).

- Step 1. Creating the pre-text;
- Step 2. Investigating the case;
- Step 3. Forming a village meeting;
- Step 4. Meeting the strange creature;
- Step 5. Interviewing the monster;
- Step 6. Report writing.

The third process drama is *The Lost Tribe* based on "The Lost Valley" in *Drama Structures: A practical Handbook for Teachers* by Cecily O'Neill and Alan Lambert.

- Step 1. Exploring the pre-text;
- Step 2. Organizing an exploration team;
- Step 3. An Official Photo;
- Step 4. Arriving in the Valley;
- Step 5. Life in the valley;
- Step 6. Returning Home.

The Third Week: Readers Theatre. The class work now turned to the construction of a readers theatre piece to be prepared for performing and videotaping as a final product. The *Seal Wife* was selected as being appropriate because the students had enjoyed working through the material in terms of the process drama and the material itself contained many moments of dramatic tension. The story is enlarged by incorporating the end of the story and additional narration written by the workshop leaders. The presentation of the *Seal Wife* was created according to the following steps:

- Step 1. The Introduction;
- Step 2. Integrating the scenes;
- Step 3. Polishing the spoken word;
- Step 4. Developing performance style;
- Step 5. Developing the flow;
- Step 6. Developing timing and flow & polishing language;
- Step 7. Performance.

Data Collection

The students were evaluated by GEPT Intermediate level test before and after the course as the pre- and post-study evaluation. The GEPT test contains four parts: listening, reading, speaking, and writing. Both the listening and reading sections contain multiple choice questions only and thus, were scored by computer. The oral test contains reading aloud, answering short questions, and describing a picture. The writing test contains a Chinese-to-English translation task and a short essay-writing task. The latter two tests were evaluated manually by two trained instructors. Two independent raters measured the quantity and quality of student oral and written production. The two raters received training regarding the scoring standards and procedures prior to their actual scoring. Their evaluation has achieved .92 inter-rater reliability. The average between the two sets of scores serves as the final scores of the students.

The Results

All analyses were carried out using S-Plus V 6.2 statistical software. Statistical significance was accepted at $p < 0.05$ for all comparisons. A student's t-test was performed to highlight score differences between pre- and post-test for each of the individual measures

(listening, reading, speaking, and writing) and for each of the sub measures (words, sentence, clause, mean length of sentence, and mean length of clause) of writing. Table 1 shows the numbers, means, standard deviations, and *p* values for each of the individual measures and Table 2 shows the numbers, means, standard deviations, and *p* values for each of the sub measures of writing.

Table 1. Means, SDs, and *P* values for each of the individual measures

	Listening		Reading		Speaking		Writing	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
n	20	20	20	20	17	17	14	14
Mean	72.6	66.8	49.6	48.2	35.1	43.6	15.8	19.4
Standard Deviation	23.2	22.8	15.2	12.9	18.3	19.1	10.1	11.4
<i>p</i> value	0.088		0.572		0.044*		0.154	

Note. **p*<.05.

The results in Table 1 show that the participants made a significant progress in speaking (*p*=0.044). Though their writing skill has improved after receiving the instruction, the progress has not reached a significant level (*p*=0.154). Surprisingly, the students' performance in listening and reading did not show any progress.

To further examine the students' progress in writing, we performed the paired t-tests on the following five subsets in writing quantity: total words, total sentences, total clauses, mean length of sentence, and mean length of clause.

Table 2. Means, SDs, and *P* values for each of the sub- measures of writing

	Words		Sentences		Clauses		MLS*		MLC**	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
n	18	18	18	18	18	18	18	18	18	18
Mean	61.1	56.8	4.9	6.8	8.9	10.4	11.1	6.8	5.6	4.3
Standard Deviation	45.4	43.5	3.2	5.2	5.8	7.5	7.6	4.5	2.7	2.4
<i>p</i> value	0.727		0.030*		0.381		0.035*		0.126	

Notes. **p*<.05.

MLS= mean length of sentence

MLC= mean length of clause

Table 2 shows that the participants had written significantly more sentences in the post-writing test (*p*=0.030) but the mean length of sentence has decreased significantly (*p*=0.035). No significant differences were found among the other three subsets of writing quality. It is noted that the participants had produced fewer words in total in the post-writing test. Since their writing scores, which measure their writing quality, are higher in the post test, we logically deduce that the students could express their ideas more effectively with fewer words and more sentences.

DISCUSSION

Although the study shows that significant improvement in speaking was made by the end of the program, the students' reading, listening and writing skills did not have any significant progress. To the surprise of the researchers, even though the students have made slight improvement in writing, their listening and reading scores have decreased in the post

tests. There are some possible explanations. First, this course emphasized fluency over accuracy in speaking and writing, and did not include grammar instruction. The reading materials for the class were selected to facilitate the development of drama. In this manner, the students were not taught specifically for taking a standardized test such as the GEPT used for the pre- and post-evaluation.

Other explanations are chiefly due to the technical difficulties and limitations of the study.

Technical difficulties

The students' English level. Because this course is a free course in the summer, it is difficult to get sufficient number of students with intermediate English level. The 32 students recruited are of high beginning to intermediate levels.

The selection of GEPT test level. Since the students' English level varied from high beginning to intermediate, the selection of GEPT test level is difficult. In this study, the intermediate level test of GEPT was implemented. However, it posed some difficulties for the students of the study, especially the reading test. It requires a certain amount of vocabulary in order to comprehend the test articles. No wonder some of them simply gave up the reading test in the middle without completing the test, especially after the physically demanding performance (the post-test was taken immediately after the final production).

Technical flaw involved in testing. Due to mechanical problem, the students' oral answers were not always recorded successfully or the students' marked answers sometimes were not registered on the computer, which resulted in incomplete data for analysis.

Limitations of the study

Students' relatively lower English proficiency. The students were of lower English proficiency, inviting the use of their first language when planning, negotiating, and designing presentations. Although the *delivery* of the exercises was done in the target language, much learning time was reduced by the student's inability to explain, instruct, direct, and question using English during the planning stages of the exercises because of their relatively lower English language proficiency.

The short duration of the study. Since the students meet every day, there is not enough time for doing research, through which the students can build reading & writing abilities.

Suggestions for further studies

Analysis might involve examining how skills develop through the use of drama starting with beginners and moving through advanced levels of teaching English through drama. In addition, more comparison studies are needed that compare the effects of the use of various applied drama techniques to those of process drama and other genres of drama in language education.

Additional Findings

In order to better understand the positive effect of this type of project, it is important to take into consideration what the participants themselves had to say about it. The following characteristic comments were given by the participants:

1. Students with low proficiency and low confidence can benefit greatly from drama activities, just like the "good" students.
2. Students can benefit from participating in group-work as well as in pairs.
3. Drama provides them with a broad range of opportunities in learning English.
4. Drama motivates them to learn English and gives them more confidence to learn English.

CONCLUSION

Learning English through drama facilitates the implement of four skills of English. In the structure of a combination of applied drama techniques, process drama and readers theatre, the participants are constantly interacting with each other and actively contributing ideas to the creative tasks on hand. In the process, they need to express themselves clearly, listen to other people attentively, reading stories and information, and writing dialogues, scripts, and journals. In the final readers theatre performance, the participants improve their fluency through repeated reading practice. Most of the participants found that a holistic English classroom through various drama methodologies is fun, motivational and can boost their confidence in learning English. As a matter of fact, the participants have improved significantly on speaking and made a slight progress on their writing within a short span of three weeks. With adjustments in course design, there is a great potential for enhancement on the other two skills as well, which awaits further research to prove.

REFERENCES

- Adams, W. (2003). *Institute book of readers theatre: A practical guide for school, theatre, & community*. Chapel Hill, NC: Professional Press.
- Flynn, R. M. (2004). Curriculum-Based Readers Theatre: Setting the Stage for reading and retention. *International Reading Association*, 360-365.
- Gerstein, M. (1986). *The Seal Mother*. New York: Dial Books.
- Haseman, B. (1991). Improvisation, Process Drama and Dramatic Art. *London Drama*, Jul, 19-21.
- Kao, S., & O'Neill, C. (1998). *Words into worlds: Learning a second language through process drama*. Stanford: Ablex Publishing Corporation.
- Kerry Moran, K. J. (2006). Nurturing Emergent Readers Through Readers Theatre. *Early Childhood Educational Journal*, 33(5), 317-323.
- Millin, S. K., & Rinehart, S. D. (1999). Some of the benefits of readers theatre participation for second-grade Title I students. *Reading Research and Instruction*, 39(1), 71-88.
- O'Neill, C., & Lambert, A. (1982). *Drama Structures: A Practical Handbook for Teachers*. London: Hutchison.
- O'Neill, C. (1995). *Dramaworlds: A framework for process drama*. Portsmouth, NH: Heinemann.
- O'Toole, J. (1992). *The process of drama: Negotiating art and meaning*. London: Routledge.
- Taylor, P. (1992). *Pretext and Storydrama: The Artistry of O'Neill and Booth*. Brisbane, Australia: NADIE Publications.
- Stinson, M., & Freebody, K. (2006). The Doll Project: The Contribution of Process Drama to Improved Results in English Oral Communication. *Youth Theatre Journal*, 20, 27-41.
- Tyler, B., & Chard, D (2000). Using readers theatre to foster fluency in struggling readers: A twist on the repeated reading strategy. *Reading & Writing Quarterly*, 16, 163-168.
- Weisner, D. (1994). *Tuesday*. New York: Clarion Books.

A Study of English Listening Proficiency and English Oral Proficiency

Yu-chieh Hsu (徐于捷)
National Chung Cheng University
stubbornhsu@gmail.com

The purpose of this study is to investigate if high English listening proficiency leads to high English oral proficiency and also to examine possible speaking difficulties for EFL speakers. The participants in this study were students from different majors, who had taken TOEIC in the past two years (2006/12-2008/12). Two sets of data were collected: listening score and speaking analysis reports. Listening data were collected by selecting participants' TOEIC listening test report and speaking data were from the oral practice reports generated by English oral training software, MyET. Through quantitatively comparing and analyzing participants' listening scores and their oral analysis reports from MyET, the research question about whether high English listening proficiency leads to high English oral proficiency and possible speaking difficulties could be verified. The results of this study showed that higher listening proficiency students demonstrated higher speaking proficiency, which implies that better oral proficiency comes after better listening proficiency. The speaking difficulties in EFL speakers are intonation and fluency, which also indicates the importance of listening ability which helps EFL learners familiarize themselves with the speech intonation and the target language patterns.

INTRODUCTION

According to Bickerton (2000), "languages are naturally acquired by people listening to language, and the human brain is built to analyze it." For L1 speakers, listening is the first learned receptive skill before speaking which is a productive skill. It shows that children can usually recognize and understand a word long before they can use it in speech production (Clark & Hecht, 1983). The idea is supported by psycholinguistic view that our brain is automatically programmed to take words and put them together. Even though each language has slightly different patterns, linguists regard them basically as dialects of one language. Hence, all EFL learners have already set a machine that is ready for acquiring for language in the brain, and it is only the sound system that learners should put efforts on.

Speech refers to the processes associated with the production and perception of the sounds used in spoken language. The process of perceptive skill and productive skill are connected ability. Fodor, Bever, and Garrett (1974) report that the improvement of learner's perception can boost up their production and consider that the process between perception and production is interrelated. Therefore, it is reasonable to assume that English learners' listening proficiency can serve as an effective indication of their speaking ability.

Controlled processing appears to play a central role in lexical access and articulation in a second language, at least until a high level of proficiency has been achieved. L2 speech tends to pause much more frequently, consist of shorter utterances, and contain many more slips of the tongue than L1 speech (Poullisse, 1997; Weise, 1984; Möhle, 1984; Lennon, 1990). However, as second language speakers become more fluent, speech rate and length of run increase and the number of filled and unfilled pauses decrease (Lennon, 1990). The assumption is that fluency is a direct function of automatic language processing ability.

According to Hudson (2006), Baddeley, A., and Logie, R. (1999), language acts as a networks with cognitive psychology and psycho-linguistics, which indicating the complex

components of language. The purpose of this research is to investigate whether English speakers' oral proficiency could be expected from their listening proficiency and could identify certain oral difficulties which might affect English learners' oral performance. The researcher assumes that even though the process of perceive a linguistic input is complex, the relationship between linguistic perception and production can be found through analyzing the observable linguistic production on certain linguistic performance. From psycholinguistic viewpoint, the relationship among listening, writing, and reading is not direct; whereas the relationship between listening and speaking shows strong correlation to some extent. The researcher hypothesizes that English learners' oral proficiency could be expected from their listening proficiency and vice versa. Furthermore, according to controlled-automatic processing model, fluency is assumed to be a difficult element for most speakers who are not likely to process English without certain controlled psychological processing. The present study aims to answer the questions of whether high speaking proficiency can be anticipated from high listening proficiency and what are the possible suprasegmental features in oral difficulties EFL speakers might encounter.

Research questions

The research of this study is conducted to answer the following questions:

1. How can speaking proficiency expected from listening proficiency?
2. What are the difficulties of oral suprasegmental features that EFL speakers might encounter?

The Study

The significance of this study lies in the application of MyET to examine EFL learners' oral proficiency. MyET can point out which word and syllable one has pronounced wrongly. It also provides the correct pronunciation with the 3D animation to show the shape of the mouth and tongue position word by word for individual learners. MyET is an English oral training software based on Dr. Pimsleur's English learning theory. Automatic Speech Analysis System (ASAS) is applied to assess speakers' oral performance. In this study, the application of MyET not only avoids the subjective rating in EFL students' oral production, but also provides EFL students a hands-on experience of using MyET. Furthermore, the findings and implications of this study might provide further ideas for both learners and teachers. Especially for learners who want to receive immediate feedback and to schedule their learning at will outside of language classrooms. To operationalize the key terms for the research, the definitions are given as following:

1. Listening proficiency

The listening proficiency is defined as the ability of English listening for understanding conversation and communication in daily life.

2. Oral proficiency

The oral proficiency here is referring to the ability of English speaking for conversation and communication.

3. TOEIC

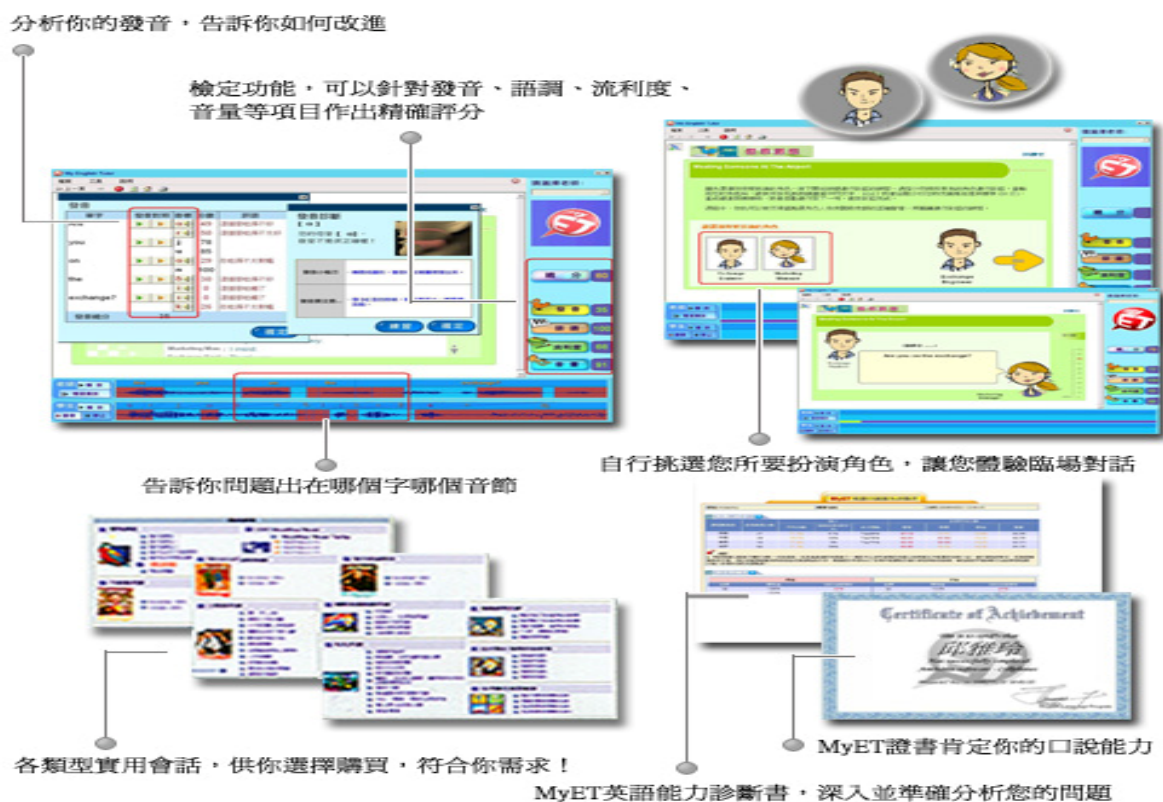
According to the official TOEIC web site³⁰, TOEIC (Test of English for International Communication) test is an English language proficiency test for people whose native language is not English. The test includes listening and reading sections, all in multiple choice questions. The total score is 990; each section is 495 with 100 multiple choice question scored 5 on each item with no deduction. A TOEIC score of 450 is frequently considered acceptable for hiring practices, with the understanding that the employee will continue English studies. A TOEIC score of 600 is frequently considered the minimum acceptable for working overseas.

³⁰ <http://www.toEIC.com>

4. MyET

MyET is an English oral practice software to analyze speakers' oral proficiency by applying Automatic Speech Analysis System (ASAS). MyET provides standard pronunciation for speakers, and then subscribes the speakers' oral speech records to ASAS for analyzing. ASAS will analyze learners' oral performance by comparing the recordings with native instructors' in terms of pronunciation, intonation, fluency, volume, and the overall oral performance. The courses in MyET range from daily conversation, business English, to academic English, and learners can select the course according to their own levels. The passing score of each unit is eighty. The oral practices in this research are from the daily conversation textbook, "Getting around in American English 3A." The level of difficulty is above upper intermediate level of GEPT. The following illustrations are MyET oral analysis reports.

The illustrations of MyET oral analysis reports



LITERATURE REVIEW

The literature review here includes general reviews on five main concepts: the relatively subjective standard setting in oral evaluation, input and output hypothesis, information processing theory, the learning model of automatic and controlled processing, and Dr. Pimsleur's oral learning principle.

The relatively subjective standard setting in oral evaluation: A review of standard setting for next generation TOEFL Academic Speaking Test (TAST)

The ETS standard setting assumes speaking competency is a continuum rather than a state, and decisions follow a judgment-empirical process (Berk, 1986). In a judgmental-empirical process, the judges' decisions are based on performance data. Standard setting is not a task for the meek, or as Berk (1986, p. 137) states, "The process of setting performance standards is open to constant criticism and remains controversial to discuss, difficult to execute, and almost impossible to defend." Since oral performance is relatively

subjective, it is difficult to obtain agreement on an exact placement. Competence has been described as a continuous variable and as such "there is clearly no point on the continuum that would separate students into the competent and the incompetent" (Jaeger, 1989, p. 492).

There are two ways to rate learners' oral performance according to Faggen (1996): a benchmark method focuses on score level descriptors; whereas others are operating on a test level pass or fail method which focuses on examinee responses. However, it is not clear whether providing the panelists with the scores for examinees' responses will serve as a biasing factor early in the judgmental process, or whether the presence of scores later in the process would help alleviate random error or unwanted bias. (Hambleton et al., 2000, p. 364). In TOFEL-iBT speaking section, the raw scores for each examinee were converted to a scale of 30 to eliminate the more difficult-to-rate examinees, whose responses fluctuated between score bands, from the pool of responses used in standard settings.

Input and output hypothesis

On the applied side of second language acquisition (SLA) theory much of the debate over what promotes competence has focused on the role of input in language learning. It has even been argued that input is the greatest sole determiner of language acquisition (Krashen, 1985). However, there is evidence that input alone is not sufficient to obtain high levels of proficiency in a second language. Swain's proposal of the Output Hypothesis places an emphasis on language learners "noticing" the gaps in their linguistic knowledge as a result of external feedback (e.g., clarification requests, modeling, and overt correction) or internal feedback (monitoring) of language they have produced. By becoming consciously aware of one's own language production, output can serve the metalinguistic function of helping to internalize linguistic forms, test hypotheses about the language, and increase control over previously internalized forms.

Information Processing Theory

Two basic psychological concepts are used to understand the mind's construction of meaning from language input: bottom-up and top-down processing. From the language teaching perspective, top-down processing enables listeners to understand the general meaning of communicative input without needing to understand all of the grammar forms or vocabulary. Top-down processing enables listeners to comprehend meanings fluently yet fails to develop accuracy in producing target language. Another information processing theory is defined by Rumelhart and Norman (1978) and Schneider & Schiffrin (1977). Rumelhart and Norman raise the idea that learners exhibit a distinct tendency to interpret new information in terms of their existing knowledge structures, often refer to as schemata. Besides, Schneider & Schiffrin propose that previously stored information can be processed either via controlled processing (i.e. processing requiring attention and awareness) or via automatic processing (i.e. processing that is not controlled, modified, or inhibited). Automatic processing differs from controlled processing since it is not capacity limited, and thus information could be carried out in parallel, whereas controlled processes can be carried out only one at a time. Therefore, automatic processing can assist speakers to perform a more fluent speech since automatic processing is built upon already familiar linguistic structure and knowledge.

Controlled-automatic processing

In 1992, de Bot employed Levelt's model together with Anderson's (1982) notions of declarative and procedural knowledge as means of analyzing the notions of the Output Hypothesis from a psycholinguistic perspective. De Bot limited his discussion to lexical access and how it relates to the shift from controlled (declarative knowledge) to automatic processing (procedural knowledge)—a process referred to as restructuring. The crux of de Bot's argument was that output plays a crucial role in the restructuring of linguistic forms into procedural forms allowing for automatic and efficient performance. As second language speakers become more fluent, speech rate and length of run increase, and the number of filled

and unfilled pauses decrease (Lennon, 1990). However, according to de Bot, output does not play a role in the acquisition of declarative knowledge itself. Unfortunately, process models that could suggest causal mechanisms have not guided research on the role of output in acquisition. Employing process-based working models has the distinct advantage of allowing researchers to make specific predictions about the performance of second language speakers under specific task requirements.

Pimsleur learning principle and principle for analyze oral proficiency

MyET is based on Pimsleur's principle. Pimsleur's learning principle was developed by using four principles, which he regarded as important principles to form memory associations and language recall. Besides, based on linguistics views, the principles for evaluating speech production are according to certain linguistic features. The following will be the four Pimsleur's principles and the linguistic features for evaluating oral production.

First, anticipation: language courses commonly require a student to repeat after an instructor, which Pimsleur argued was a passive way of learning. Pimsleur developed a "challenge and response" technique, where a student was prompted to translate a phrase into the target language, which was then confirmed. This technique creates a more active way of learning, requiring the student to think before responding. Pimsleur pointed out that the principle of anticipation reflected real life conversations where a speaker must recall a phrase quickly. Second, graduated-interval recall: it is a method of reviewing learned vocabulary at increasing longer intervals. It is a version of retention through spaced repetition. The goal of this spaced recall is to help the student to move vocabulary into long-term memory. Third, core vocabulary: the Pimsleur method focuses on teaching commonly used words in order to lead to a comprehensive understanding of a "core vocabulary". In the typical Simon & Schuster 60 cassette/CD course (four modules of 15 cassette/CDs each) this does not provide a large breadth of vocabulary. However, word-frequency text analyses indicate that a relatively small core vocabulary accounts for the majority of words spoken in a particular language. The Pimsleur method never teaches grammar explicitly, instead leaving the student to infer the grammar through common patterns and phrases repeated over and over. Fourth, organic learning: The program uses an audio format because Dr. Pimsleur argued that the majority of students of languages wanted first and foremost to learn to speak and understand. This auditory skill, learned through their ears and mouths, is a very different skill to the visual one of reading and writing. Pimsleur argued that these two independent skills - audition and vision - should not be confused. He referred to his auditory system as "organic learning," which entails studying grammar, vocabulary, and pronunciation simultaneously. Learning by listening is also intended to teach the proper accent, which of course cannot be learned through written material.

Based on linguistic views, the features which analyze oral production are being organized into systems for contrasting audition terms of pitch, loudness, duration, and silence, either singly or in combination; and they expound contrasts in meaning of an attitudinal, grammatical, or social kind. The prosodic systems recognized are pitch-direction, pitch-range, loudness, tempo, rhythm, and pitch.

METHODOLOGY

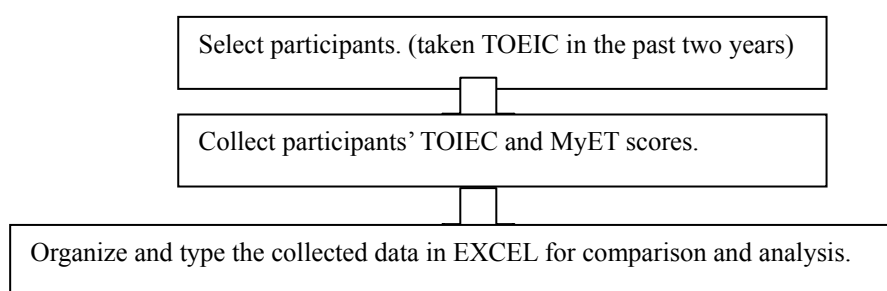
This section describes the overall research design of the study. First, the research method of this study will be described. Second, a chart will display the research procedure to show how the research was being conducted. Third, the participants' backgrounds and their level of listening and oral proficiency will be described, and the instrument applied in this study will also be introduced. Fourth, the scores of participants' oral practice average score and listening score will be compared to examine whether speaking proficiency can be expected from listening proficiency. Besides, the subset scores in participants' oral performance will be

displayed in figures to discover the difficulties of oral suprasegmental features that EFL speakers might encounter.

Research method

Listening proficiency is assessed by participants' TOEIC listening scores, and the oral proficiency is assessed by participants' MyET scores. By collecting participants' TOEIC listening scores and oral practice scores, researcher first categorize the level of participants' listening ability and then compare their listening proficiency to their oral performance. The subset scores in oral practice reports by MyET will be demonstrated with bar chart to show how participants perform in each subset of oral practices. The data of this study are collected, displayed, and analyzed with basic descriptive statistic calculation with Excel.

Research procedure



Data collection

Participants

The participants of this research are ten adults from different institute and departments: one graduate student, two post-graduates, and seven undergraduate students. All participants took TOEIC (Test of English for International Communication) in the past two years. Since eight out of ten participants majored, majoring, or double-majoring in the department of foreign languages and literature, their level of English listening and oral proficiency are estimated to be above the average university students. According to the listening score report from TOEIC official website, the average listening score in Taiwan is 285. In this study, five participants got full credits in TOEIC listening (495), three above 400, and two above 360.

Instruments

One instrument was applied in this study: MyET English oral practice software. The levels of the two selected units in MyET are from GEPT high-intermediate level in accordance to the participants' English proficiency. The topics of the two units are: "Using active listening" and "Getting recommendations 2," which are common topics in real life conversation with long dialogues. The study applies average score of the two units to reduce possible variables or distracters in participants' speaking performance.

Data analysis

The data of TOEIC listening scores and the average score of two MyET total oral scores were analyzed by applying descriptive statistics to find the relationship between students' listening proficiency and oral proficiency. Since the standards of oral rating are rather subjective, the rating system of MyET, ASAS, set a standardized rating scale to mark learners' oral proficiency on the same criteria. In order to make participants' oral proficiency more validated, the average score of two oral practices was used in the analysis instead of only one oral practice score to reduce possible variables or distracters.

RESULTS

To answer the research question on the expectation of oral proficiency from listening proficiency and the difficulties in EFL oral performance, the result of this study shows that high listening proficiency leads to high oral proficiency. The result also shows that there are

indeed certain speaking difficulties for EFL learners, which impeded English learners to reach better oral proficiency. The participants' scores on listening and oral practices are presented in the following table. (See Table 1). Furthermore, Table 1.1 is the subset scores from the oral practice report, which demonstrates participants' speaking difficulties. The scores in bold are those below the passing score of MyET (score 80). (See Table 1.1)

Table 1. *The scores of oral practices and TOEIC listening & their correlation*

	MyET (Book 3A) Unit 1.5	MyET (Book 3A) Unit 3.1 B	Average of the two scores	TOEIC listening score
Participant 1	81.57	84.43	83	495
Participant 2	82.07	85.07	83.57	455
Participant 3	88.36	88.86	88.61	495
Participant 4	80.86	84.9	82.88	495
Participant 5	73.71	80.64	77.175	380
Participant 6	82.14	82.5	82.32	475
Participant 7	76.93	76.48	76.705	385
Participant 8	82.79	85.86	84.325	430
Participant 9	81.43	78.14	79.785	495
Participant 10	85.71	80.07	82.89	495

Figure 1 The distribution of listening and speaking proficiency

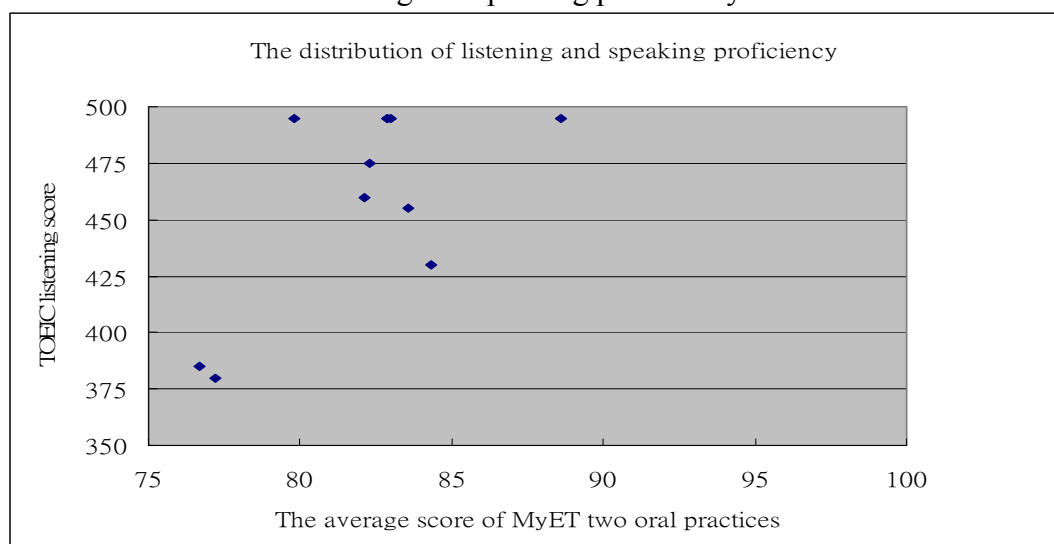


Table 1.1 *The subset scores of the oral practice report*

Select unit: Using active listening										
Pronunciation	87.5	86.57	87.64	81.1	86.57	88.21	86.93	83	89.64	82.5
Intonation	75.57	75.71	54	64.14	84.36	84.21	74.29	68.57	82.57	79.71
Fluency	72.21	79.5	66.64	75.52	71.29	84.14	73.14	71.93	79.07	74.36
Volume	87.21	80.71	88.93	80.33	83.93	80.21	83.5	83.86	85.79	78.21
Total score	82.14	82.5	76.93	76.48	82.79	85.86	81.43	78.14	85.71	80.07
Selective unit: Getting recommendations 2										
Pronunciation	86.21	86.5	85.64	86.57	87.86	89.38	90	89.81	86.86	89.71
Intonation	71.5	84.79	75	83	74.79	75.62	87.21	84.48	51.5	63.29
Fluency	78.79	80.79	79.93	83.43	66.57	83.62	85.57	91.57	57.14	75
Volume	84.86	81.07	82.21	85.43	84.71	83.29	86.14	85.57	84.71	81.29
Total score	81.57	84.43	82.07	85.07	80.86	84.9	88.36	88.86	73.71	80.64

Figure 1.1 Subsets oral score distribution 1

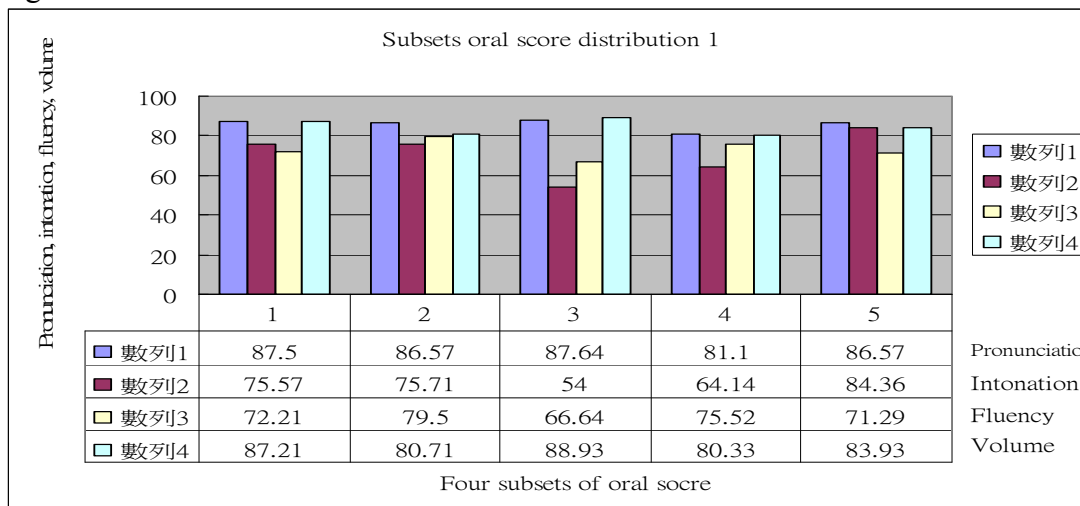


Figure 1.2 Subsets oral score distribution 2

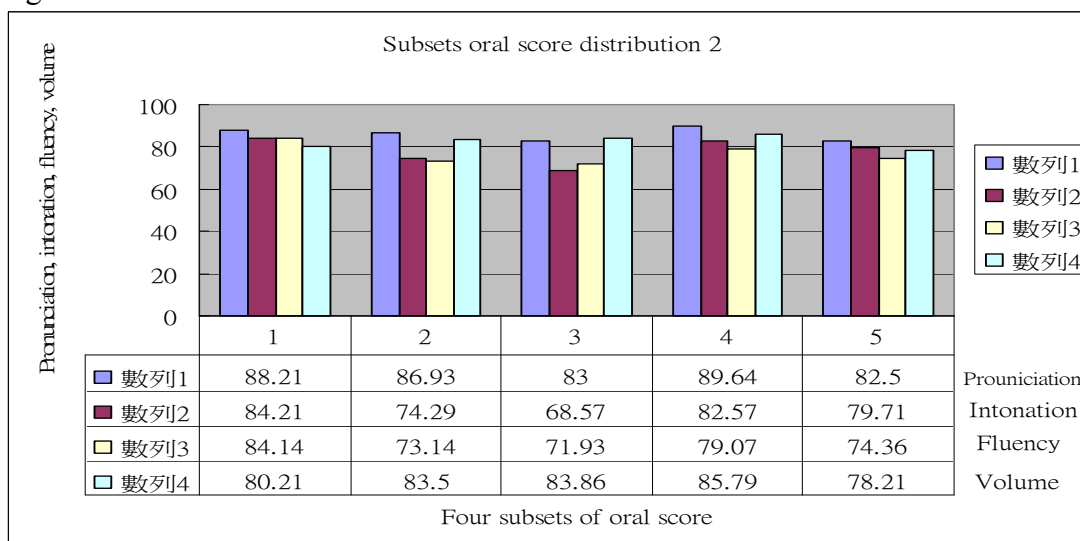


Figure 1.3 Subsets oral score distribution 3

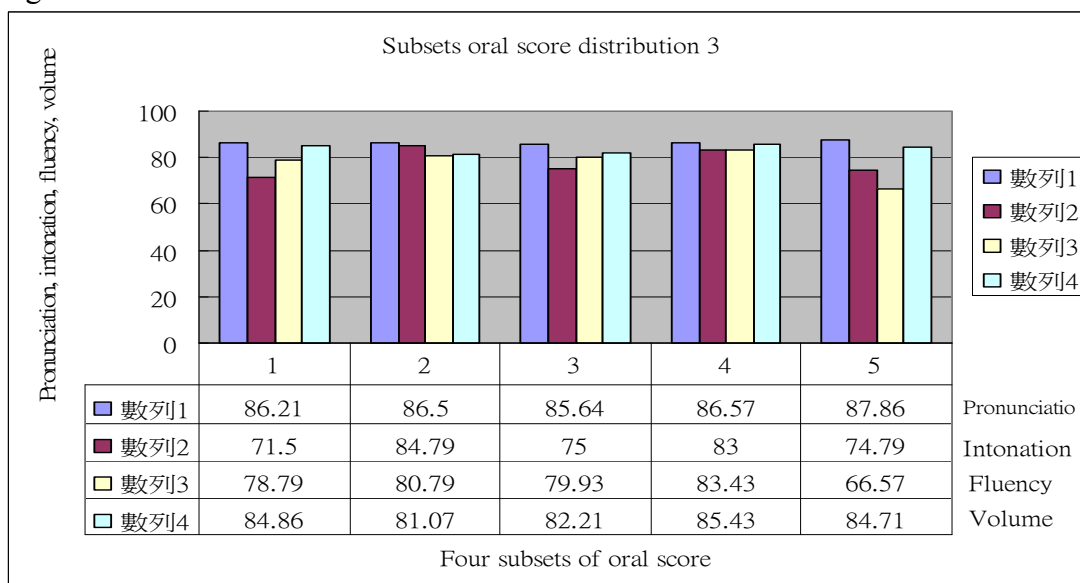
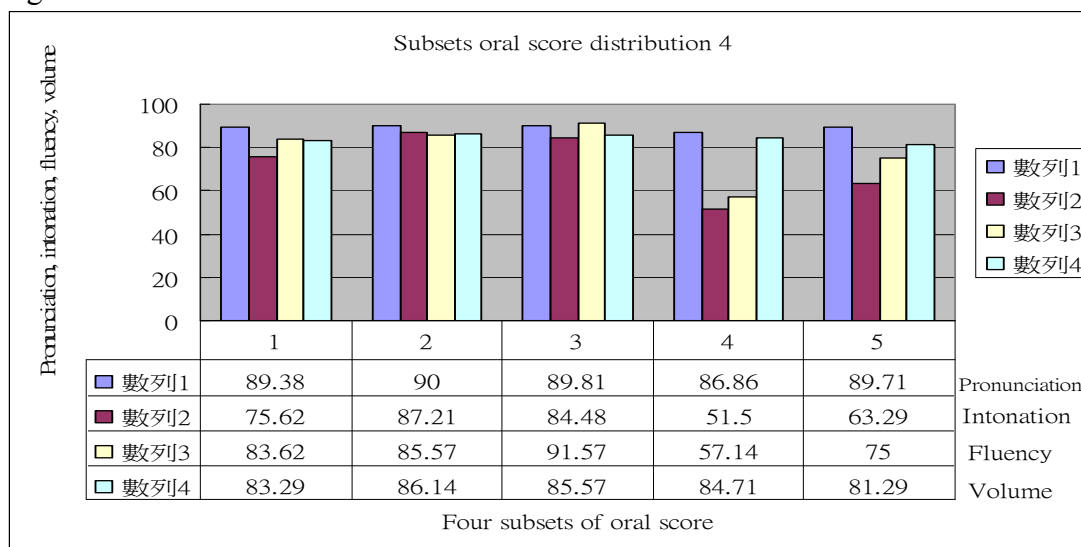


Figure 1.4 Subsets oral score distribution 4



DISCUSSION

The collected data on individual scores of listening and oral practices shows that participants who received higher score on TOEIC listening section also got higher scores on MyET oral practices. On the contrary, participants who got lower score on TOEIC listening received relatively lower score on MyET oral practices. Therefore, high oral performance can be expected from high listening proficiency.

According to participants' oral practice reports, the most difficult parts in speaking are intonation and fluency. In the intonation section, the passing rate is only 35 percent. The low passing rate also shows in fluency, which is only 30 percent. The oral report demonstrates that most English learners have no problem on pronunciation and volume, however, the failing scores (below 80) on intonation and fluency sections show that intonation and fluency might be difficult for ESL speakers.

Another discussion here will be the advantages of using MyET as the criteria for rating oral proficiency. Owing to the subjective standards for rating oral performance, the application of MyET, which evaluates students' oral performance by Automatic Speech Analysis System (ASAS), can provide a more objective scale of rating learners' oral proficiency.

CONCLUSION

This section contains two parts: findings and implications, and limitations and suggestions.

Findings and Implications The findings of this research indicate that if English learners aim to improve their oral proficiency, the first step is to enhance their listening ability; such sequence of learning is also corresponding to input and output hypothesis. As for ESL teachers, this study provides evidences to show that instructions on oral training should incorporate both listening and speaking instruction, and listening instruction should be before oral training. For both English learners and teachers, in order to reach better English oral proficiency, the practice of listening and the frequency of using English to communicate are suggested methods.

The subset oral practice scores suggested that there are two difficult oral features while speaking English: intonation and fluency. Echoing the controlled-automatic model of learning, fluency appears in more advanced speakers, whose language is processed automatically in mind. Hence, in order to reduce hesitation in oral performance, more language input is

suggested to cultivate an automatic language mechanism for preparing later speech. As for the intonation difficulty in oral performance, since intonation is different between Chinese and English, intonation shows less satisfying results than the other subset items in oral practice. Even though intonation and fluency might not impede understanding in daily communication, they indeed affect English learners' oral performance. Intonation and fluency are two possible suprasegmental features that influence speakers' oral proficiency as well as English learners' ability to comprehend native speakers (Miller, 2000). Hence, the importance of intonation and fluency in oral proficiency should not be underestimated if English learners want to sound more like a native speaker. Lastly, as the research results indicate, listen more carefully and be familiarizing with native speakers' intonation will not only help improve EFL learners' intonation, but also avoid misunderstanding while communicating with native English speakers.

Limitations and Suggestions

The limitations of this research are, first, even though the total score of TOEIC listening test remain the same, the difficulty of TOEIC listening test had been slightly modified since March, 2008; second, the total number of participants should be increased to make the research result more validate and reliable; lastly, a follow-up study should be added in the future to see how learners apply MyET to identify their oral suprasegmental and segmental difficulties and to improve their English speaking and listening abilities.

REFERENCES

- Berk, R. A. (1986). A consumer's guide to setting performance standards on criterion-referenced tests. *Review of Education Research*, 56(1), 137-172.
- Bickerton, D. (2000). *Behavioral and brain sciences*. Cambridge University Press.
- Brown, G., & Yule, G. (1983). *Teaching the spoken language*. Cambridge: Cambridge University Press.
- Brown, H. D. (2001). *Teaching by principles: an interactive approach to language pedagogy*. New York: Longman.
- Canale, M., & Swale, M. (1980). Theoretical bases of communicative approached to second language teaching and testing. *Applied linguistics*, 1, 1-47.
- Clark, E. V., & Hecht, B. F. (1983). Comprehension, production and language acquisition. *Annual Review of Psychology*, 34, 325-349.
- Clark, H. H., & Chase, W. G. (1974). Perceptual coding strategies in the formation and verification of descriptions. *Memory & Cognition*, 2, 101-111.
- Faggen, J. (1994). Research Memorandum. *Setting standards for constructed-response tests: An overview*. Princeton, NJ:ETS.
- Fodor, J. A., Bever, T. G. & Garrett, M. (1974). *The psychology of language*. New York: McGraw-Hill.
- Garman, M. (2002). *Psycholinguistics*. Cambridge: Cambridge University Press.
- Goh, C. (1998). How ESL learners with different listening abilities use comprehension strategies and tactics. *Language Teaching Research*, 2, 124-147.
- Hambleton, R. K., Jaeger, R. M., Plake, B. S., & Mills, C. (2000). Setting performance standards on complex educational assessments. *Applied Psychological Measurement*, 24(4), 355-366.
- Hudson, R. (2007). *Language Networks*. New York: Oxford University Press.
- Jarvella, R. J. & Collas, J. G. (1974). Memory for the intentions of sentences. *Memory & Cognition*, 2, 185-188.
- Krashen, S. (1985). *The Input Hypothesis: Issues and Implications*. Longman Group United Kingdom.

Miller, F. S., (2000). *Targeting Pronunciation: The Intonation, Sounds, and Rhythm of American English*. Boston, MA: Houghton Mifflin Company.

Corpus-Based Remediation for Technology-Students' Vocabulary and Reading Comprehension

Chin-cheng Huang (黃金誠)
Fooyin University
cchuang@mail.npust.edu.tw

This study aims to explore a more appropriate teaching-learning method to better help technology-students more efficiently and effectively develop their vocabulary knowledge and reading comprehension. The subjects were 139 sophomores from a management college at a technological university in southern Taiwan. This study divided the subjects into three groups: (1) control group (CG) (n = 52), (2) vocabulary training (VT) group (n = 48) and (3) extensive-reading (ER) group (n = 39). Each subject in VT group received vocabulary training based on the Academic Word List for Commercial and Management Students (AWLCMS), and they had to receive a vocabulary quiz every next week. Each subject in the ER group had to read 10 marketing-related articles selected in one semester. All participants had to take Schmitt's (2000) Vocabulary Levels Test (2000-, 3000-, and academic word) and read a marketing-related article and do a recall protocol in Chinese after reading as pretest and posttest. The results showed that although three treatments could improve students' vocabulary knowledge, vocabulary training and ER training could improve the subjects' reading comprehension. However, the subjects receiving ER training performed significantly better in vocabulary knowledge and reading comprehension than those of other two groups.

INTRODUCTION

In Taiwan, though EFL students at technological university are now recognized as lower English-proficiency learners than those at comprehensive university (Lin, 1995; Huang, 2004a, 2004b), in order to read English-printed textbooks and materials in the special areas of their studies and English courses, nevertheless, their English skills have to reach a certain level of reading comprehension during/post their reading. If they fail to attain an appropriate reading comprehension, they then will also fail to understand the information contained in the text. If college students cannot obtain sufficient and accurate information and/or facts drawn from their set course-textbooks they will encounter serious problems in their effort to learn special knowledge in their special field. Instructors at technological university often complain about this issue when they have little option but to employ English-printed textbooks and/or materials to teach technology-students, this is especially the case for students learning to specialize in management studies and this issue is further highlighted by their need to translate theory into practice, sometimes involving hazardous procedures. Instructors fully understand using textbooks and materials written in English are necessary if not vital to inform and develop their students' knowledge of their study areas, and are equally aware of students' capabilities or limitations in English, yet they have to resort to the use of Chinese-translated textbooks or instructor-translated handouts. They are eager to find a teaching strategy which offers them a more appropriate information-exchange approach to help their students to gain more information or ideas from English-printed textbooks.

Corpus-based Teaching

Since, 1980s, with the benefits of advance technology, corpora are developing fast, broadly, and largely (Rundell, 2008), a lot of studies (Gavioli and Aston, 2001; Moudraia, 2003, 2004) apply corpora in language instruction. However, some researchers (Cook, 1998; Owen, 1996; Prodromou, 1997; Seidlhofer, 1999; Widdowson, 2000) indicated the problems

of using corpus-based materials in language pedagogy, O'keeffe and Farr (2003) argued "that in the end it is teachers who will engage in the process of recontextualising corpora and any useful findings from corpus-based description. It is teachers who will mediate between corpus-based content and the needs of the learners in their individual classroom contexts" (p. 391). Salsbury and Crummer (2008) and Rundell (2008) advocates that the chief advantages of corpus-based teaching can offer EFL learners authentic texts and concentrate on the most usual, frequent, typical, and useful materials.

Recently, some researchers conducted corpus-based studies to help students learn authentic language. Mudraya (2006) compiled a nearly 2,000,000 word corpus, called the Student Engineering English Corpus (SEEC), to conduct data-driven instructional activities. The results showed that this lexical approach can increase engineering-major students' language experience and improved their English awareness. Liu (2003) analyzed 3 contemporary spoken corpora to gain 302 frequently used idioms. He then classified those idioms in 3 bands based on teaching college students. Simpson and Mendis (2003) analyzed the Michigan Corpus of Academic Spoken English and found 238 idioms types. They recommended that corpus-based idioms can help students not only learn unknown idioms, but also raise their awareness of the speech contexts. Frazier (2003) analyzed 3 corpora to make sure that in counterfactual or hypothetical environments, would-clauses occur more often quite distant from or entirely without any corresponding if-clauses. He suggested that teachers should use this finding in the instruction of language structure. Bernardini (2003) used the corpus of English X Italian (CEXI) to explain how to use corpus-based findings to instruct socio-cultural insights, discourse-structuring expressions, and lexical patterns. Yoon (2005) used a corpus to teach six students to learn academic writing and found that corpus-base pedagogy could help students not only grammatical and vocabulary problems, but also promote their awareness of language and perception of grammar. Since most of these studies focused on the analysis of corpora, the present study utilizes the frequent words of a 2,000,000-word corpus to train students to memorize vocabulary and develop reading comprehension.

Extensive Reading

Recently, some studies (Day, 2002; Dupuy, 1997; Mason & Krashen, 1997; Renandya & Jacobs, 2002; U, 2001) in a consensus of their findings revealed extensive reading was a more effective method to facilitate learners' reading comprehension than other pedagogical methods because it can assist readers to develop reading habit and accelerate their level of autonomy in learning. In order to investigate a more effective method to instruct EFL readers, Rodriguez and Sadoski (2000) randomly assigned 160 9th graders to 4 learning situations: (1) rote rehearsal, (2) context, (3) context and keyword, and (4) keyword. Their outcome presented the students in context and keyword situation recalled most cues than did students in the other 3 situations. Mason and Krashen (1997) asked their university students in Japan to self-select reading materials and write brief summaries and/or comments on reading texts to investigate the influence of extensive reading on EFL students' reading proficiency. Their results revealed extensive reading was a more effective approach to facilitate reading comprehension and reading proficiency than a traditional approach.

Horst, Cobb, and Meara (1998) conducted a study to investigate the influence of exposure to extensive reading and frequent words on Jordanian university students' recognizance of the meanings of unknown words and the association between words. Their findings showed using frequency words to teach students vocabulary could not increase their students' vocabulary knowledge, but extensive reading could help students gain more words through incidental learning. Like Horst, Cobb, and Meara's findings, Cobb (1999) found that Omani university students who learned English for academic purposes could not increase their vocabulary knowledge through studying word list and dictionary, but rather, improved

through extensive reading. In addition, Lituanas, Jacobs, and Renandya (2001) found that extensive reading led remedial students in a Philippines secondary school to achieve greater gains in reading comprehension than via a traditional reading situation. Leung (2002) also found that extensive reading could help Japanese adults enhance not only vocabulary knowledge and reading comprehension, but also foster a positive attitude toward reading.

Research Questions

In order to explore a more effective method to help English-limited freshmen from agricultural college facilitate their vocabulary knowledge and reading comprehension, the researcher set two research questions for this study: Do no treatment, vocabulary training, and extensive-reading training have the same efficiency on improving vocabulary knowledge and reading comprehension of technology students of management major?

METHOD

Participants

In this study, all subjects came from three classes. The researchers assigned one class as the control group (n = 52), another class as the vocabulary training (VT) group (n = 48), and the other class as the extensive-reading (ER) group (n = 39). According to the curriculum regulation of the Department of Running Management, the subjects in three groups had to take marketing management as their required course. All participants who took this course used the same English textbook, had the same time periods (three hours per week), share the same teaching schedule, and had the same midterm and final test. The textbook used in this course was 樓永堅 and 方世榮 translated Kotler Keller's (2005) *Marketing Management*. In class, since students' textbook was printed in Chinese, the teacher explained the content knowledge and communicated with his students in Chinese. In their midterm and final examination, they answered questions in Chinese, too.

Instruments

The vocabulary levels test (VLT) in the present study was derived from the Schmitt's (2000) *Vocabulary Levels Test: Version 1* (pp. 192-200). The Schmitt's Vocabulary Levels Test (SVLT) is a paper-and-pencil test that includes the following five levels of word-frequency—the 2,000-word level, the 3,000-word level, the 5,000-word level, the 10,000-word level, and the Academic Vocabulary level (including general academic words used in university texts). Since technology-students in Taiwan only have a limited vocabulary size (Huang, 2004b, 2004c), the vocabulary level test in this study only included the first two levels and Academic vocabulary level of the Schmitt's Vocabulary Level Test. Based on Schmitt, Schmitt, and Clapham (2001), "If a shorter test is desirable, it may be reasonable to exclude certain sections (such as the 10,000 section for beginning learners,...) rather than shortening any particular section" (p. 71). Therefore, the present study excluded Schmitt's 10,000-word level. Meanwhile, in Schmitt, Schmitt, and Clapham's study, Schmitt's VLT showed excellent reliability and validity, in which, test takers who knew more low-frequency words also knew more high-frequency words. The reliability (Cronbach Alpha) of Schmitt's VLT: Version 1 was .920 for 2000-word level, .929 for 3,000-word level, .927 for 5,000-word level, and .958 for Academic vocabulary level. The reliability of vocabulary level test for the present study was .84 for the pretest and .90 for the posttest.

Reading Text

The researcher quoted English reading passage from the book, *Marketing* (5 ed.), published by Lamb, Hair, and McDaniel, C. (2000). Its content involves marketing management philosophies (Appendix B). The readability of the text was at the 17th level according to Fry's (1977) Readability Graph. The total running words in this article is 302. In this article, although there were only 13.91% running words out of West's (1953) General Service List, most of the words excluded in West's List frequently appeared in the reading

passage, such as orientation, aggressive, philosophy, battery, and mercury. Since the content knowledge of the English text was familiar to every participant, they may not feel difficult to read this passage since almost every university student of a commercial or a management college had the knowledge that the marketing activities of an organization are influenced production, sales, market, and societal marketing orientation.

Recall Protocol

The recall protocol (RP) consisted of a free written recall of a passage, after it had been read. After the subjects had read the English passage, they were given ten minutes to write down in Chinese as much as possible all they could remember from the text, without direct reference to the reading passage. In their writing, it was unnecessary for them to follow the sequence of idea units of content knowledge. The reliability (Cronbach Alpha) of the recall protocol score for the pretest was .24, and .37 for the posttest. This reliability was moderate.

Treatment

Because the students in three groups (classes) took the same course, Marketing Management, used the same textbook, and received the same English curriculum and teaching schedule, the experimental training depended much on students' study beyond designated course hours, meaning students could read outside the hours set for the course. In class, the instructors who taught in Chinese employed text-summary to build up students' content knowledge in their major field. They did not instruct their students' English, nor helped the researchers to review vocabulary or reading materials. The instructor of the vocabulary training group helped the researchers to delivery a word list for memorizing and monitored a 15-word quiz in the next week, while the instructor of the extensive reading group helped deliver an article for outside reading only and collect students' comments in the next week. The instructor of the control group did not care about this experiment, neither asked his students to memorize vocabulary, nor ask them to read management-related passages and to write English comments.

In terms of training materials, in order to probe the impact of corpus-based vocabulary training on technology-university students, the researcher spend one year collecting 102 academic textbooks of commercial and management areas, scanned their texts, and construct a 2,535,226-word corpus, named the Academic Word List for Commercial and Management Students (AWLCMS) (Huang, 2007). This corpus consists of 37,317 types and 23,339 word families. The researchers applied this corpus as the base of remedial materials. First, the researcher deleted West's (1953) General Service List (GSL) out from the word families because the target words for English remediation were the high frequent academic words, not the high-frequency words. Second, based on the principle that the higher frequency a word showed, the earlier it should be memorized, the researcher selected 56 high-frequency academic words and 10 technical words of marketing as the target words every week to compose a vocabulary profile. Then, the researchers used West's high-frequency words to make sentences or adapt sentences from dictionaries for interpreting the meaning of the target words. Fourth, the instructor of the VT group delivered the every-week word profile to students and asked them to read all sentences and memorize the target words at home. Meanwhile, he monitored a 15-question quiz which was based on the previous target words. Excepting the weeks for preparation, the pretest, the midterm test, holiday, the last two weeks for posttest and the final test, every student in the VT group received 10 word profiles, 10 quizzes, and totally they had to memorize about 660 target words.

In order to investigate the influence of the intensive reading on vocabulary knowledge, the researcher decided to arrange extensive reading as the other group's task. Taking note of Mason and Krashen (1997), the researcher composed a reading profile which included 10 articles. All of the passages were selected from three marketing-related books. The length of each passage and essay was between 500 and 1,000 words and the topic of each passage and

essay involved marketing field. After completing the pretest, the subjects in the ER group were assigned a passage as outside-reading task per week. They had to read the assigned passage, look up the dictionary to find the meaning of unknown words, underline the main ideas, and write down their comments or critiques based on the texts they read. However, copying from the reading text was forbidden. In the next meeting in-class, they had to turn in their writing. After correction, the researcher would return their comments with some feedback. Excluding the first three weeks for the pretest, holiday, the week before the midterm test, the last two weeks for the posttest and the final test, every student in the ER group received 10 articles totally.

Data Collection Procedures

Vocabulary test and reading comprehension test were conducted in the third week as the pretest and the sixteenth week of the fall semester as the posttest. All subjects had to complete the tests in the same time period without the use of dictionaries and cheating. They took about 43 minutes to complete the pretest and the posttest: (1) 25 minutes to complete the vocabulary test, (2) 8 minutes to read the English text and memorize the messages, and (3) 10 minutes to do the RP.

Data Analysis

In the present study, reading comprehension score involved the total score of the three parts on the recall protocol—including the scores for main ideas, supporting ideas, and details. The scores of reading comprehension were accumulated under the formula: the RP score = $5 \times$ the total number of main ideas + $3 \times$ the total number of supporting ideas + $1 \times$ the total number of details. Since the content of the reading passage included 5 main ideas, 4 supporting ideas, and 6 details, the total possible score of the reading comprehension was 43. The vocabulary scores were judged with reference to three levels, the 2,000-word, 3,000-word, and academic word level. Because one correct answer stands for 1 point in each level test, the highest possible score was 90 (1 point \times 30 items \times 3 levels).

The researcher used the Statistical Package for Social Science (SPSS) 15 version to analyze the collected data to examine the correlation between the dependent and the independent variables. This study applied Levene's test to check the equality of variances among three groups and then dependent *t* test to evaluate if the progress among groups significant or not after the experimental training. All hypotheses were tested at the $\alpha = .05$ level of significance using a two-tailed test.

RESULTS AND DISCUSSION

Placement Test

Since the major purpose of this study was to investigate the progress of vocabulary knowledge and reading comprehension, the researchers applied the independent *t* test to analyze the data of the pre-test and check if the subjects in each group had equivalent vocabulary knowledge and reading comprehension before instruction. Table 1 displays the results of comparing the means of the vocabulary knowledge score and reading comprehension score of each group in the pre-test. Although the subjects in the VT group gained a little higher means of the vocabulary knowledge scores than did those in the ER group and the CG group, the level of significances of Revene's test for equality of variances among three groups were not statistically significant $F(98) = .64, p > .05, F(89) = .04, p > .05,$ and $F(85) = 1.03, p > .05.$ In other hand, the subjects in the ER group recalled more information from the reading text than did those in the VT group and the CG group and the mean differences among three groups were very slight, between .07 and .31. The level of significances of Revene's test for equality of variances among three groups were statistically insignificant, $F(98) = .05, p > .05, F(89) = .10, p > .05, F(85) = .01, p > .05.$ The above findings indicate any pre-test differences in the subjects' vocabulary knowledge and reading

comprehension among each group were statistically insignificant.

Table 1. *Revene's Test for Equality of Variances by Comparing Control, Vocabulary and Reading Group on Pre-test of Vocabulary Knowledge and Reading Comprehension (N = 139)*

Variables	Groups	n	Mean	SD	df	F	Mean difference
Vocabulary	CG	52	27.60	10.37	98	.64	-.46
	VT	48	28.06	11.69			
	CG	52	27.60	10.37	89	.04	-.37
	ER	39	27.97	11.35			
	VT	48	28.06	11.69	85	1.03	.09
	ER	39	27.97	11.35			
Reading	CG	52	1.10	1.89	98	.05	-.07
	VT	48	1.17	1.73			
	CG	52	1.10	1.89	89	.10	-.31
	ER	39	1.41	1.82			
	VT	48	1.17	1.73	85	.01	-.24
	ER	39	1.41	1.82			

Note: * $p < .05$.

Findings of the Post-test

The findings in comparing the CG, VT and ER group on post-test of the vocabulary knowledge score and the reading comprehension score is presented in Table 2. The means of the vocabulary knowledge scores of the CG, VT, and ER group were 29.10 with a standard deviation of 10.57, 35.19 with a standard deviation of 12.66, and 42.00 with a standard deviation of 13.85. The mean difference between the CG group and the VT group was 6.09, between the CG group and the ER group was 12.9, and between the VT group and the ER group was 6.81. In the pretest, the VT group's mean of vocabulary score was .46 higher than that of the CG group and .09 higher than that of the ER group (Table 1); however, the ER group gained 12.90 higher than that of the CG group and 6.81 higher than the VT group in the posttest (Table 2). Like the results of vocabulary knowledge, the ER group gained 4.37 higher than that of the CG group and 3.35 higher than that of the VT group in the posttest.

Table 2. *Minimum, Maximum, Means, and Standard Deviations of Vocabulary Knowledge Scores and Reading Comprehension Scores of Each Group in the Posttest (N = 139)*

Variables	Groups	n	Minimum	Maximum	Mean	SD
Vocabulary	CG	52	14	67	29.10	10.57
	VT	48	16	75	35.19	12.66
	ER	39	22	67	42.00	13.85
Reading	CG	52	0	10	1.40	2.18
	VT	48	0	9	2.42	2.23
	ER	39	0	13	5.77	3.35

Although the students in the ER group gained higher score in vocabulary knowledge and reading comprehension in the posttest than those in the VT group, while the students in the VT group performed better in vocabulary score and reading score in the posttest than those in the CG, it is necessary to apply ANOVA to make sure whether the mean differences of vocabulary scores and reading scores among three groups are significant or insignificant. Table 3 presented the outcomes of applying one-way ANOVA to test the variances of the vocabulary score and reading score in the posttest by three groups. The results show that the

mean difference of vocabulary scores among three groups was statistically significant, $F(2, 136) = 12.33, p < .01$, while the mean difference of reading scores by three groups was also significant, $F(2, 136) = 33.75, p < .01$. After finding both F values of vocabulary score and reading score were significant, the researcher utilized Post Hoc to explore the mean difference between which groups was significant. Table 3 shows that the mean differences of vocabulary scores between the ER group and the VT group, the ER group and control group, and the VT group and the control group were all statistically significant. Meanwhile, the mean differences of reading scores between the ER group and the VT group, and the ER group and the control group were statistically significant. However, the mean difference of reading score between the VT group and the control group was not significant. Since the subjects in the ER group had a significantly higher reading comprehension score than did those in the VT group and the control group, the findings of Table 3 mean that in the post-test, the subjects in the ER group could recall more information from the reading passage than did those in the VT group and the control group.

Table 3. *Applying ANOVA to Analyze the Variances of Vocabulary Knowledge and Reading Comprehension in the Post-test by three groups (N = 139)*

Items	Variances	Sum squares	df	Mean square	F	Post Hoc
Vocabulary	Between groups	3720.67	2	1860.34	12.33**	3>2>1
	Within groups	20515.83	136	150.85		
	Total	24236.50	138			
Reading	Between groups	447.83	2	223.91	33.72**	3>2; 3>1
	Within groups	903.11	136	6.64		
	Total	1350.94	138			

Note. * $p < .05$, ** $p < .01$. 1 refers to the CG group, 2 refers to the VT group, 3 refers to the ER group.

Table 4 presents the findings of paired sample t test for comparing the vocabulary knowledge score and the reading comprehension score of the pre-test and post-test by each group. In term of the vocabulary score, the mean gained scores of the subjects in the control group, the VT group, and the ER group were 1.50, 7.13, and 14.03. Every group presented their progression. Gains in the vocabulary knowledge scores in the control group, the VT and the ER group were statistically significant, $t(51) = 2.40, p < .05$, $t(47) = 10.20, p < .01$, and $t(38) = 13.08, p < .01$. It means students in three groups significantly gained more vocabulary knowledge score in the post-test.

In term of the reading comprehension score, the mean gained score of the subjects in the control group, the VT group, and ER group were .30, 1.25, and 4.36. The differences of means between the pre-test and post-test in the VT group and the ER group were statistically significant, $t(47) = 2.96, p < .01$ and $t(37) = 8.06, p < .01$. However, the difference of means in the control group was insignificant, $t(51) = .91, p > .05$. It implied that both the VT group and the ER group significantly obtained more reading comprehension score in the post-test, but not the control group.

Table 4. *Paired Sample t Test for Comparing the Vocabulary Knowledge Score and the Reading Comprehension Score of the Pre-test and Post-test by Each Group (N = 139)*

Groups	Tests	<i>n</i>	Mean	<i>SD</i>	<i>t</i>	Gain
CG	Pre-vocabulary	52	27.60	10.37	-2.40*	-1.50
	Post-vocabulary	52	29.10	10.57		
	Pre-reading	52	1.10	1.89	-.91	-.30
	Post-reading	52	1.40	2.18		
VT	Pre-vocabulary	48	28.06	11.69	-10.20**	-7.13
	Post-vocabulary	48	35.19	12.66		
	Pre-reading	48	1.17	1.73	-2.96**	-1.25
	Post-reading	48	2.42	2.23		
ER	Pre-vocabulary	39	27.97	11.35	-13.08**	-14.03
	Post-vocabulary	39	42.00	13.85		
	Pre-reading	39	1.41	1.82	-8.06**	-4.36
	Post-reading	39	5.77	3.35		

Note. * $p < .05$, ** $p < .01$. 1 refers to the CG group, 2 refers to the VT group, 3 refers to the ER group.

Discussion of Findings

Since the subjects in the control group, the VT, and ER group had close to equal/homogeneity in the placement test (Table 1), the significantly different progress of vocabulary knowledge among three groups implied that their different progress of vocabulary knowledge was due to the different training, and not attributed to chance or accident. According to table 4, even the students got no treatment can slightly improve their vocabulary knowledge. The reason is that university students may incidentally acquire English words through contacting English materials or doing their assignments of other subjects. The students in the VT group could significantly increase their vocabulary knowledge. That is, rote rehearsal of high-frequency academic words could facilitate technology-sophomores' vocabulary knowledge. The students in the ER group could meaningfully promote their vocabulary knowledge, too. In terms of developing vocabulary knowledge of technology-sophomores of the management college, extensive-reading method was proved as better than general academic words and technical words method.

The findings of the present study corresponded with Mason and Krashen's (1997), Lituanas and his colleagues' (2001), and Rodriguez and Sadoski's (2000) study. These previous studies revealed extensive readers gained more vocabulary knowledge score than readers via traditional approach, cloze test, keyword training, or rote. However, since their studies did not have placement test, they could not make sure how much their subjects' progress was impacted by receiving other training or course.

In this study, although gains made by the subjects receiving frequent words, academic words and technical words were inferior to those made by extensive students, the rote of high-frequent words academic words and technical words did help technology-sophomores grow their vocabulary knowledge. This result does not correspond with Horst, Cobb, and Meara's (1998), Cobb's (1999), and Watt's (1995) findings that isolated-learning unfamiliar words could not develop learners' vocabulary knowledge. The reason why rote rehearsal could improve technology-students' vocabulary knowledge is that most general academic words and technical words are easy to remember since those words involve their study field and they often confront those words in other courses. Meanwhile, they could incidentally learn most of high-frequent words through context sentences when they studied those frequent academic words and technical words. In addition, since the instructor held a quiz every week, students had to understand and memorize every target word outside class, they

could gain higher scores. In this teaching approach, students do gradually remember some words for use in their academic study.

On a different concern, extensive readers have to look up a dictionary to find the meanings of unknown words and they then should better understand the meaning of a whole paragraph and/or passage. When they read a passage, a lot of words repeatedly appear, including high-frequency words, academic words, technical words, even low-frequency words. Therefore, they could remember a lot of words co-incidentally. Some students in the ER group even became autonomic readers who visited the researcher, their instructor or other teachers to discuss their comments and the researchers' feedbacks. Currently, despite the fact that they are no longer compelled to pursue an English course, they continue to request from the researcher and instructor for further English passages to read.

Based on Table 3 and 4, it is clear that the performance of students of the ER group was significantly better than that of students using the word-training method or no treatment after experimental training for a semester. We saw the subjects in the ER group promoted their reading comprehension score from 1.41 to 5.77 which was more than a three times improvement. This impressive progress, the researcher would like to stress, might be attributed to the following reasons. First, a lot of reading in turn fostered extensive readers' reading motivation and attitude. In the beginning, most students usually complained they had more extra reading than their counterparts who took the same course now or before. After the experimental instruction, several students appeared in the instructors' office to require for more reading passages or essays. The researchers could well appreciate the critical change in students' reading motivation and pro-learning attitude. This phenomenon is in accordant with Leung's (2002) and Day and Bamford's (1998) findings that extensive reading can positively enhance a reader's attitude toward reading. In Nuttall's (1982) virtuous circle, she has clearly explained this phenomenon, namely; one reads more and understands better, he/she then enjoys reading and anticipates reading more.

Second, extensive reading offers EFL readers more opportunities to exposure in the English texts. In reading, extensive readers repeatedly recognized a variety of high-frequency words, general academic words, and some technical words. Although they were not asked to memorize any word list of high-frequency academic words and technical words, they could learn those words from the context of the reading texts incidentally. This learning-situation is, according to Rodriguez and Sadoski's (2000), Horst, Cobb, and Meara's (1998), and Cobb's (1999) results, consistent with the claim that, extensive reading can help readers acquire more words through incident-learning.

Finally, the students in the ER group had to write a comment or a criticism in English according to the content of the reading text, but avoid copying the sentences as printed in the text. In order to criticize the content of the reading passage, they had to read carefully, look dictionaries up, analyze sentence structures, understand the gist, take notes, and imitate author's writing skills. The most important point of this activity was that every extensive reader should focus on a learned topic and try to use words already familiar or known to express their ideas and viewpoints. When confronted with unfamiliar words or sentence structures, they would look up dictionaries or ask others. In addition, they also learned something from the researchers' feedback. The connection of reading and writing made extensive readers learn more reading skills and promote their reading comprehension. This outcome corresponds with Mason and Krashen's (1997) study, in that those who write summaries after reading achieve more gain in reading comprehension than did those who did not write. Extensive reading, for this research, facilitates readers' reading proficiency of academic texts, and paraphrasing writing provides the readers more opportunities to re-think and put to practice, use what they have learned.

ACKNOWLEDGEMENTS

This paper is drawn from the final report for the project supported by National Science Council (NSC96-2411-H-127-001).

REFERENCES

- 樓永堅, 方世榮譯 (2006)。行銷管理學。台灣：東華書局。
- Bernardini, S. (2003). Designing a corpus for translation and language teaching: The CEXI experience. *TESOL quarterly*, 37(3), 528-537.
- Cobb, T. (1999). Breath and depth of lexical acquisition with hands-on concordancing. *Computer Assisted Language Learning*, 12(4), 345-360.
- Cook, G. (1998). The uses of reality: A reply to Ronald Carter. *ELT Journal*, 52(1), 57-63.
- Day, R. R. (2002). Top ten principles for teaching extensive reading: A response. *Reading in a Foreign Language*, 14(2).
- Day, R. R., & Bamford, J. (1998). *Extensive reading in the second language classroom*. New York: Cambridge University Press.
- Dupuy, B. C. (1997). Voices from the classroom: Intermediate-level French students favor extensive reading over grammar and give their reasons. *Applied Language Learning*, 8(2), 285-293.
- Frazier, S. (2003). A corpus analysis of would-clauses without adjacent if-clauses. *TESOL Quarterly*, 37(3), 443-466.
- Fry, E. B. (1977). Fry's readability graph: Clarifications, validity, and extension to level 17. *Journal of reading*, 21, 242-252.
- Gavioli, L., & Aston, G. (2001). Enriching reality: Language corpora in language pedagogy. *ELT Journal*, 55(3), 238-246.
- Horst, M., Cobb, T., & Meara, P. (1998). Beyond a clockwork orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language*, 11(2), 207-223.
- Huang, C. C. (2004a). *High school and university students' vocabulary knowledge, content knowledge and differential reading comprehension*. Taipei, Taiwan: Crane.
- Huang, C. C. (2004b). A vocabulary gap between agricultural freshmen and textbooks. *Proceedings of the 21st International Conference on English Teaching and Learning in the R.O.C.* (pp. 205-217). Taipei, Taiwan: Crane.
- Huang, C. C. (2004c). 科技大學學生英文字彙量及大一應原文教科書字彙量研究. *Proceedings of the 21st International Conference on English Teaching and Learning in the R.O.C.* (pp. 647-659). Taipei, Taiwan: Crane.
- Huang, C. C. (2007). A corpus of textbooks in English for students of commercial and management college at technological university and using this corpus as a basis of remedial reading program. A final report for a project of National Science Center. NSC95-2411-H-020-001.
- Iwahori, Y. (2008). Developing reading fluency: A study of extensive reading in EFL. *Reading in a Foreign Language*, 20(1), 70-91.
- Keller, K. (2005). *Marketing management*. New York: Pearson Education.
- Lamb, C. W. Jr., Hair, J. F. Jr., & McDaniel, C. (2000). *Marketing* (5 ed.). Cincinnati, OH: South-Western College Publishing.
- Leung, C. Y. (2002). Extensive reading and language learning: A diary study of a beginning learner of Japanese. *Reading in a Foreign Language*, 14(1),
- Lin, M. S. (1995). 中華民國高中高職英文教學現況比較研究。第十二屆中華民國英語文教學研討會論文集，頁59-79。台北：文鶴。
- Lituanas, P. M., Jacobs, G. M., & Renandya, W. A. (2001). An investigation of extensive reading with remedial students in a Philippines secondary school. *International Journal*

- of *Educational Research*, 35, 217-225.
- Liu, D. (2003). The most frequently used spoken American English idioms: A corpus analysis and its implications. *TESOL Quarterly*, 37(4), 671-700.
- Mason, B., & Krashen, S. (1997). Extensive reading in English as a foreign language. *System*, 25(1), 91-102.
- Moudraia, O. (2003). The student engineering corpus: Analysing word frequency. In D. Archer, P. Rayson, A. Wilson, & T. McEnery (Eds.), *Proceedings of the corpus linguistics 2003 conference* (pp. 552-561) UCREL technical paper number 16. UCREL, Lancaster University.
- Moudraia, O. (2004). The student engineering English corpus. *ICAME Journal*, 28, 139-143.
- Mudraya, O. (2006). Engineering English: A lexical frequency instructional model. *English for Specific Purposes*, 25, 235-256.
- Nuttall, C. (1982). *Teaching reading skills in a foreign language*. London: Heinemann.
- O'keeffe, A., & Farr, F. (2003). Using language corpora in initial teacher education: Pedagogic issues and practical applications. *TESOL Quarterly*, 37(3), 389-418.
- Owen, C. (1996). Do concordances need to be consulted? *ELT Journal*, 50, 219-224.
- Prodromou, L. (1997). Corpora: The real thing? *English Teaching Professional*, 5, 2-6.
- Read, J. (1988). Measuring the vocabulary knowledge of second language learners. *RELC Journal*, 19, 12-25.
- Renandya, W. A., & Jacobs, G. M. (2002). Extensive reading: Why aren't we all doing it? In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 295-302). Cambridge: Cambridge University Press.
- Rodriguez, M., & Sadoski, M. (2000). Effects of Rote, Context, keyword, and Context/keyword methods on retention of vocabulary in EFL classrooms. *Language Learning*, 50(2), 385-412.
- Rundell, M. (2008). The corpus revolution revisited. *English Today*, 24(1), 23-27.
- Salsbury, T., & Crummer, C. (2008). Using teacher-developed corpora in the CBI classroom. *English Teaching Forum*, 2, 28-37.
- Schmitt, N. (2000). *Vocabulary in language teaching*. New York: Cambridge University Press.
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the vocabulary levels test. *Language Testing*, 18(1), 55-88.
- Seidlhofer, B. (1999). Double standards: teacher education in the expanding circle. *World Englishes*, 18, 233-245.
- Simpson, R., & Mendis, D. (2003). A corpus-based study of idioms in academic speech. *TESOL Quarterly*, 37(3), 419-441.
- U, A. (2001). Opening up the minds of ESL learners to read the minds of English writers. *English Teacher: An International Journal*, 4(3), 206-216.
- Watts, S. M. (1995). Vocabulary instruction and meaningful learning. *Journal of Reading*, 29, 193-107.
- West, M. (1953). *A general service list of English words*. London: Longman.
- Widdowson, H. G. (1978). *Teaching language as communication*. Oxford: Oxford University Press.
- Widdowson, H. G. (2000). On the limitations of linguistics applied. *Applied Linguistics*, 21, 3-25.
- Yoon, H. (2005). *An investigation of students' experiences with corpus technology in second language academic writing*. Unpublished doctoral dissertation, The Ohio State University, Columbus, OH.

The Effects of Storytelling Approach on English Learning Motivation of Low Achievers in Junior High School

Hui-ling Huang, Wen-hui Chen

National Yunlin University of Science and Technology

huangje@yuntech.edu.tw
g9341718@yuntech.edu.tw

The study aims to explore the effects storytelling on English learning motivation of underachievers in junior high school and to understand in what way and to what extent storytelling affects teaching and learning. Based on the methods of action research, 21 English underachievers from a Junior High School were selected to participate in the English club for 9 weeks. Data collected from classroom observations, teacher's journals, after-class questionnaires, and students' works were analyzed to assess the effects of storytelling approach. The results indicate that storytelling alone cannot improve the participants' learning motivation although it did change their attitude toward English learning. Their increased learning interest was baffled by the required learning form their regular classes and thus the motivation that should have driven them to learn more by themselves did not develop. Four factors were identified that influenced the participants' engagement in the storytelling while two major difficulties related to employing storytelling in the class were discussed.

INTRODUCTION

The application of storytelling as a scaffolding approach to improve ESL/EFL learning has been validated through a few empirical studies (Cary, 1998; Huang, 2006). However, in terms of learning motivation, the relation between storytelling and learning motivation seems to invite more hypothetical approvals than empirical examination. It is assumed that if learners love stories, they will be more motivated to learn. Yet, lack of evidence from empirical or experimental studies, at least in Taiwan, storytelling as an approach will still be questioned regarding its constraints in actual practice and the extent to which it could benefit learners of different levels. To supplement the study of storytelling on English learning, particularly on the aspect of motivation, the aim of this study is to explore the effectiveness of storytelling on motivating underachievers, who usually are considered deficits in learning motivation. Moreover, the plausibility of integrating storytelling will be examined from the teacher's perspective to uncover its merits and weaknesses when being applied in EFL classrooms.

LITERATURE REVIEW

Storytelling and EFL Learning

The merits of storytelling as an educational tool have long been recognized. First, storytelling in essence aims to connect the speaker and the listener by means of telling stories (Hendrickson, 1992), which, as Mello (2001) pointed out, may have a beneficial impact on students' interpersonal relationships and also sense of connectedness in the classroom—implying that storytelling can be an effective communication medium. For

language learning, Trousdale (1990) noted that storytelling provides a structure of scaffolding in which children's language development could be built up through the interaction with adults. Also, storytelling can activate listening and conversation between the teacher and the students so that s/he can appraise their level of comprehension and provide comprehensible input accordingly for them to improve in this interaction process (Koenig & Zorn, 2002). Storytelling can incorporate various props and performances such as dramas, music, pictures, puppets, and other forms of multi-sensory input to enhance EFL comprehension (Hendrickson, 1992; Huang, 2006).

Moreover, Trostle-Brand and Donato (2001) proposed that storytelling may impact on the development of multiple intelligence proposed by Gardner (1993). They stated that storytelling not only enhance students' linguistic intelligence, but also logical mathematical intelligence, as storytelling provides not only natural sequencing of events, cause-and-effect patterns, but also riddles and mysteries, to facilitate the ability to comprehend long chains of reasoning and so to exercise logical-mathematical skills. Visual-spatial, musical, bodily-kinesthetic intelligences can also be elevated through stories, using it to provide visual perception, refrains, chants, songs, or dance and movements for students to practice and acquire skills needed to understand a story. Also, storytelling activities can allow the listener and teller to become aware of the story characters' feelings and may therefore help to develop abilities in perceiving one's own and other's emotions, on one hand promoting students' interpersonal intelligences and on the other enhance intrapersonal intelligences.

Not surprisingly, the merits of storytelling appeal language educators and have been applied to the learning of foreign languages. In Taiwan, for instance, recent studies on storytelling and EFL learning have revealed the plausibility of applying storytelling to grade-level classrooms and the specific areas that it could contribute to the learning. In Yao's study (2003), storytelling could motivate grade students' English learning, and enhance their creativity and knowledge. Tsai (2004) indicated that storytelling was effective in enhancing the third-grade students' listening and speaking skills. Storytelling helped to improve the students' comprehension capacity of an English story (Chou, 2006). In Huang's (2006) study, story listening plus illustrated-text reading can enhance EFL young learners' reading comprehension.

Increasing studies have been conducted to provide empirical and experimental evidence to examine the actual effect of storytelling in the classrooms. However, in terms of learning motivation, although there seems an assumption that storytelling can improve learning motivation since the playfulness of storytelling involved a variety of performing techniques such as props, pictures, puppets and dramatic improvising that could arouse learners' interest in learning process (Trostle-Brand & Donato, 2001). Yet, little empirical or experimental evidence has been provided to prove the link. Thus, there is a need for the research on storytelling and its relationship to learning motivation, which to many EFL learners is a key factor of learning success.

Motivation and Underachievement

Li (2004) found that students' English learning motivation had a significant and positive correlation with their English achievement. Specifically, students who have higher expectation of learning English (higher expected value) and consider English useful in work (higher instrumental utility) are more likely to achieve better learning outcomes. Her findings are in line with the previous literature (Deci & Ryan, 1985; Gardner, 1972, 1985; Oxford & Shearin, 1994) that motivation is a strong factor of learning success. As the double-peak phenomenon of English achievement in Taiwan has become a concern of teachers and parents (Chou, 2006), students who belong to the peak of low achievement were mostly underachievers, which can be defined as having "performance which does not measure up to the individual's level of aptitude or the performance below the expected level indicated by

that individual's performance on ability and aptitude tests (Mandel & Marcus, 1988, p. 4)."

Through survey and in-depth interviews conducted in a junior high school, Liou's (2002) study identified several characteristics of underachievers as follows: (a) underachieving students generally possess negative attitudes toward English learning; (b) they were low in learning motivation, and tended to be distracted; (c) they were lacking of clear goals, and usually did not have learning plans; (d) they generally thought that school life was boring; (e) they were inclined to ask help from good friends and families rather than from teachers in terms of school work; (f) they failed to make use of time to review school work, and preferred other activities. Echoing Liou's study, Kao (2003) added that underachieving students are passive in English learning, and attend classes simply because they are required to rather than because they are motivated.

As the lack of learning motivation play a key role to the failure of underachiever's English learning, the current teaching approach obviously fail to motivate the learning. One may wonder whether the well-recognized merits of storytelling can help motivate the low achievers and extend its contribution to the affective part of language learning, which is the aim of this study.

METHODS

In order to test the plausibility of applying storytelling as an approach to motivate low achievers, this study employed the methods of action research to realize and examine the idea of "teachers as storytellers" and to reflect on the process of the storytelling implementation as well as the motivation changes of the participants during the research time. In order not to interfere the on-going curriculum of the school, a storytelling club was arranged as an extra-curriculum activity of a junior high school for this research to proceed. Thus, instead of using a regular classroom, the storytelling class was conducted in the place where the space was bigger than the regular classrooms and the arrangement of the tables and seating was suitable for the activity-based program, such as group works, discussions, and games.

Participant

Twenty-one first year students of a junior high school in middle Taiwan, whose English Grade-Point Average (GPA) fell into the lowest 20% of their classes, were sought to participate in the study. These students who were selected as the participants were confirmed by their homeroom teachers with no learning disability but were those with low English learning motivation. Eight girls and 13 boys were selected to join the English storytelling club.

The second author of this paper was the teacher researcher, who had been a volunteer storyteller at a university library from 2006 to 2007; yet she had never had the experience of adopting storytelling approach for junior high school students.

Data Collection

To increase the validity of the research, multiple data should be collected to support the explanations given by the researchers in this study (Burns, 1999; McNiff, 2002). Thus, six kinds of data were included: (a) classroom observations, (b) teaching journals, (c) informal interviews, (d) participants' homework and assignment, (e) questionnaires, and (f) web log. Besides filed notes, a digital video was utilized to record classroom activities, thus, both verbal and non-verbal behaviors and interactions would be documented for a more in-depth analysis. Teaching journal was kept to reflect on teaching performance and incidents related to the effect of storytelling. Informal interviews were conducted during the break time between classes to enquire the participants' comments and opinions about the class. Their opinions or feelings thus served as references to improve the teaching materials, activities, procedures, so that new strategies could be adapted for the following classes.

For the questionnaire, four types of the questionnaires were utilized in the research. They were (a) an EFL Learning Motivation Scale (Appendix A), (b) a classroom questionnaire (Appendix C), (c) participants' letters, and (d) a questionnaire about attitudes toward the overall project of storytelling (Appendix D). Hsu's (2003) "EFL Learning Motivation Scale" was adopted to measure the changes of the participants' learning motivation before and after the storytelling project. This scale involves three dimensions: the extrinsic motivation containing, the intrinsic motivation, and the effort in English learning process. In total, there were twenty-three 4-point Likert-scaled question items.

A web log (<http://blog.xuite.net/lemonpip/story>) was created for this study to serve as an exchange platform for teacher-and-student, or student-and-student to share their feelings and opinions about the class. The web log features no limitation of time, space, and distance so that it can be treated as an effective medium for the teacher to post class announcements of the class and also some materials such as the context of the story, or reflections. Also, the participants could post any messages and ideas on the web log, and receive an immediate response from the teacher.

The participants were also required to write letters in midterm-storytelling class to tell the teachers their reflections of the class. By the end of the research, the participants were asked to fill out the "EFL Learning Motivation Scale" again to see the overall motivation change after the storytelling sessions.

Storytelling Program and Materials

The storytelling program was held on Thursday for nine weeks from March 27, 2007 to June 21, 2007. Each session lasted about ninety minutes with a ten-minute break after the first forty minutes. In each session, a class pattern including three phases was employed as follows.

5. *Introduction.* In this phase, the story was introduced through keywords and experience-related topics to activate the participants' prior knowledge and to build up the anticipation of the story.
6. *Telling time.* Time telling was used to comprehend and appreciate the story. Thus, the story was rendered in a slow pace with rich intonation and nonverbal communication such as miming, props, drawings, or flash cards. Refrains were massively used for the participants to follow the story.
7. *Extending activities.* The extending activities were applied to recall the story and strengthen the language learning, but they also made storytelling more interesting (Collins & Cooper, 1997). In this study, several extending activities were designed according to the story told in the class. The extending activities included (a) role-plays in which the participants were encouraged to act out the story, (b) the design of a storybook that the participants in groups worked together to draw their own storybook by adapting the story, and (c) story maps, that is, to put the pictures or sentences to show the sequence of the story plot.

Nine stories chosen were:

1. One Wish (Yolen, 1986)
2. The Sausage (Yolen, 1986)
3. A Little Red House (2004)
4. *The Seven Chinese Brothers* (Mahy & Tseng, 1990)
5. *The Skeleton Hiccups* (Cuyler & Schindler, 2005).
6. *Problems, Problems, problems* (Hamilton & Welss, 2005).
7. *An Old lady who Swallowed a Fly* (Taback, 1997)
8. *Tikki Tikki Tembo* (Mosel & Lent, 1989).
9. *One Witch* (Leuck, 2006).

These stories were chosen for their complete story grammar and interesting or repetitive plots, which are universal appeal to audience of various ages.

Research Procedures

Upon the framework of action research, the research procedure was conducted through a constant cycle in order to evaluate the practice. Figure 2 illustrates the research procedure.

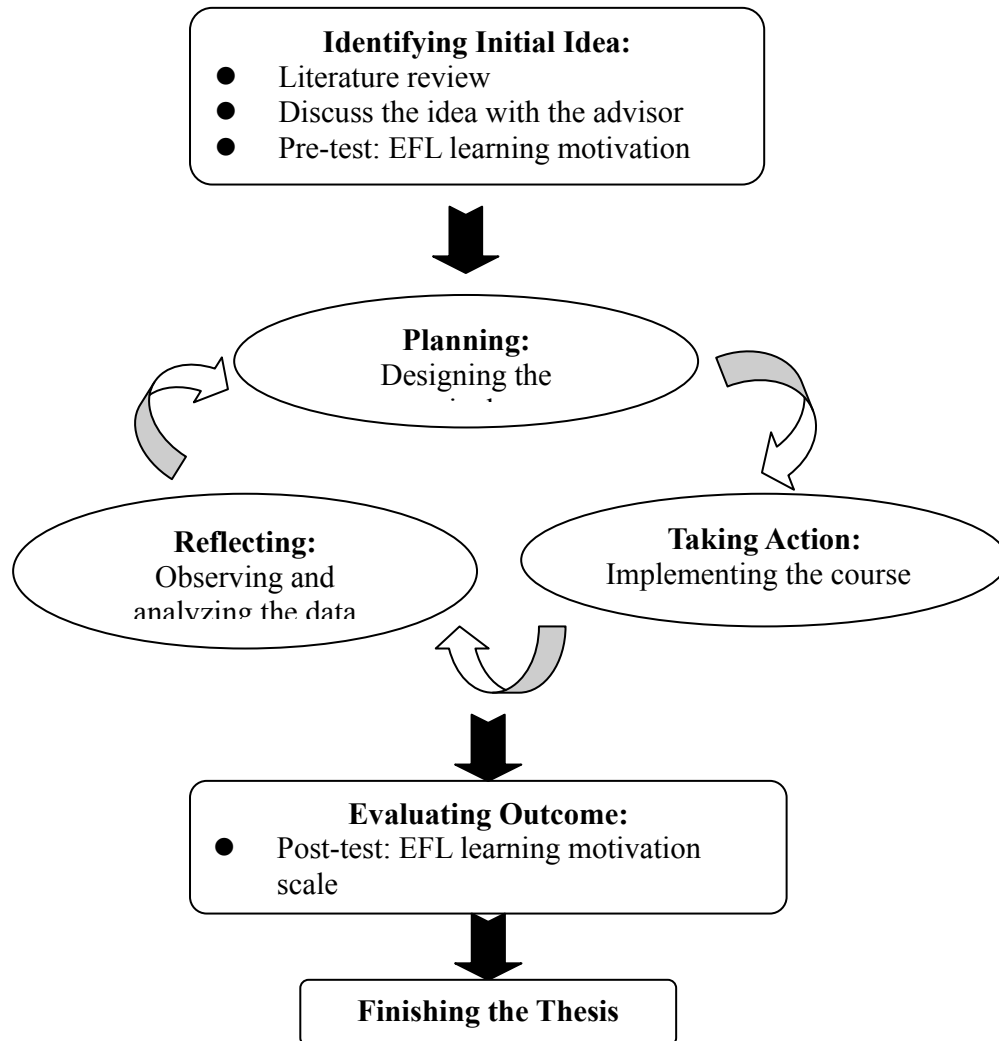


Figure 1. The procedure of the research.

Data Analysis

This research employed the framework of data analysis from McKernan (1996, as cited in Burns, 1996, p. 156-160), including assembling, coding, comparing, and reporting the outcomes.

RESULTS AND DISCUSSION

The results are divided into four sections with the aim to answer the four research questions of this study. They include firstly the effects of storytelling on participants' English learning motivation; secondly, the major factors that have influenced the participants' engagement in English storytelling; and lastly, the difficulties encountered and corresponding solutions taken. In order to protect the privacy of the participants, pseudonyms are used to

identify individuals in the related results.

The Effects of Storytelling on Participants' English Learning Motivation

The results show that there was no significant change between the pre-test and the post-test. Although the means of the post-test had increased, the difference was not statistically significant (see Table 1).

Table 1. *Means and Standard Deviations of the Pre-Test and the Post-Test*

	M	SD	T	P
The Pre-Test	53.05	11.029	-1.259	0.215
The Post-Test	57.81	13.378		

p<.05

The possible reason could be that it was hard to extend learning outside the storytelling club. Although the storytelling club had lasted for nine weeks, it was only conducted once a week. Except for the storytelling class, these participants' English learning was still based on their regular English classes. English instruction in regular class appeared to directly influence on the participants' English learning. Although the participants liked the storytelling class and they hoped that storytelling approach could be employed in their regular English class, they did not think their English teachers in the regular classes would be willing to take this kind of approach. A representative comment addressed by these participants is as follows:

Learning English through storytelling is not bad, but it hard to use in our regular English class. Our regular English class will use much time to give quizzes rather than storytelling, so I still hate English. [Jacky]

Table 2. *Frequency and Percentage of the Responses to the Attitudes toward the Whole Storytelling Session*

Statements	Frequency (%)		
	Agree	Neutral	Disagree
Do you like the way of learning English through storytelling?	15 (71.42)	6 (28.57)	0 (0.00)
Will you be willing to participate in the storytelling club in the future?	16 (76.19)	1 (0.04)	4 (19.04)
Do you expect that the storytelling approach can be conducted in regular English class?	15 (71.42)	6 (28.57)	0 (0.00)

t appeared that due to the lack of faith that this new approach would be taken by their English teachers; they did not see the changes and chances for their motivation to develop.

Table 3 *Major Factors Influencing the Participants' Engagement in Storytelling*

Major Factors	Subdivided factors
1. Story Types	<ul style="list-style-type: none"> ● Absurd and humorous story plots. ● Vivid characters. ● Story structures.
2. The Use of Props	<ul style="list-style-type: none"> ● To arouse students' attention. ● To enhance students' comprehension.
3. Extending Activities	<ul style="list-style-type: none"> ● Student-centered learning. ● Positive teacher-student interaction
4. Affective Influences	<ul style="list-style-type: none"> ● Relationship among classmates. ● Interaction between teacher and students.

The Factors that Influenced the Participants' Engagement in English Storytelling

Four factors that influenced the participant's learning have been identified. They are (a) story types, (b) the use of props (c) extending activities, and (d) affective influences. The four factors, together with their subcategories, are shown in Table 3.

Story Types

Among the 9 stories told, the participants' favorite story types were those with absurd or humorous plots as shown in Table 4. Those stories usually are noodlehead or tall tales with idiosyncratic characters whose conducts were nonsensical or hilarious. Jacky reflected, "*An Old Lady Who Swallowed a Fly* is my favorite story. I love this funny story, although it is really stupid and impossible for a person to swallow so many things like that!" Jim revealed that his favorite story was *Tikki Tikki Tembo* and commented, "How could a mother name her son such a long name? It was very exhausting to call this long name. By the way, it was funny when these two sons fell into the well."

In addition to the story plots, the participants were enthralled by characters with unique personal traits, and as a result, these characters became a key for them to recall the story plots. When being asked which story was their favorite one, Lee responded, "I don't remember the name of the story, and I don't know how to describe the story plot clearly. What I still remember is that there were an old husband and wife, and a sausage was stuck to the wife's nose." Kino also said, "My favorite story is *One Witch*. No particular reasons but many interesting characters were in this story such as the witch, the animals..."

The stories with accumulative structures and repetitive sentence patterns directly enhanced the participants' comprehension of the story and then indirectly benefited their involvement in English learning. The participants were more willing to give feedback when they could follow the repetition and gradually grasp the accumulative plots as the story progressed.

The Use of Props

The use of props aroused the participants' curiosity instantly. During the storytelling time, their eyes lit up when the props were shown. And, just in the beginning of the storytelling, some participants could not wait to criticize the props while some wanted to

know how the props were made. Table 4 is the dialogues that show the participants' reaction to the props during the storytelling process of *The Sausage*:

Table 4 Classroom Dialogues in the Storytelling Process of *The Sausage*

Teacher: One day, a fine lady came to the old woman's house. (Showed the prop of the fine lady)
Jacky: OLD (Said loudly)
Teacher: Right! She is old. Jacky, you are so great! The fine lady is old. Is she poor? (Pointed to the fine lady)
Joy: 很明顯的，身價差太多了。
Students: (Chatted about something.)
Teacher: Um... Any questions?
Sean: 為什麼她那麼胖? (Pointed to the fine lady.)
Teacher: Oh! Because she is a fine lady. She is rich so she has delicious food to eat. 她比較有錢，所以可以買比較好吃的東西吃。有可能是這樣所以她比較胖!
Students: (Laughed)
Teacher: (Continued to tell the story.) One day a fine lady came to the old woman's house. (Put the prop of a humble house on the blackboard)
Jim: (Pointed to the humble house) 為什麼屋頂會歪歪的?
Teacher: Yeah! Don't forget the old woman is poor so her house is also very humble!

Table 6 (Continued)

Teacher: The old woman said, "I wish I could have a sausage." Thump! A sausage dropped onto the table. (Stuck the prop of a sausage on the blackboard.)
Jacky: 那個香腸怎麼長得像蚯蚓?
Students: (Laughed loudly.)
Teacher: Ha! 不要懷疑! It's truly a sausage. ...Then the old woman's husband came home and found his wife had a wish for a sausage. (Showed the prop of the husband.)
Lee: 他是和尚嗎? (Pointed to the husband's bald head.)
Teacher: Ha! He is just bald not a monk. 他只是禿頭而已啦!
Students: (Laughed)

The participants expressed their high interest in every detail of the props, which excited them to keep asking many questions and joked about the props. Although some questions seemed to interrupt the storytelling, the props had succeeded in directing the participants' attention in the storytelling process.

In addition to arousing the participants' attention, the use of props played a crucial role in enhancing their comprehension in English storytelling. As Rainbow commented on the storytelling of *The Seven Chinese Brothers* in the fourth week:

I like this story (*The Seven Chinese Brothers*). Although I don't have a good understanding in English, I could guess the meaning from the props of the seven Chinese brothers. They were so great in helping each other with their extraordinary powers.

It showed that the props gave clues for her to follow the storyline and guessed the possible messages. In other words, the use of props as nonverbal cues facilitated comprehension even though they did not fully understand the language, they could grasp the message.

Extending Activities

Extending activity was also a crucial factor influencing the participants' involvement in learning English through storytelling. After storytelling time, diverse extending activities based on the core of the stories were being carried out, such as role playing, games, brainstorming, storybook rewriting, and group competitions. The purpose of the implementation of extending activities was not only for pleasure but also for practicing the target language that the participants had learnt from the storytelling time. It was found that the storytelling itself provides a meaningful context where learners absorb the target language through interesting narratives rather than memorization of the linguistic forms. And as the extending activities were carried out as a form of student-centered learning, it not only fostered the absorption of the target language but also aroused their interest in learning.

Many participants expressed that the extending activity was one of their favorite parts in the storytelling class. Gino said that "I like the storytelling class, because we can listen to interesting stories and play exciting activities." Mike also commented that "The storytelling class was great! I love to come here. All of the stories and the extending activities are interesting to me." The activity for the story, *The Skeleton Hiccups*, was a good example to illustrate how the participants were engaged in the learning from the activity. The story was about Skeleton waking up with the hiccups and the hiccups were not making things easy for him. When brushing his teeth, his jaw flew off! While polishing his bones, the hiccups caused his arm to come undone! His friend, Ghost, suggested several ways Skeleton could do to get rid of them, but nothing worked until Ghost found a mirror and held it up to Skeleton's face. When seeing his reflection, Skeleton screamed in fright and those hiccups jumped away from Skeleton.

After telling the story, the participants were asked to rewrite the story by changing the characters and making a storybook for a new story. A rewritten example, *Spiderman with Hiccups*, was shown to them (See Appendix G), then forty minutes were given to them. It was rewarding to see the participants could create and deploy their new ideas and topics into their stories but still keep the structure of the original story. They also tried to put in as many English words as they had learned. The extending activity enabled them to learn how to apply the already-learnt vocabulary to a real task.

Impact of Affective Factors

According to the questionnaires that enquired the participants' attitudes toward the storytelling club, 95% of the participants agreed that good interaction with the teacher was also a factor that engaged them in the storytelling. They commented that being able to interact with the teacher was a relaxing and interesting experience, which was different from their regular English classes. Peter commented that, "I feel that the storytelling class was more interesting than the regular class, because the English teacher in the regular class is very serious and rarely has interaction with us." And Jack said, "Coming to the storytelling class was great, since Miss Chen allowed us to share opinions freely in class. It was freer than our regular class." As noted above, in the participants' past experiences, English teachers have been figures of authority so the teaching flow was mostly one way traffic from teacher to

students without much feedback generated from students, which may result in a communication gap between students and teacher and make students anxious and stressful.

On the contrary, the intimate interaction between the students and the teacher in the storytelling club was warm and comfortable. Rainbow put it this way: “Each time I came here I always felt warm, because I could leave my daily pressure behind and pay careful attention to listen to the teacher’s storytelling.” Indeed, the role of a teacher in storytelling class is not an authority but a storyteller to mediate the story and to facilitate the learning. The good interaction between teacher and students had transformed their relationship and also the dynamics of the classroom.

The Difficulties Encountered while Employing Storytelling in Class and the Corresponding Solutions

It was expected that to implant storytelling as a new approach to improve underachievers’ English learning, certain difficulties would be encountered. Table 5 summarizes the difficulties encountered and solutions adopted during the storytelling class.

Table 5. Difficulties Encountered and Possible Solutions

Type	Problems	Solutions
Participants’ Low English Proficiency	The participants were unable to comprehend the storytelling in whole English.	<ul style="list-style-type: none"> ● Lower the anxiety ● Focus the use of English on key phrases and sentences. ● Use props to aid the participants’ comprehension.
Participants’ Short Attention Span	Storytelling longer than 15 minutes resulted in less attention from the participants.	<ul style="list-style-type: none"> ● Control the storytelling time. ● Provide extending activities for the participants.

The first was the participants’ low English proficiency that resulted in their learning anxiety. Two solutions were attempted. One was to focus on key English words and limited phrases in order to lower their anxiety; the other was to incorporate props into the storytelling process to support their comprehension. Most of the participants’ anxiety about the English comprehension difficulty was alleviated after a few weeks, because they were able to grasp the gist of the story content by the various types of non-verbal input, which included the storyteller’s body languages and dramatic intonations. When the participants understood that their comprehension could be sustained, they became increasingly comfortable and thus more willing to participate in the storytelling class. Most of the participants expressed that learning English through storytelling was totally different from their past learning experiences. As Rainbow commented:

I like the way of learning English through storytelling, because I can comprehend more what the teacher taught in English storytelling club than in regular class. In regular English class, I always feel very nervous, since the teaching material is too difficult for me and I cannot catch up with my peers. However, in the storytelling class, all of the stories are interesting and the teacher uses different ways to make it easier for me to understand the English stories.

The second difficulty identified was the participants' short attention span during the storytelling process. Two solutions applied were the control of storytelling time and providing extending activities for them. It was observed that the proper storytelling time should be controlled within 15 minutes; otherwise the participants' enthusiasm for the storytelling would diminish. A story that is longer than 15 minutes could gradually wore out the participants' patients; yet, a story shorter than 10 minutes might be too short to maintain the beauty of its narrative essence. Thus, how to shorten a long story but still keep its narrative essence would be a key to sustain the participants' attention span during the storytelling process.

As for the extending activities, as aforementioned, the participants enjoyed these activities, and during the period of storytelling class, the role of storytelling is similar to an appetizer to evoke these young adults' appetites to learn. Therefore, conducting extending activities may be the main course to maintain their attention and accordingly, enhance the learning.

CONCLUSION

This study has examined the process of applying storytelling as an approach and how effective it was to improve the motivation of underachievers. It reveals certain elements that could have improved the success of story teaching and also some student-centered activities to gear toward language learning. Although all of the participants expressed a positive attitude and high interest toward learning English through storytelling; however, they were not motivated to learn English more because the high interest did not automatically transfer into learning motivation if what and how they were required to learn in the regular class remained the same. That is, storytelling alone cannot change the motivation or even the learning results if other influential factors such as test-driven and performance-based language teaching remain the same.

While it is still encouraging that storytelling had brought some positive impacts to the participants, the long-term effective of storytelling remains uncertain. Thus, longitudinal studies are needed, particularly the integration of storytelling in regular curricula to examine the extent to which it influences learners' long-term motivation. Besides, the students in a regular class are composed of different proficiency levels. This phenomenon is different from this study where the participants were limited to low achievers. So, additional research focusing on heterogeneous groups would contribute to the understanding of storytelling as a regular approach in classroom teaching.

Further studies could also consider verifying the different effects between all-English storytelling and Chinese-English storytelling. The findings of this study showed that all-English storytelling gave rise to the underachievers' anxiety because of their comprehension difficulty. For this reason, Chinese-English storytelling was employed in the study. It is worth further comparing the effects of all-English and Chinese-English storytelling on learning of EFL students across different levels.

REFERENCES

- A little red house. (2004, March 26). Retrieved March 26, 2004, from <http://www.thevirtualvine.com/thelittleredhouse.html>.
- Burns, A. (1999). *Collaborative action research for English language teachers*. Cambridge: Cambridge University Press.
- Cary, S. (1998). *The effectiveness of a contextualized storytelling approach for second language acquisition*. (UMI Microform 9828452)
- Chou, C. T. (周中天)(2002)。英語成績差距懸殊的危機「The crisis of the distinctive gap between English learners」。康軒教育雜誌第三期。2002年10月11日, 取

- 自 <http://www.knsh.com.tw/epaper/m/003/22-1.htm>
- Chou, T. Y. (2006). *Effects of English storytelling on English learning of elementary school students*. Unpublished master's thesis, Kaohsiung Normal University, Kaohsiung, Taiwan.
- Collins, R., & Cooper, P. J. (1997). *The power of story: Teaching through storytelling*. Needham Heights, MA: Allyn & Bacon.
- Cuyler, M. & Schindler, S.D. (2005). *The Skeleton Hiccups*. Fort Worth, TX: Aladdin Publishing.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.
- Gardner, H. (1993). *Multiple intelligence: The theory in practice*. New York: Basic Books.
- Gardner, R. C. (1972). *Attitudes and motivation in second language second language learning*. Rowley, MA: Newbury House.
- Gardner, R. C. (1985). *Social psychology and second language learning*. London: Edward Arnold.
- Hamilton, M. & Welss, M. (2005). *Children tell stories: Teaching and using storytelling in the classroom*. Katnah, NY.: Richard Owen.
- Huang, H. L. (2006). The effects of storytelling on EFL young learners' reading comprehension and word recall. *English Teaching and Learning*, 30(3), 51-54.
- Hendrickson, J. M. (1992). *Storytelling for foreign language learners*. (ERIC Document Reproduction Service No. ED 355824)
- King, C. (2001). Many storytellers use props: In ins and outs of using props. Retrieved July 6, 2008, from <http://www.suite101.com/article.cfm/storytelling/77144/2>
- Koenig, J.M., & Zorn, C.R. (2002). Using storytelling as an approach to teaching and learning with diverse students. *Journal of Nursing Education*, 41(9), 393-399.
- Lee, F. H. (2004). *An investigation of middle school EFL teachers' beliefs in and attitudes toward English teaching in Southern Taiwan*. Unpublished master's thesis, Southern Taiwan University, Tainan, Taiwan.
- Leuck, L. (2006). *One witch*. London: Bloomsbury.
- Mahy, M. & Tseng, J. (1990). *The seven Chinese brothers*. New York: Scholastic Inc.
- Mandel, H.P., & Marcus, S.I. (1988). *The psychology of underachievement: Differential diagnosis and differential treatment*. New York: John Wiley & Sons.
- McNiff, J. (2002). *Action research: Principles and practice*. London: New York Routledge.
- Mello, R. (2001). The power of storytelling: How oral narrative influences children's relationships in classrooms. *International Journal of Education and the Arts*, 2(1).
- Mosel, A. & Lent, B. (1989). *Tikki Tikki Tembo*. New York :H. Holt and Co.
- Oxford, R., & Shearin, J. (1994). Language learning motivation: Expanding the theoretical framework. *The Modern Language Journal*, 78(1), 12-28.
- Taback, S. (1997). *There was an old lady who swallowed a fly*. New York : Viking.
- Trousdal, A. M. (1990). Interactive storytelling: Scaffolding children's narratives. *Language Art*, 67(2), 164-173.
- Trostle-Brand, S. L., & Donato, J. M. (2001). *Storytelling in emergent literacy: Fostering multiple intelligences*. Albany, NY: Thomson Delmar Learning.
- Yao, S. H. (2003). *Application of storytelling in an English teaching program in elementary school*. Unpublished master's thesis, Kaohsiung Normal University, Kaohsiung, Taiwan.
- Yolen, J. (Ed). (1986). *Favorite folktales from around the world*. New York: Random House, Inc.

- Hsu, Y. T. (徐玉婷) (2003)。國中生英語焦慮、英語學習動機與英語學習策略之相關研究。國立成功大學教育研究所，未出版，嘉義。
- Kao, W. C. (郭文泉) (2003)。高/低成就學生在克漏字測驗作答策略上的差異：個案研究。國立高雄師範大學英語學系碩士班碩士論文，未出版，高雄。
- Liou, M. C. (廖茂村) (2002)。高低成就學生學習活動之研究：以高雄市英明國中為例。國立高雄師範大學教育研究所碩士論文，未出版，高雄。
- Tsai, C. L. (蔡瓊玲) (2004)。TPR 策略應用對國小學童英語聽說學習成效之影響：以說故事教學為例。國立台北師範學院兒童英語教育研究所，未出版，台北。

APPENDIX A: EFL Learning Motivation Scale

英語學習動機問卷調查

姓名：_____

親愛的同學您好：本問卷能幫助您了解自己的英語學習狀況，並將結果做為未來英語教學參考。請在作答前先寫上您的姓名，並仔細閱讀作答說明後，再開始作答。作答時請看清楚每個題目的意思，再勾選出您認為最適合的答案；答案絕無所謂對與錯，只要選出您個人的感受即可。而這些資料純屬學術用途之用，因此您的回答也將會被完全的保密，請安心作答。

作答說明：請根據您的「實際狀況」來回答下列問題，並在您認為最適合的答案方框「打勾」選出。

		非 常 符 合	符 合	全 不 符 合	完 不 符 合
1. 我覺得學習英文是非常重要的。-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. 我學習英語，是因為可以考上好學校。-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. 我學習英語，是為了將來可以找到好工作。-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. 我學習英語，是為了在學校考試得高分。-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. 我學習英語，是因為能在國外旅行遇到困難時有助於溝通，例如：迷路。-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. 我學習英語，是因為未來有許多機會會用到它，例如：國外留學、工作或旅行	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. 學習習英文對我來說是非常有趣的。-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. 我學習英語，是因為可以了解不同的文化。-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. 我學習英語，是因為英語能力好會讓我跟別人比起來比較酷。-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10.我學習英語可以使我比較有自信。-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11.我學習英語，是因為嚮往國外文化。例如：喜歡國外的電影、歌曲或喜歡和外 國人相處。-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12.在英語科方面我比其他人用功。-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13.我會仔細做英語作業。-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14.我會仔細訂正作業、考卷。-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15.我平常會主動唸英語，不會等到考試才唸。-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16.我每次英語考試，都會認真準備。-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17.我常練習英語課文，對話以提升說的能力。-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18.在聽英語廣播、新聞或歌曲時，我會仔細去聆聽。-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19.我常閱讀英語雜誌、書籍或歌詞。-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20.在閱讀英語雜誌、書籍或歌詞，我會努力去看懂其中的意思。-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21.我會找機會和外國人或外國老師交談，以增強說的能力。-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
22.我會以英語和家人或同學校談。-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
23.我會試著將學過的英語單字或句型用在日常生活中。-----	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

◎本問卷的作答到此結束，感謝您的耐心與合作。祝您有個美好的一天！✿

APPENDIX B: Classroom Questionnaire

課堂意見表

1. 你喜歡今天這個故事嗎？為什麼？故事的哪個部分最吸引你？
2. 你喜歡今天的活動嗎？為什麼？
3. 有沒有什麼要給老師的建議呢？

★感謝妳／你的寶貴意見喔～

APPENDIX C: Questionnaire about the Attitude toward the Overall Project of Storytelling

英語故事社學員心得回饋單

首先，在此謝謝你熱請參與本次英語故事社活動。為了讓老師能夠在未來的課程上能有更完善的計畫，請同學能幫老師具體回答下列問題，做為老師未來教學上的參考。

1. 參加完這一學期的英語故事社活動，你最大的收穫是什麼？

2. 我們所講過的故事中，你最喜歡哪一個？為什麼？

3. 我們在英語故事社所做的活動中，你最喜歡哪一些？為什麼？

4. 你覺得來英語故事社哪些地方是吸引你的？
 - 聽到有趣的故事
 - 和同學課堂交流（例如：和同學哈拉、認識新同學…）
 - 和老師上課互動
 - 毫無吸引我的地方
 - 其他

_____（請具體說明）

5. 透過聽故事來學英文，你覺得這樣的方式你喜歡嗎？為什麼？

6. 你希望班上的英語課有機會也可以用說故事的方式來上嗎？為什麼？

7. 如果以後還有英語故事社的活動，你會想來參加嗎？

English Preservice Teachers' Perspectives, Computer Self-Efficacy and Computer Anxiety between Cultures

Wen-chi Huang (黃文志)
National Taiwan Normal University
897210220@ntnu.edu.tw

The current study examined and compared the English preservice teachers' perspectives and choices about using computer to teach English, computer self-efficacy and computer anxiety between eastern and western cultures. In other words, it put stress on investigating the correlation and coefficient of the variables mentioned above between eastern and western English preservice teachers. Furthermore, the study tried to find out the preservice teachers' characteristics of influencing the frequency of computer use during their teaching internship. Participating the study were fifty Taiwanese English preservice teachers who have received the educational programs of teaching high school students. As for the western subjects, forty American English preservice teachers, graduated from the educational department and most of them would teach English to international students who treat English as a second or foreign language. Two sets of questionnaire based on related studies (Capel, 1997; Marcoulides, Stocker, & Marcoulides, 2004) were designed to test and evaluate the participants' perspectives and choices about using computer to teach English, computer self-efficacy and computer anxiety. Findings indicated that quality and quantity of prior experience influenced teachers' perspectives and choices about using the computer to teach English. Frequency of computer use in the classroom was related to computer anxiety and number of types of computer training. Generally, the English preservice teachers were most concerned and worried about tasks demands and damaging the computer hardware and software, rather than socially anxious.

INTRODUCTION

It is widely believed that computers can make positive contributions to school curriculum, especially in the upper grades. Since the early 1980s, technology has been a focus for educators and teacher training programs. In the late 1980s, studies began to indicate how effective teacher training programs were in preparing teachers to implement computers (Whetstone & Carr-Chellman, 2001). As a matter of fact, the pervasion of computer not only accelerates the information exchanging, but also has great impacts on people's life style. Gradually, the ability of using computer has already turns out to be the critical skills for our teachers and students.

The positive goal envisioned for computer technology cannot, however, be realized by computers alone; computers are only a part of complicated scenario of educational change. The key element in the change process is the teachers and also, they need to be considered not only as objects of change, but also as change agents who can transform their interpretation for their practices through the use of computer technology (McGrail, 2005). Usually, teachers are regarded as the bridge among the students and knowledge. In this way, if they do not have enough computer equipment and capability of using computer to cope with their job, the teachers definitely have no proper means to make use of the computer to assist their teaching, not mention enriching the teaching content and materials. Moreover, in the study by Rosen and Weil (1995), they found that teachers' negative attitudes for using computer tend to have depressing influences on students' attitudes to use the computer to learn. In other words, no matter what the reason is, if the teachers do not consider the computer as an effective tool for assisting students' learning, then they are inclined to convey the pessimistic ideas to their students unconsciously.

As mentioned above, sufficient computer capability for classroom instruction has already been regarded as the basic ability for a qualified teacher. In reality, the application of computer during teachers' instruction does help to the efficiency of teaching and learning. On the one hand, the computer can be treated as a medium for combining all the teaching materials together; on the other hand, surfing on the Internet may provide teachers the opportunities to search more useful resources to assist students' learning and make their teaching materials more colorful.

Little attention has been given to teachers' beliefs about computer technology and their experiences with it in their practices (Becker & Ravitz, 1999; Dodson, 2000), especially for preservice teachers. However, preservice teachers' pedagogical perspectives and emotional attitudes of using computer to teach English may be interesting factors to influence their classroom instruction, especially in Taiwanese context. In this current study, English preservice teachers' perspectives are brought into the discussion on computer application in the classroom. By exploring preservice teachers' perspective and their origins as determined by their positions and responsibilities during the teaching practicum, the study contributes to developing an understanding of English preservice teacher perspectives and teaching anxiety. This understanding is significant in informing research and practice on how to synchronize educational change with English preservice teachers' real perceptions, their actual capabilities, emotional attitudes, and needs for desire for change.

Research questions

1. What are the relationships between computer self-efficacy and computer anxiety level of Taiwanese and American English preservice teachers?
2. What are Taiwanese and American English preservice teachers' perspectives and emotional attitudes toward computers in the classroom?

LITERATURE REVIEW

English preservice teachers' perspectives on computer in the classroom

Little is still understood about English preservice teachers' perceptions of computer application as reported by preservice teachers themselves. To date, only a few empirical studies have explored English preservice and inservice teacher perspectives on computer integration into classrooms (Barrell, 1999; Traubitz, 1998). A recent qualitative study in several high schools in San Francisco Bay Area by Peck, Cuban, and Korkpatrick (2002), reported a continuing situation that many English language Arts (ELA) teachers were reluctant to integrate computer technology into their practices. They also pointed out two major reasons: first, the teachers were unconvinced that computer technology would aid their primary mission, that is, educating students in their particular subject matter; second, they perceived computer technology integration as a secondary rather than primary goal in their instruction due to more pressing priorities and concerns they felt they needed to attend to, such as producing higher graduation rates and improving scores in standardized testing.

Goodwyn, Adams, and Clarke (1997), in an interview study on preservice and inservice English teachers' perceptions of computer technology, found attitudes to be an obstacle to computer technology integration in some ELA classrooms. According to the researchers, 32% of teachers studied were ambivalent about information technology. Goodwyn, Adams, and Clarke argued that the ambivalence of "The Unresolved" stemmed from their concerns about information technology's negative influence on student culture in and outside school. These teachers blamed computer technology, especially games, for moving students away from literature, defined by them as book-based culture, to what they called a leisure-based culture, where reading, writing, and literature exploration were portrayed as difficult rather than attainable tasks (McGrail, 2005). At the same time, these teachers were eager to obtain further support and training in technology, for they saw its potential to motivate the less able

or to promote interactivity in English.

On the contrary, another group of researchers proposed that some preservice and inservice teachers considered computer technology as a positive element in the classroom. McDevitt (1996) described using interactive videoconferencing between preservice teachers and a formal classroom teacher, in which preservice teachers observed a lesson and then interacted with the teacher after the lesson. McDevitt (1996) found that, provided opportunities, preservice teachers would like to share a common virtual field experience and reflect by way of interactive videoconferencing with formal teachers. By structuring videoconferencing experience for all participants, preservice teachers actively interacted with students and their teachers both during and after the lessons they were “observing”.

Researchers (Balli & Diggs, 1996) proposed that, given the constructivist belief that people learn by doing, preservice teachers were expected to infuse technology into their teaching and learning. In other words, technology infusion activities were driven by a constructivist pedagogy where preservice teachers were constructors of their own knowledge, as the author aimed to expand the preservice teachers’ vision of constructivist teaching and technology infusion.

English preservice teachers’ self-efficacy and anxiety on computer in the classroom

Talab and Newhouse (1993) proposed that the application of multimedia, especially the technology-integration stuffs, plays significant role during the instruction, and teachers who believe that they have successfully integrated new technology tend to be teachers who successfully integrate technology into their instruction (Riggs, 1988). Ropp (1999) also shared the same idea that teachers who believe in the utility of technology in their instruction may preserve through challenges that face the novice technology users. Based on this point we may theoretically state that inexperienced teachers cannot make use of the equipment or multimedia very well to facilitate their instruction because it’s not possible for them to be familiar with a new environment and some external factors in the classroom at the beginning year. Definitely, those pre-service teachers seem to face the same situation with the inexperienced ones. McInerney, McInerney, and Sinclair (1994) also pointed out that negative cognition for using computers may accompany the feelings of anxiety, including worries about embarrassment, looking foolish or even damaging the computer equipment.

According to Beach and Pearson (1998), preservice teachers vary considerably in their ability to grapple with the conflicts and tensions of applying technology in the classroom (Furlong & Maynard, 1995). Some simply avoid or minimize conflicts and tensions by conforming to the system. Others may be so overwhelmed they perceive conflicts and tensions as unmanageable, leading to a sense of resignation, futility, or flight from the system (Beach & Pearson, 1998). Or, they may openly acknowledge conflicts and tensions, but attempt to resolve them from the perspective of short-term expedencies, as opposed to reflecting critically on their own beliefs and theories about teaching and learning (Kettle & Sellars, 1996). They may assume, for example, that they can resolve a tension between their didactic instructional style and their pupils’ disinterest by making their lectures more “lively,” as opposed to considering the larger issue of their basic instructional approach.

Horwitz (1996) had mentioned that English teachers have many things to be anxious about: unruly students, challenging to their authority and competence, inflexible performance standards, unfamiliar with the equipment and environment, a complaining public, and unfortunately, many others. Based on this idea, it can be speculate that when English teachers’ general teaching efficacy is lower, they may mutually feel anxious for not knowing what to do in the classroom. The whole situation is like a vicious circle because when the external factors start to influence teachers’ instruction, they tend to feel deeply nervous and worried if they do not have enough experience or knowledge to deal with the problems. Finally, the feeling or attitude of anxiety will continue to become a great obstacle which

distracting their efficacy during the instruction.

Horwitz (1996) also stated that English teachers who suffer higher levels of foreign language anxiety will have a tendency to use English less. In this way, for one thing, students would lack the opportunities to train their listening ability during the instruction; for another, students would consider their teachers as a model of feeling reluctantly to using English to communicate with their peers in the classroom. As said by Horwitz (1996), moreover, English anxiety can inhibit a teacher's ability to effectively present the target language, interact with students, and serve as a positive role model as a Language learner. Finally, it will turn out to be that anxious English teachers may tend toward linguistic interactions that are predictable and more easily controlled. Also, based on my personal opinion, English teachers with higher levels of anxiety may communicate negative messages about language learning to their students.

Parker and Guarino (2001) found that among internship, preservice and inservice teachers, all of them tend to have lower anxiety and feel more confident when the Personal Teaching Efficacy (PTE) is high. Based on Parker and Guarino's research, it can be interpreted that higher Personal Teaching Efficacy (PTE) plays a positive role to reduce the feeling of anxiety during the English instruction regardless of the extent of teaching experience. Furthermore, English teachers who are not familiar with the environment or the equipment of the classroom have more possibilities to feel anxious when giving the instruction. Thus, Ross (1995) had some suggestions for improving preservice and inservice teachers' GTE including establishing mentoring programs and strengthening the preparation in the area of classroom discipline.

Ropp (1999) proposed that attitudes and beliefs about teaching in the use of computer in a subject matter area can vitiate well-planned instruction in teacher preparation. In other words, English teachers' negative attitudes or feelings of anxiety when using computer may affect their self-efficacy toward English instruction. In fact, an English teacher completes the required course of study to become certified to teach English does not mean that his or her language learning is complete. On the contrary, upon experiencing the demands on their linguistic skills posed by classroom teaching, many new coming teachers immediately recognize the need to improve their English proficiency. It is almost as significant as the teachers' sense of confidence in using the English in the classroom. As long as the teachers feel anxious or have not enough confidence in giving the English instruction, their self-efficacy tends to be low because they might be afraid of not knowing what to do to deal with the problems they encounter.

METHOD

Participants

In the current research, 50 Taiwanese English preservice teachers were invited to join the study, and all of them had already finished the teacher training program and the teaching practicum. Some of these Taiwanese English preservice teachers worked to teach English in the cram school. The participants' age ranged from 20-25, and some of them were still majoring the graduate school.

As for the USA English preservice teachers, 40 teachers participated in the study. Also, all of them had already finished the teacher training program. The participants' age ranged from 20-30 because some of them had been a full-time worker for several years before they finished the teacher training program.

Instrument

A set of Questionnaire with six-point liker scale was applied to the participants to evaluate their computers anxiety level in the classroom, and their computer self-efficacy when using the computers for instruction. Data analysis of this part would mainly be the

quantitative statistics to examine the correlation and coefficient between English preservice teachers' computer self-efficacy and computer anxiety level.

Moreover, an informal interview with the participants was applied to evaluate their personal perspectives toward the use of computer technology on English instruction. To be more specific, the researcher did an interview with those participants respectively who were willing to share their perspectives toward the use of computer during instruction. In this way, data analysis of this part would be mainly qualitative interpretation.

Procedure

The Taiwanese English preservice teachers came from the north and central part of Taiwan. Thus, it took a whole week for all of the Taiwanese teachers to finish the questionnaire. Moreover, it took almost one month to finish the interview with the participants who were willing to share their perceptions.

In terms of the USA English preservice teachers, they were given the questionnaire to finished next semester and it took three days to collect the quantitative data. Also, those who were willing to share their perceptions of using computer to teach English in the classroom were given the interview respectively, and it took two weeks to collect the qualitative data.

RESULTS

English preservice teachers from Taiwan

With little training and experience with computers, Taiwanese English preservice teachers resented what seemed to be a substantial disparity between confidence level and actual usage (20%). Only 10% of Taiwanese English preservice teachers reported having computers integrated into their curriculum, and 4% of them even mentioned that they never applied the computer into their curriculum during the teaching practicum.

According to the results, it indicated that the coefficient of correlation of Taiwanese English preservice teachers' computer self-efficacy and computer anxiety level is $-.767$ ($p < .01$). In this way, in relation to the data, it can be interpreted that there exists high negative correlation between Taiwanese English preservice teachers' computer self-efficacy and computer anxiety. That is to say, among the participants, if the English preservice teachers tend to have high computer self-efficacy, then it has lower possibility that they will feel nervous or anxious during the English instruction through the use of computer. On the contrary, if the English preservice teachers have higher anxiety when using the computer to teach English, their personal computer self-efficacy is inclined to become lower.

English preservice teachers from the USA

With little training and experience with computers, The USA English preservice teachers resented what seemed to be a substantial disparity between confidence level and actual usage (30%). In addition, 40% of the USA English preservice teachers reported having computers integrated into their curriculum, and most of them (95%) had the experiences to use computer during the instruction. The data showed that the coefficient of correlation of American English preservice teachers' computer self-efficacy and computer anxiety level is $-.638$ ($p < .05$). The results were quite similar with that of Taiwanese English preservice teachers, with slightly lower negative relationship.

Some participants from the USA would like to question the usefulness of computer application in the classroom. They also admitted that they had fun making the PowerPoint, but their teaching did not really differ. One teacher also remarked, "I haven't seen major improvement because students may still be distracted and cannot really concentrate." What's more, both Taiwanese and the USA English preservice teachers all indicated one critical problem that computer application cannot overcome, that is, the proficiency levels of English abilities, and this made the English preservice teachers feel less willing to use the computers during the instruction.

DISCUSSIONS

English preservice teachers from Taiwan

By and large, the data demonstrated that Taiwanese English preservice teachers felt anxious about using the computer to teach English in the classroom because they were afraid of making mistakes and even damaging the computers. However, they also mentioned that some related computer training or workshop may be helpful for them to be more willing to make use of the computer technology during the instruction. Furthermore, these English preservice teachers had negative relationships between their computer self-efficacy and computer anxiety level, and most of them had low computer self-efficacy and high computer anxiety level by the time this study was carried out. Although they may not be willing to use the computer, most of the Taiwanese English preservice teachers considered computer application can trigger students' motivation which also helped them make the teaching process more excellent. More or less, several of the participants mentioned that the support from the school they finished the teaching practicum was limited although they had the enthusiasm to utilize the computer to fulfill the goal of using computer to teach English.

Moreover, Taiwanese English preservice teachers seemed to be more cooperative and collaborative in terms of sharing the teaching materials and experiences when using the computer to teach English because the sharing of materials and experiences can help them overcome the problem of low computer self-efficacy and high computer anxiety level. Also, it was discovered out that teachers who usually did not use the computer were likely to not encourage their students to do so. In this way, it can be concluded that teachers' computer capability and self-efficacy have great influence on their teaching in the classroom.

English preservice teachers from the USA

Generally, the American English preservice teachers showed less anxiety toward computer application for teaching English, compared with Taiwanese English preservice teachers. The former saw computers as an agent of change for future schools and more than 90% of them considered computers very important in classroom instruction. However, interestingly, the results also indicated that only 20% of these English preservice teachers voluntarily completed an English course with computers. If we consider the experiences of the English preservice teachers, however, a reasonable hypothesis can be suggested. Due to their limited experience, they knew very little about how to successfully implement computers into English instruction. What they saw as implementing computers and what the experts saw as implementing computers may be two extremely different things (Carr and Psztajn, 1996). Similar with Taiwanese English preservice teachers, the American English preservice teachers also indicated negative relationships between the computer self-efficacy and computer anxiety level. In other words, the more they thought of themselves as capable in using computers to teach English, the less they might feel anxious in using computers during English instruction.

The data of interview also showed that the American English preservice teachers usually worked independently during and after the classroom period. To be more specific, this group of participants usually worked on their own to find some teaching materials and to implement some teaching software through computers. This was quite different from the situation of Taiwanese English preservice teachers who usually would like to share some teaching materials and activities with other teachers. Moreover, based on the data gathered in this research, the underestimation of difficulties seemed to be present in the USA English preservice teachers' perceptions of computers. Because they did not know the pedagogy of implementing computers into their instruction, these English preservice teachers did not comprehend the training that is necessary. In fact, they need to understand the various methods of integrating hardware and software into the curriculum in powerful ways (Whetstone & Carr-Chellman, 2001). English preservice teachers also need to understand the

limits, extensions, and future of computer technology for classroom instruction (Whetstone & Carr-Chellman, 2001).

REFERENCES

- Balli, S., & Diggs, L. (1996). Learning to teach with technology: A pilot project with preservice teachers. *Educational Technology*, 36(1), 56-61.
- Barrell, B. (1999). Technology and change in Atlantic Canada's new secondary English language arts curriculum. *English Education*, 3(13), 231-247.
- Beach, R., & Pearson, D. (1998). Changes in preservice teachers' perceptions of conflicts and tensions. *Teaching and Teacher Education*, 14(3), 337-351.
- Becker, J., & Ravitz, J. (1999). The influence of computer and internet use on teachers' pedagogical practices and perceptions. *Journal of Research on Computing in Education*, 31(4), 357-382.
- Beyerbach, B., Walsh, C., & Vannatta, R. (2001). From teaching technology to using technology to enhance student learning: Preservice teachers' changing perceptions of Technology infusion. *Journal of Technology and Teacher Education*, 9(1), 105-127.
- Carr, A., A., & Psztajn, P. (1996). Hearing unheard voice: Teacher education as conceptualized by preservice teachers. *Teacher Education Quarterly*, 23(2), 35-44.
- Dodson, M. M. (2000). Monologic and dialogic conversations: How preservice teachers socially construct knowledge through oral and computer-mediated classroom discourse. *National Reading Conference Yearbook*, 49, 137-152.
- Doering, A., Hughes, J., & Huffman, D. (2003). Preservice teachers: Are we thinking with technology? *Journal of Research on Technology in Education*, 35(3), 342-361.
- Furlong, J., & Maynard, T. (1995). *Mentoring student teachers: The growth of professional knowledge*. New York: Routledge.
- Goodwyn, A., Adams, A., & Clarke, S. (1997). The great God of the future: The views of current and future English teachers on the place of IT in literacy. *English in Education*, 31(2), 54-62.
- Horwitz, E. K. (1996). Even teachers get the blues: Recognizing and alleviating language teachers' feelings of foreign language anxiety. *Foreign Language Annals*, 29(3), 365-372.
- Kettle, B., & Sellars, N. (1996). The development of student teachers' practical theory of teaching. *Teaching and Teacher Education*, 12(1), 1-24.
- McDevitt, M. A. (1996). A virtual view: Classroom observation at a distance. *Journal of Teacher Education*, 47(3), 213-222.
- McGrail, E. (2005). Teachers, technology, and change: English teachers' perspectives. *Journal of Technology and Teacher Education*, 13(1), 5-24.
- Peck, C., Cuban, L., & Kirkpatrick, H. (2002). Techno-promoter dreams, student realities. *Phi Delta Kappan*, 83(6), 472-480.
- Ropp, M. M. (1999). Exploring individual characteristics associated with learning to use computers in preservice teacher preparation. *Journal of Research on Computing in Education*, 31(4), 402-424.
- Rosen, L. D., & Weil, M. M. (1995). Computers availability, computer experience and technophobia among public school teachers. *Computers in Human Behavior*, 11(1), 9-31.
- Sudzina, M. R. (1993). Technology, teachers, educational reform: Implications for teachers preparation. Paper presented at the Annual Meeting of the Association of Teacher Educators, Los Angeles, CA.
- Traubitz, N. (1998). A semester of action research: Reinventing my English teaching through technology. *English Journal*, 87(1), 73-77.
- Whetstone, L., & Carr-Chellman, A. A. (2001). Preparing preservice teachers to use

technology: Survey results. *TechTrends*, 45(4), 11-18.

Beyond Communicative Competence: Using Instant Messenger in Intercultural Exchange

Shu-mei Hung (洪淑美)
Cheng-Shiu University
shumay.hung@gmail.com

Theory of Communicative Language Teaching (CLT), though being widely accepted and applied in foreign language education, has been criticized to neglect second language learners' identity, hold a homogeneous and fixed view on culture, and over-emphasize the "effective exchange of information" in communication at the expense of overlooking the role of "establishment and maintenance of human relationship" in it. The idea of "Intercultural Communicative Competence" was thus brought up as a construct beyond CLT by emphasizing second language learners' culture and identity and their skills of discovery and interaction in intercultural communication.

This study looks at how Web2.0 technologies (instant messengers and WIKIs) can be used to develop language learners' intercultural communicative competence, particularly their skills of discovery and interaction.

Five pairs of foreign language learners (five English as Foreign Language learners from Taiwan and five Chinese as Foreign Language learners from UK) were recruited for participating in the telecollaborative dialogue on each other's cultures. The analysis started with an ethnographic approach to realize how the specific characteristics of instant messengers can affect learners' practice of "skills of discovery and interaction" with their online interlocutors. Next, learners' instant messenger (IM) chat recordings were analyzed in a fine-grained way by using discourse analysis in combination with theories of question types, politeness, and social presence to examine learners' intercultural questioning skills and their strategies of negotiating interactional conventions in the communication.

In the presentation, I am going to discuss three findings: 1) effective intercultural questioning skills, 2) learners' different types and strategies of negotiating interactional convention, and 3) pros and cons of using instant messengers in intercultural exchange.

COMMUNICATIVE LANGUAGE TEACHING AND ITS PROBLEMS

The rise of communicative language teaching (CLT) since the late 1970s has moved the emphasis in foreign language education from grammatical or structural approaches to a more communicative and functional emphasis (Doye, 1996) and has focused on developing the learners' skills in communicative situations. However, its pedagogical framework has been criticized to be constructed on the basis of assimilating native speakers' ability to use and interpret language appropriately in the process of interaction thereby over-emphasizing the 'effective exchange of information' in communication at the expense of neglecting the role of 'establishment and maintenance of human relationship' in it (Byram, 1997).

Problems of CLT Byram's (1997) argued that the direct transfer of Hymes' theory of communicative competence to second or foreign language education was problematic. The reason is that Hymes' theory was constructed within the context of first language acquisition, in which only learners' heritage culture is involved. However, in the context of second or foreign language acquisition, learners are introduced to a new linguistic and cultural system after they have already acquired the cultural system embedded in their mother tongue. How to deal with the contact of two cultural systems within such a context was not taken into

consideration when applied linguistic theorists (e.g. Canal & Swain, 1980) adopted Hymes' framework in the formulation of a theory for second or foreign language learning. As a result, the theoretical basis of 'communicative language teaching' narrowly focuses on the assimilation of native speakers' linguistic and cultural competence, ignoring the significance of the social identities and cultural competence that learners already bring with them in the intercultural interaction. Another problem embedded in 'communicative language teaching', as argued by Byram (1997), is its sole focus on the effective exchange of information. According to Byram (ibid), successful communication is not judged solely in terms of the efficiency of information exchange. It is focused on establishing and maintaining relationships. He argued: "*The efficacy of communication depends upon using language to demonstrate one's willingness to relate, which often involves the indirectness of politeness rather than the direct and efficient choice of language full of information.*" (1997:3)

The difficulty of effective communication in the intercultural context is increased particularly because of the invisible cultural faultlines (Kramsch, 1993; 2003) or rich points (Agar, 1994) inherent in each other's system of linguistic expressions that can hinder the understandings between the interlocutors. In order to remove these barriers, we need something more than the linguistic, sociolinguistic and discourse competences for communication as claimed by theories of communicative language teaching.

Intercultural Communicative Competence Byram (1997) suggested that the more desirable outcome is a learner with the ability to see and manage the relationships between themselves and their own cultural beliefs, behaviours and meanings, as expressed in a foreign language, and those of their interlocutors, expressed in the same language. He thus proposed a model of 'intercultural communicative competence' by adding the element of 'intercultural competence' into the communicative competence framework. There are five components in Byram's framework of intercultural communicative competence (ICC).

Byram (1997: 50-53) defined each of the components as follows:

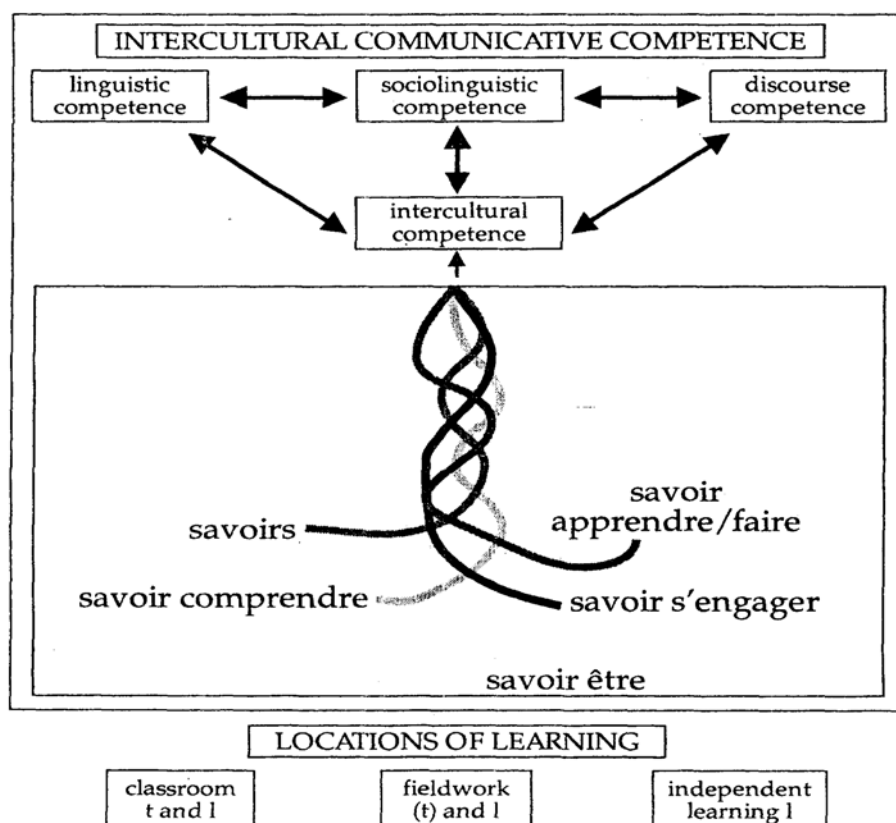
“Attitude (*savoir être*): curiosity and openness, readiness to suspend disbelief about other cultures and belief about one's own;

Knowledge (*savoirs*): knowledge of social groups and their products and practices in one's own and in one's interlocutor's country, and of the general processes of societal and individual interaction;

Skills of interpreting and relating (*savoir comprendre*): abilities of identifying ethnocentric perspectives in a document or an event; identify areas of misunderstanding and explain the pre-suppositions in a statement in order to reduce the dysfunction they may cause;

Skills of discovery and interaction (*savoir apprendre/faire*): abilities of acquire new knowledge of a culture and the operating knowledge, attitudes and skills under the constraints of real-time communication and interaction;

Critical cultural awareness (*savoir s'engager*): an ability to evaluate, critically and on the basis of explicit criteria, perspectives, practices and products in one's own the other culture and countries.”



Byram's (1997) framework of intercultural communicative competence

According to my own understandings of Byram's ICC framework and the rationale behind this framework, 'intercultural communicative competence' differs from 'communicative competence' in at least the following three ways.

Emphasizing the 'relationship establishment' in communication: The intercultural speaker-model emphasizes more 'establishing relationship' in the communication, not just 'the exchange of information'. Its objectives emphasize learners' ability to see through and mediate the differences between them and their interlocutors so as to negotiate and establish a mutually acceptable interactional convention, thereby establishing a mutually satisfactory relationship between each other.

Emphasizing the skills in learning about new culture: In order to address the 'multifaceted and fluid' nature of culture, the ICC framework emphasized 'method' or 'skills' in addition to the 'content' in learning about culture. Byram (1997) asserted that foreign language teaching should not attempt to provide representations of other cultures or introduce learners to a 'culture', to a particular combination of beliefs, behaviours and meanings "dominant" in a specific society but should concentrate on equipping learners with the means of accessing and analyzing any cultural practices and meanings they encounter.

This approach aims to provide learners with critical tools to explore and to develop their critical understanding of their own and other societies. In Byram's ICC framework, this aspect is seen in the component of "skills of discovery", the objectives for which is to "elicit from an interlocutor the concepts and values of documents or events and develop an explanatory system susceptible of application to other phenomena" and to "identify significant references within and across cultures and elicit their significance and connotations" (1997: 52-53). The importance of "skills of discovery" lies in its role of transforming language learners from passive receivers of cultural facts into active agents in negotiating and constructing cultural meanings. In a similar vein, Kramsch (1998) proposed

the need for the norm of the native speaker to be replaced by that of the ‘intercultural speaker’ who is not bound to fluency in the standard form of a language, but is instead able to negotiate and adapt to differing standards of appropriateness in order to engage in successful communication with others.

Balanced treatment of learners’ own culture and the target culture: Finally, the ICC framework has a balanced treatment between the learners’ own culture and the target culture. This can be seen in the components of “knowledge” and “skills of interpreting and relating”, in which learners are required to acquire knowledge not only about the target culture but also their own cultures so that they can have the ability to identify the ethnocentric perspectives as well as the areas of misunderstandings in both cultures. Learners are thus no longer expected to reject their own culture and take on the target culture, but rather to find what Kramsch (1993) describes as a ‘third place’. Kramsch suggests that learners need to locate themselves in a place which “grows in the interstices between the cultures the learner grew up with and new cultures he or she is being introduced to” (p. 236). This concept describes language learners’ newly achieved distance from both the home and target cultures and refers to the multiplicity of cultural identities which belong to all of us, thereby rejecting the fallacy of the essentialist idea that one nation equals to one culture.

Online Foreign Language Intercultural Exchange

The advent of internet technology has significantly enriched the teaching and learning methods a foreign language teacher can bring to his or her classroom. Particularly, it extends the scope of intercultural language learning from single local classrooms to the global educational settings. As Kern (2006:198) mentioned, “A recent development in network-based language teaching is a shift in focus from single classrooms to long-distance collaborations involving two or more classrooms, often in different countries. This shift expands the focus from language learning to an emphasis on culture (i.e., intercultural competence, cultural learning, and cultural literacy).” The ‘virtual connection’ enables learners to have an authentic experience of communicating with learners from other countries. Because of the linguistic and geographical differences, these two groups of learners are usually assumed to be of different cultural backgrounds as well. By actually interacting with people of different cultural backgrounds, learners are expected to develop their intercultural communicative competence from the experience of managing to communicate effectively with their interlocutors.

A number of studies have reported learners’ growth in intercultural competence while participating in tele-collaborative foreign language learning (Furstenberg et al., 2001; Müller-Hartmann, 2000; von der Emde, Schneider, & Kötter, 2001; Liaw, 2006). However, most of the previous studies tended to directly apply the ICC framework without further examining the framework itself. In addition, previous studies selected their participants mainly from the US or European countries such as Germany, France, Spain, Italy, UK, and Ireland. Few telecollaborative studies have been found to conduct the intercultural exchanges between European countries and Asian countries. Besides, the previous studies mainly used asynchronous tools such as emails and discussion boards for learner communication. Few studies were found to use synchronous tools such as instant messengers or videoconferencing.

Because of these issues, this study aimed to connect foreign language learners in the UK (Europe) with those in Taiwan (Asia) and attempted to investigate learners’ development of intercultural communicative competence (with a focus on the skills of discovery and interaction) when they interacted with target language speakers through synchronous tools – instant messengers.

Research Questions

The research focus was put on the use of discourse analysis to uncover learners' interactional and questioning/discovering strategies in their dialogue. In addition, as the intercultural exchange was conducted with the instant messenger-facilitated learning context, the effects of the technologies on intercultural tele-collaborative learning were also examined. The three research questions are:

1. How did the instant messenger-mediated learning context influence the online intercultural dialogue?
2. How did these learners negotiate and co-construct interactional conventions?
3. What questioning techniques are used by the learners to discover each other's culture?

Data Collection

Five pairs of university students (five from Taiwan and five from the UK) were connected by internet technologies including instant messengers (IM) and Blackboard WIKIs to explore each other's cultures through dialoguing. The procedure is as follows:

Step One (Pre-exchange questionnaire): Before the exchange activities were conducted, the participants were asked to fill in a pre-exchange questionnaire for me to collect the data about these participants' background and their availability of time for the intercultural exchange.

Step Two (Pairing the participants): Five pairs were formed mainly based on the participants' availability of exchange dates. I then emailed each participant the information about their partner and the dates arranged for them to conduct the online chat.

Step Three (Setting up project webpage): A course webpage was constructed on the Blackboard System of the UK university, in which I placed the course information and instruction and created WIKI group links for each pair which allowed them to construct WIKI pages with ease.

Step Four (WIKI page construction): Before the start of the chat, learners were required to construct two WIKI pages. On the first page, they described their home culture with a special emphasis on describing their current life (school and leisure) as well as their ambitions. On the second page, they were asked to describe their current experience, impression or understandings toward the target culture and mention the questions or areas they would like to explore regarding the target culture. The construction of the WIKI pages provided learners with the opportunity to reflect on their views toward their own culture and the other's culture. The content in these pages served as the resource for learners to find topics in their first chat.

Step Five (Read partner's pages and form ideas for talk): The participants formed their ideas for the online chat by reading their exchange partner's two WIKI pages. From their exchange partner's WIKI pages, learners framed their understanding regarding what the target culture was like from their exchange partner's point of view and also what questions or understandings their exchange partners held for their home culture.

Step Six (Conducting the online chat): After reading the WIKI pages, the participants conducted their first text-based talk session with their exchange partners. From the talk, they discussed and negotiated understandings about their home culture and the target culture. By doing this, they developed different perspectives toward their own culture and new understandings toward the target culture.

Step Seven: After each talk session, the participants were asked to forward the chat history to the researcher. At the same email, the participants were required to describe any problems they had encountered in the talk.

Repeat Step Four to Step Seven for five times: After the talk, the participants were asked to revisit their WIKI website and revise the content of their two WIKI pages according to the new understandings they had developed from the online talk. They were asked to modify the content on their home culture page by adding in the aspects they had been asked by their exchange partner, to discuss their new understandings on the target culture page, and to increase the breadth and depth of the questions they wanted to explore further toward the target culture. These changes then served as the new input for the next talk session, for which they could either continue the previous talk by discussing the old information or starting new topics from the new input they have added to the WIKI pages. Step Four to Step Seven was repeated for five times so in total there were five talk sessions conducted by each pair and each WIKI page had been revised for five times – the process of learners’ revision and reflective writing in these WIKI pages provided important data for me to confirm the findings generated from the analysis of learners’ IM chat transcripts.

Step Eight (Post-exchange questionnaire):

The post-exchange questionnaire aimed to investigate how the participants perceived this exchange project in terms of its course design and the cultural and technological impact it brought to them.

Data Analysis

The data collected included the students’ five-hour text-based chat recordings on IM software, their writings about their understandings of their own culture and the other’s culture on Blackboard WIKIs, pre-project questionnaire, email communication between the students and the researcher, and post-project structured interview via questionnaire.

The analysis started with an ethnographic approach to realize how the specific characteristics of instant messengers can affect learners’ practice of “skills of discovery and interaction” with their online interlocutors. Next, learners’ instant messenger chat recordings were analyzed in a fine-grained way by using discourse analysis in combination with theories of question types, politeness, and social presence to examine learners’ intercultural questioning skills and their strategies of negotiating interactional conventions in the communication.

Questions as an analytical tool: In seeking answers to the research question “*how did learners use questioning strategies to discover about the other’s culture*”, constructs and categorization of question types played important roles. The operationalisation of different question types helped me to characterize what the learners were doing in each turn of their conversation, which then formed the basis for me to interpret the strategies the learners were applying in the process of pursuing their understanding of their interlocutor’s culture. Following the framework used by Belz (2003) and Ware (2003, 2005) in their telecollaborative studies, the questions used by the learners in this study were initially analyzed with three main types of questions: information-seeking, information-checking, and rhetorical. According to Schiffrin (1994:165),

Information-seeking questions are asked when the speaker lacks knowledge of a particular state of affairs and wants to gain that knowledge by eliciting information from the hearer. They are divided into five subcategories based on lexical, syntactical and functional criteria: why –questions, opinion –question, what/how –questions, yes/no –questions, either/or –questions.

Information-checking questions are designed to confirm or disconfirm the hearer’s interpretation of a particular utterance. The information being sought is not the completion of a proposition but reception of a referent or proposition (1994:169-170).

Rhetorical questions are not designed by the questioner to elicit novel information – it is an illocutionary act that has the direct illocutionary force of a question and is not generally used with the expectation of an answer but with some different indirect force, such as a

command, a tentative statement, and an evaluation.

Social interaction strategies from politeness and e-learning theory: To address the research question “*how did learners establish relationship and establish interactional conventions*”, the categories for analyzing learners’ skills of interaction were borrowed from two frameworks: Brown and Levinson’s politeness theory (1987) and Garrison and Anderson’s (2003) construct of social presence in e-learning theory. The reason for adopting concepts from these two frameworks is because of their suitability for the type of data collected in this study. Particularly, the research followed Vinegre’s (2008) method of using Brown and Levinson’s framework for exploring learners’ social interaction as Vinegre (2008) argued that the strategies mentioned in Brown and Levinson’s politeness theory seemed to make a unique contribution to the construction of co-operative social interaction (Watts, 2003:47), which is exactly the focus of the analysis in this study. In addition to Brown and Levinson’s framework, I also drew on Garrison & Anderson’s framework of social presence in e-learning context. This framework is widely used in e-learning research (for example, Arnold & Ducate, 2006; Pawan et al, 2003), which is also the context of this study, and I found that some constructs in this framework can better represent the interactional strategies the learners have applied in this exchange, which Brown and Levinson’s framework does not address.

FINDINGS AND DISCUSSIONS

How the IM-mediated learning context influenced the intercultural exchange: The findings showed that synchronous mediums such as IM is suitable for learners to practice their “skills of discovery and interaction” because of its instant turn-taking feature, which enables learners to initiate immediate and sustained negotiation and repair in the dialogue for adjusting their communication styles and modifying their questioning strategies. Most learners reported that IM made them feel less threatened in the learning process, which allowed them to discuss what they really wanted to talk about instead of what they supposed they should say. Learners were, thus, engaged in “real” and “authentic” dialogue via instant messengers instead of the more monologue-like interaction on discussion boards. The “authenticity” induced by the use of IM in intercultural exploration encouraged learners’ release of their agency in collaborative learning in the following ways.

Firstly, they were found to actively make use of multi-modal functions of expression via IM channel to articulate their own identities. They made use of different online resources to help express aspects of their own cultures and to learn from or scaffold and negotiate with each other new meanings and methods of self-representations under technology-facilitated context. *Secondly*, the instant turn-taking mechanism made the establishment and maintenance of relationship an essential part in the interaction. Unlike asynchronous communication, real-time interaction has higher demands in “relationship building” between the interlocutors. Learners tended to start their conversation by a brief exchange of what was happening to them in their daily life. As a result, these real-life daily events served as another rich source for learners to get a glimpse of their interlocutor’s culture, through which learners actively identified the significant references and initiated negotiation in order to construct or reconstruct their understandings of their interlocutors’ identities.

Limitations of IM were also identified such as the lack of visual expression and being unsuitable for providing lengthier information in the conversation, which suggested that an integration of IM with other asynchronous tools such as WIKIs is necessary and other technologies that enable face-to-face or voice communication should be tried.

How learners established relationship and interactional convention with each other: Two types of conversational styles exhibited in two pairs were compared: task-oriented vs. personal-oriented. The ‘reciprocity’ in the use of interactional strategy” was identified as a

method that these learners relied on to negotiate the conversational styles with their interlocutors. The two pairs of learners displayed very different interactional patterns. The first pair tended to keep a formal and task-oriented style of interacting with each other while the second pair's talk appeared to be more personally oriented. The former pair normally started their conversation with a brief greeting and then the focus was shifted to the tasks in the exchange. Their closing also only comprised of the negotiation of topic and time for the next chat session, except that in the closing of the last chat session, they have given each other more compliments and appreciations for what their partner has helped them to learn in the exchange. The second pair, on the other hand, shared with each other a lot of details of their daily life in their conversation. They normally started their conversation by telling each other what had happened to them recently and there were more reciprocal responses and mutual appreciation in their closings. Therefore, a lot of "self-disclosure" was evident between the second pair, which contributed to the establishment of their more friend-like relationship in their exchanges, while the first pair did not talk to each other much about the details of their personal life; instead, they normally exchanged their opinions on the topics they have agreed to discuss about only.

Moreover, the first pair normally decided on a particular topic for each session and stuck to it throughout the whole session of interaction (although some flexibility can also be seen). The second pair appeared to be more casual in their conversation; they jumped from topic to topic; they did not negotiate to decide the topic for each session in advance; instead, the topics of discussion just emerged with the conversation flow; that is, the topics for discussion were normally generated from what they brought into the openings in their conversation by telling each other what had happened to them in their everyday life.

Previous telecollaborative studies indicated that national differences in communication genres could be a salient factor that caused miscommunication. By comparing the two pairs of learners' negotiating interactional conventions, the findings of this study suggest that individual differences in learners' expectations and purposes for this exchange activity may be more influential in the interaction than culturally or nationally-based discrepancies. The findings indicated that the interactional conventions were negotiated by the learners according to the specific context co-constructed between them instead of being pre-fixed or pre-given before the exchange.

How learners used questioning techniques to discover about their interlocutors' cultures: It was observed that learners with higher intercultural competence displayed their questioning strategies in this intercultural exchange in at least the following four ways. First of all, they tended to ask a series of questions focusing on a particular theme so that they could get deeper and more comprehensive understandings of a particular cultural phenomenon. They did this by expanding the scope of probing and trying to cover different aspects of a topic in his questioning. In addition, these learners were able to elicit the information they needed by moving from specific questions to general questions.

Secondly, these learners were able to make their questions more precise and easy to answer by adopting different question types in their probing. Previous studies (Jones et al, 2006; Black et al, 2003) mentioned that referential questions or open-ended questions are more able to elicit rich information than display questions or closed questions. However, it appeared that different types of questions are needed under different contexts of interaction. Yes/no questions can be used as a prompt for further discussion or a parameter to evaluate a situation while either/or questions can be used to clarify some possible different interpretations of the meaning in certain expressions. The use of particular question types depends on the context and the purposes of the questioning.

Thirdly, these learners managed to make their questions as understandable as possible by operating "skills of interpreting and relating" (Byram, 1997) to mediate the possible

misunderstandings created by the questions. This was demonstrated by their sensitivity in the possible differences in the linguistic expression of a particular concept used under different cultural contexts (e.g. the use of the word “gig” for “concert” in UK context and translating UK pounds into US dollars), which showed their critical awareness in the cultural diversities and their ability to de-centre themselves from the ethnocentric discourse (i.e. the way of expression from their own culture) and mediate between these differences in interaction (Byram, 1997).

Fourthly, these learners were able to reduce the possible face-threatening effects caused by particular questions by using mitigating strategies (Ware, 2003: 261) so that his questions could encourage more feedback and responses from their interlocutors. This was demonstrated by some of their statements they added before their probing questions, for example: *“i’m always wondering how much western music is listened to in other countries. i don’t like to presume that everyone listens to the same thing..”* Another example is their following up their request with further explanation when they realized that the expression or question they used might carry negative connotations; for example: *“is it easy to visit there (China)? i don’t really understand the relations between the countries; my apologies for that”* This echoes O’Dowd’s (2003) argument that when the topics involve more sensitive issues, it is important that learners can use proper strategies to mitigate the possible misunderstandings and to express their opinions in a less threatening but more interesting and understandable way for their interlocutors.

CONCLUSION

In summary, this research has contributed to the telecollaborative studies by examining the use of instant messenger in intercultural exchange, connecting European students with Asian students, and exploring learners’ skills of discovery and interaction, the pivotal component in Byram’s framework of intercultural communicative competence.

In terms of learners’ skills of discovery and interaction, I have found that learners could establish their relationships with each other through the use of different interactional strategies and negotiated their interactional conventions and their interactional conventions were constructed by both parties in the process of interaction, with the differences appearing to be more individually-based instead of culturally-based. In the process of analyzing learners’ questioning behaviours, I identified the following questioning techniques for exemplifying and operationalising the concept of “skills of discovery”:

- 1) Using a series of questions focusing on a particular theme instead of a bunch of unrelated questions
- 2) Making use of different types of questions to make the questions precise and easy to answer instead of using broad and general questions
- 3) Operating “skills of interpreting” to mediate the possible misunderstandings embedded in the questions so that the questions can be understandable
- 4) Operating “skills of interacting” to reduce the face-threatening acts caused by particular questions so that the questions can encourage feedback and responses from the interlocutors

The findings also suggest that IM-mediated learning context appeared to encourage learners’ active engagement (Ware, 2003) in their intercultural exchange with their learning partners. The synchronous feature of IM makes socialisation a natural part of interaction. In addition, its instant turn-taking feature appeared to facilitate learners’ interaction and the process of discovery. These seemed to suggest that IM could be an ideal platform for learners to practice the skills of discovery and interaction in the intercultural exchange in terms of the “learning agency” embedded within it.

REFERENCES

- Agar, M. (1994). *Language Shock: Understanding the Culture of Conversation*. NY: HarperCollins Publishers Inc.
- Alptekin, C. (2002). Towards intercultural communicative competence in ELT. *ELT Journal*, 56 (1), 46-54.
- Arnold, L. & Ducate, N. (2006). "Future foreign language teachers' social and cognitive collaboration in an online environment." *Language Learning and Technology* 10(1): 42-66.
- Atkinson, D. (1999) TESOL and culture. *TESOL Quarterly* 33 (4), 625–54.
- Bauer, B., deBenedette, L., Furstenberg, G., Levet, S, & Waryn, S. (2006). The Cultura project. In J. A. Belz & S. L. Thorne. (eds.) *Internet-mediated Intercultural Foreign Language Education* (pp. 31-62). Boston: Thomson Heinle.
- Belz, J. A. (2005) Intercultural questioning, discovery and tension in Internet-mediated language learning partnerships. *Language and Intercultural Communication* 5(1), 3-39.
- Brown, P. and Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge, UK: Cambridge University Press.
- Byram, M. (1997) *Teaching and Assessing Intercultural Communicative Competence*. Multilingual Matters: Clevedon.
- Canale, M., & Swain, G., (1981) Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1,1:1-47.
- Corbett, J. (2003) *An Intercultural Approach to English Language Teaching*. Multilingual Matters.
- Crozet, C., & Liddicoat, A.J. (1999) The challenge of intercultural language teaching: Engaging with culture in the classroom. J. Lo Bianco, A.J. Liddicoat & C. Crozet (eds), *Striving for the Third Place: Intercultural competence through language education* (pp. 113-126). Canberra: Language Australia.
- Doye, P. (1996). Foreign language teaching and education for intercultural and international understanding. *Evaluation and Research in Education*, 10 (2&3): 104-112.
- Furstenberg, G., Levet, S., English, K. and Maillet, K. (2001). Giving a virtual voice to the silent language of culture: The Culture Project. *Language Learning & Technology*, (1), 55-102. Retrieved 17 August, 2003: <http://ilt.msu.edu/vol5num1/furstenberg/default.html>
- Garrison, D. R. & Anderson, T. (2003). *E-Learning in the 21st Century*. London & New York: RoutledgeFalmer
- Herring, S. C. (2004). Computer-mediated discourse analysis: An approach to researching online behavior. In: S. A. Barab, R. Kling, and J. H. Gray (Eds.), *Designing for Virtual Communities in the Service of Learning* (pp. 338-376). New York: Cambridge University Press.
- Holliday, A., Hyde, M., & Kullman J. (2004) *Intercultural communication: an advanced resource book*. London : Routledge
- Hymes, D.H. (1974) *Foundations in sociolinguistics: An ethnographic approach*. Pennsylvania: University of Pennsylvania Press.
- Jin, L. & Erben, T. (2007). Intercultural learning via instant messenger interaction. *CALICO Journal* 24(2).
- Kern, R. (2006). Perspectives on technology in learning and teaching languages. *TESOL Quarterly*, 40(1), 183-210.
- Kern, R., Ware, P. & Warschauer, M. (2004). Crossing frontiers: New directions in online pedagogy and research. *Annual Review of Applied Linguistics*, 24, 243-260.
- Kramsch, C. (1993a) *Context and Culture in Language Education*. Oxford: Oxford

- University Press.
- Kramsch, C. (1998) The privilege of the intercultural speaker. M. Byram & M. Fleming (eds), *Language Learning in Intercultural Perspective: Approaches through Drama and Ethnography* (pp. 16-31). Cambridge: Cambridge University Press.
- Kramsch, C., and Thorne, S. (2002). Foreign language learning as global communicative practice. In D. Block and D. Cameron (Eds.), *Language learning and teaching in the age of globalization* (pp. 83-100). London: Routledge.
- Kubota, R. (1999). Japanese culture constructed by discourses: implications for applied linguistics research and ELT. *TESOL Quarterly*, 33 (1), 9-35
- Kubota, R. (2001). Discursive construction of the image of U.S. classrooms. *TESOL Quarterly*, 35(1), 9-38.
- Kubota, R. (2006). Critical approach to culture in English language teaching. In B. Beaven (ed), *IATEFL 2006 Harrogate Conference Selections*.
- Kumaravadivelu, B. (2007). *Cultural Globalization and Language Education*. Yale University Press.
- Lee, Y. (2006). Respecifying display questions: Interactional resources for language teaching. *TESOL Quarterly*, 40(4).
- Liaw, M. (2006) E-learning and the development of intercultural competence. *Language Learning and Technology*, 10(3):49-64.
- Markee, N. (2000). *Conversation Analysis*. NJ: Lawrence Erlbaum Associates, Inc.
- Mazur, J. M. (2004). Conversational analysis for educational technologists: theoretical and methodological issues for researching the structure, processes, and meaning of online talk in Jonassen, D.H. (ed) *Handbook of research on educational communications and technology*, Laurence Erlbaum Associates, Mahwah, New Jersey.
- O'Dowd, R. (2000). Intercultural learning via videoconferencing: A pilot exchange project. *ReCALL* 12 (1), 49-63.
- O'Dowd, R. (2003). Understanding 'the other side': Intercultural learning in a Spanish English e-mail exchange. *Language Learning and Technology* 7 (2), 118-144. Retrieve on 17 August, 2003: <http://ilt.msu.edu/vol7num2/odowd/default.html>
- O'Dowd, R. (2005). Network-based language teaching and the development of intercultural communicative competence. Unpublished doctoral dissertation: University of Essen, Germany.
- O'Dowd, R. (2006). The use of videoconferencing and e-mail as mediators of intercultural student ethnography. In J. A. Belz & S. L. Thorne. (eds.) *Internet-mediated Intercultural Foreign Language Education* (pp. 86-120). Boston: Thomson Heinle.
- Pillay, H., Clarke, J., & Taylor, P. (2006). Learning agency in new learning environments: an Australian case study of the influence of context. In A. Figueiredo & A. Afonso (eds.) *Managing Learning in Virtual Settings: The Role of Context*. London: Information Science Publishing.
- Pomerantz, A. & Fehr, B. J. (1997). Conversation analysis: an approach to the study of social action as sense making practices. In T. A. van Dijk (ed.) *Discourse as Social Interaction*. London: Sage.
- Schiffrin, D. (1994). *Approaches to discourse*. Oxford, England: Blackwell.
- Sercu, L. (2002) Autonomous learning and the acquisition of intercultural communicative competence: some implications for course development. *Language, Culture, and Curriculum*, 15 (1), 61-74.
- ten Have, P. (1999). *Doing Conversation Analysis*. London: Sage.
- Thorne, S. (2003). Artifacts an cultures-of-use in intercultural communication. *Language Learning and Technology*, 7(2):38-67
- Tudini, V. (2007). Negotiation and intercultural learning in Italian native speaker chat rooms.

- The Modern Language Journal*, 91(4): 577-601.
- Van Ek, J. A. (1987) *Objectives for Foreign Language Learning. Volume 2: Levels*. Strasbourg: Council of Europe.
- Vinagre, M. (2008). Politeness strategies in collaborative email exchanges. *Computers & Education*, 50(3):1022-1036.
- Vinagre, M. (2007). Social interaction in telecollaborative language learning. UNTELE conference selections.
- Ware, P. D. (2003). *From involvement to engagement in online communication: Promoting intercultural competence in foreign language education*. Unpublished doctoral dissertation, University of California, Berkeley.
- Ware, P. D., & Kramsch, C. (2005). Toward an intercultural stance: Teaching German and English through telecollaboration. *Modern Language Journal*.
- Ware, P.D. (2005). Missed communication. *Language Learning and Technology* 9 (2), 64-89.
- Warschauer, M. (1999). *Electronic Literacies. Language, culture and power in online*. New Jersey: Lawrence Erlbaum Associates.
- Warschauer, M. and Kern, R. (Eds.) (2000). *Network-based language teaching: Concepts and practice*. Cambridge: Cambridge University Press.
- Warschauer, M. (2006). Information Communication Technologies, Agency and Empowerment: An Interview with Mark Warschauer. *IATEFL CALL Review*, Summer 2006

Teaching Strategies in Online Writing Instruction by Pre-service Teachers

Chia-chi Kuo (郭佳淇), Charles Brown (白任遠)
National Chiayi University
jannyqh@yahoo.com.tw

According to the Report of the National Commission on Writing (2003) indicated that more than 50% students' writing ability is not sufficient for them to achieve in higher education or work places. While with the advance of the technology, it may provide the solution for the problem. Owing to Lyon County school students' unfavorable performance on state writing exam and its remote location, the online assistance project initiated to help students improve their writings, especially on revising. As Stallard (1974) mentioned a good writer would do more revisions, because it is a crucial step in the writing process that refines a work. In this research, the online project the researcher participated utilized the feature of the Internet which allowed pre-service English teachers in Murray State University to assist students improving their writing skills. The focuses of this research are on the pre-service teachers' teaching strategies, especially on giving feedbacks in online environment, and the impact of teacher training on their teaching practice. To discover these questions, the researcher did a case study with the utilization of content analysis and interview. After the project started, the researcher collected tutors' replies from the blog and categorize responses into different types. The interviews were conducted after every tutor posted their replies. The findings showed that the focus of tutors' response is on the form of the writing, and their criteria of making such decision basically depend on students' writing abilities. The inquiry of tutors' perception showed the influence of teacher training, but the findings also indicated there are other factors such as their personal experiences and believes. While, which factor has greatest influence on teaching practice does not discuss in this paper.

INTRODUCTION

Background

As the popularity of using the Internet, the new writing style has emerged. This new style has had a great impact on young generation's writing ability, and this phenomenon is now prevalent worldwide, rendering many countries to start making efforts in promoting students' writing ability. Some teachers blame the decline in writing skills on technology especially chat rooms; however, Grejda and Hannafin's (1992) research showed that computers actually make students produce more texts and revisions in their writings. Moreover, a recent report from an online discussion forum, organized by the International Association of Teachers of English as a Foreign Language (IATEFL) Learning Technologies Special Interest Group, mentioned that Computer Assist Language Learning (CALL) would become normalized in the future for all classrooms (Ioannou-Georgiou, 2006).

To improve students' writing ability, educators identify problems in schools' writing courses. Examining these courses, especially in the United States, Harper (1997) and Yoder (1993) indicated that students' inability to produce good works is due to not knowing how to revise their writings. Most students just want to complete assignments quickly without the thought of revising, even though revision is the most important step in the writing process which refines a piece of writing. Therefore, to promote students' writing ability, teaching revision is an important instructional part in the writing classes.

In spite of the negative influences that technology may cause, the uses of computers indeed provide an easier and "less painful" way for students to draft and revise. Even though

the Internet is a controversial tool in the language learning area for its impact of chat rooms, it indeed provides more opportunities for students and teachers to obtain the helps they need. As the new technology continuously appears and is popularized, this challenges educators' abilities to utilize them in their teaching.

Context

In the fall of 2006, the researcher obtained the opportunity to study in Murray State University (MSU), which located in the state of Kentucky, the United States. The researcher enrolled in Ms. Deborah Bell's course, Teaching English in Middle and High school. In this course, all of students are pre-service English teachers, and provided the opportunity to assist actual middle and high school students in revising their writings. This writing assistance project is in cooperation with the school district in Lyon County, Kentucky. This project was initiated because Lyon County school students performed lower than expected scores in the state writing exam, also, the school district is located in a rural place, where is no easy access to resources. Therefore, the Lyon County School initiated the project and cooperated with MSU, but because of the distance between these two schools, the project implemented only online, through a blog created by Murray State. Privacy is a primary concern so all the tutors and tutees are anonymous.

This project lasted through the spring semester 2007, in a course named Teaching Writing in Secondary Schools. The participants in this online project are different in these two courses, because of different course styles. During the fall semester 2006, there were approximately five classes from Lyon County joining this project, including two twelfth grade and three seventh grade classes. One of the twelfth grade classes is an AP class, which means these students' writing abilities are better than other twelfth graders and qualify for Advanced Placement. To implement the online tutoring session, Lyon County students followed their teachers' instruction posted writings on the blog. MSU students downloaded those students' writings, and then provided responses and questions to help them improve their works through revision.

Prior to revising on the blog, Ms. Bell informed MSU students that they must conform to the regulations in the Kentucky Writing Handbook regarding ethical responses, which mentioned that providing help in the wrong way would be against the state law.

Ms. Bell's Instruction about Responding to Students' Writings

Because the researcher had enrolled in both courses, which Ms. Bell required students to participate in the Lyon County online writing assistance project; therefore, the researcher had a thorough understanding of the content Ms. Bell had taught the participants about responding to students' writings.

Because of the different purposes of two courses, the class in the 2006 fall semester focused on the broad aspect of teaching English language art, and teaching writing is just one part of it. Therefore, in that course, Ms. Bell only introduced students to the regulations in Kentucky Writing Handbook about the ethical rules of providing responses. While, the course in 2007 spring semester is completely focus on teaching writing in secondary schools, so Ms. Bell taught students lots of notions of giving feedback before and during the Lyon County project proceeded.

The guidelines that Ms. Bell provided to students are, first, do not overwhelm students with all errors and mistakes being pointed out. Second, focus on the content of the writing, because revising does not equal to editing. Ms. Bell specifically pointed out the difference between revising and editing, and used plenty of examples to help student understand them. Third, try to point out the good part in students' writings. Fourth, though teachers should focus on the content while revising; however, when students' writing is unable to be understood for surface problems, in this situation teachers should deal with that first.

In terms of methods, Ms. Bell taught students, first, select only a few problems in a

paper that need to be dealt with first. Second, not to revise the whole paper at once, dealt with the first one or two paragraphs and ask students to look for the same patterns in the following parts. Third, because Kentucky state law restricts teachers' responding method, which are ask questions or give tips; however, sometimes these methods are not so clear to students, and some techniques can be used to clarify tutors' meaning, such as circle or underline the part they wish students to focus on. Most responding methods are written in the Kentucky Writing Handbook, and Ms. Bell provided some other thoughts. All of the guidelines and methods the researcher mentioned obeyed the state law in Kentucky.

Kentucky Writing Handbook

Kentucky Writing Handbook is published by Kentucky Department of Education (2006), which is used as the guideline for writing teachers in Kentucky to help students develop as proficient writers and learners, and to teach in the way that conforms to the state law. In Kentucky, the writing assessment is applied to the elementary, middle, and secondary school students. One of its features is it includes the writing portfolio assessment. According to the information from participants of this research (Personal Conference), except the specific grade levels are required to make portfolios, other grade levels have to do the same thing to get used to the procedure of making portfolios, which shows that the writing instruction in Kentucky emphasizes the writing process.

The content of the Kentucky Writing Handbook is divided into two parts, one is writing development and the other is scoring, the latter part is not the concern of the project. In case of violating the state law, all of the tutors are required to know how to respond to students' writing ethically. Therefore, the researcher would introduce the first part of the handbook, especially on the code of ethics. There are two purposes of responding to students' writing in the handbook, first, helping students to become independent thinkers and writers. Second, teaching the writer to internalize the concepts of audience awareness, focused purpose, idea development, and/or organization. The handbook specifically points out the revision means add, delete, or rearrange content. It does not mean to correct grammatical errors.

As for the code of ethics, the handbook specifically indicates that teachers can only ask questions to clarify student's purpose or meaning, and they may indicate the position of error and give hints for student to correct it. Another regulation is no one other than students shall make direct corrections or revisions on student's work. The handbook provides almost all kinds of situations that may happen while responding to students' writing, and provide the sample questions and methods to solve the problems. The researcher thinks this handbook would be very useful for the participants in this research, because they are still under teacher training, and the handbook clearly points what to do and what not to.

Research Question

Recently, students' writing ability is declining and this situation has caught educators' attention, and they tried to solve the problem from writing instruction. In this research, the Lyon County Middle School faced the problem of students' performance on writing, and MSU students had the chance to do this relatively new online assistance writing project, which the use of Internet is still controversial in language teaching field. Besides, all of the MSU students are still under teacher training, and their performance would be worth observing for educators, because they can display the impact of teacher training. Therefore, the research questions the researcher wanted to discuss are:

What kind of the responses will pre-service English teachers make and focus on in the online circumstance?

How do they determine which aspect of the paper are more important for students to deal with while they are revising?

How does teacher training impact on teaching practice?

Purpose

According to the Report of the National Commission on Writing (2003), more than 50% students' writing abilities are in the basic level, which means these students can only produce the writing merely to be understood. However, the basic writing ability is not sufficient for achieving in higher education or in the workplace nowadays, and this is a serious problem. One of the solutions has been proposed is to modify writing instruction (Report of the National Commission on Writing, 2003). As Stallard (1974) mentioned a good writer would do more revising, but in real life, teachers hardly have time to assist students to revise individually. With the progress of technology, distance learning becomes one of solutions, although this is still quite new in language teaching area. But the new online project provides students the opportunity to get the help from professionals, and the researcher would like to know how it would affect tutors' behavior. Therefore, the purpose of this research is to discover in the online context what kind of suggestions pre-service teachers would make on students' writings and their perception of teaching revision. The findings could indicate the efficacy of teacher training program, and further reveal the aspects the educators should focused on in training teachers for promoting students' writing ability.

LITERATURE REVIEW

Writing, considered as the most sophisticated form of transmitting information and expressing oneself, allows people to look more in depth into their thoughts and refine the information which being conveyed. Even in our modern society, when there are so many different ways of communication, writing ability is still required. With the invention of new technology, writing has changed. Teachers find that students can not compose a good writing anymore, due to their involvement of chat rooms and MSN messengers; furthermore, they also worried about students' attitude toward writing.

In the report of the National Commission on Writing (2003) suggested that to improve students' writing ability one of the approaches is started from reforming writing instruction. To look back on the history of writing instruction during the twentieth century, the major change is the instruction focus shifted from the product to process (Smith, 2000). The features of traditional writing approach, which stress on the correct usage and mechanics, looking at human discourse problems, emphasizing on parts to whole learning, and asserting language is learned through skills (Applebee, 1986 & Varble, 1990 & Minot, 1994). Traditional writing instruction concerns only about whether the final products students make reach the standard that teachers set up, while the writing process had been neglected.

In around 80s, the writing process had become popular in teaching writing. Osterhaus and Milner (1997) mentioned that educators thought a process-oriented approach would be more beneficial to students than a product-oriented approach, for the reason that it provides students the opportunity to develop their ideas and revise their works. Minot (1994) also mentioned that the earlier scholars such as Janet Emig had done the research to investigate students' actual writing process, and showed that students can learn how to write well through writing step by step. Nowadays, there are more and more teachers and educators who believe that the process-oriented instruction can facilitate students' writing abilities. The features of process writing approach, as Graves (1977) pointed out are to help students develop their authority as a writer, and teacher is to assist them to discover the knowledge on their own. White and Arndt (1991) indicated that the biggest difference on evaluation is a piece of writing would be evaluated during the process more than once.

Writing is never a simple task that students just write down whatever they think and hand it over. Writing is much more complicated than that; it is a recursive process, which includes prewriting, drafting, revising, editing, and publishing. As Donald Murray (1989) stated that a piece of writing is never finished, but for students, the last sentence to be written

down is the end without any thought for revision. Stallard's (1974) finding indicated that good writers make more revisions, which explains why students do not score well in writing, and also shows that revising is a crucial step to make a work better. Harper (1997) and Yoder (1992) both described that students do not know how to revise or misunderstand it. The reason for this situation is teachers often assume that students know how to revisit a writing critically. Harper (1997) and Yoder (1992) also mentioned that students are commonly mistaken revising for proofreading. Yet, even if teachers wanted to teach revision strategies, there is not enough instruction time in real teaching circumstances.

Tompkins (2004) mentioned that the purpose of revision is to let students clarify and refine ideas in their writings. Apparently, revising is not equal to proofreading. Except for the misconception of revising makes students could not succeed in doing revision, another reason for students' performance is that revising is a complex task itself. Hillocks Jr. (1986) proposed the notion "revision is a process" and concluded that there are many different levels of revision, from simply correcting the misspelling to changing the information. Graves (1983) inferred that children's development in revision moves from the cosmetic to information level, and Sommers (1979) mentioned that experienced writers do more in-depth revision as well. The aforementioned findings (Hillocks, 1986 & Graves, 1983 & Sommers, 1979) sounded like integrating revising and editing; however, writing process is not linear but can be proceed in anyway as long as it suits for writers' needs (Tompkins, 2004).

According to Tompkins' (2004) the definition of the word "revise" is "seeing again". Tompkins (2004) indicated that writers have to treat themselves as a reader to see their work from another view point, this revision skill sounds very simple but many researchers found that when implementing, it is very difficult. But Scardamalia and Bereiter (1983) found that after training, students can use better revision strategies, which proved that revision skills can be promoted by teaching.

To assist students in revising their writings, Sowers (1982) believed that in responding to students' writing during conference, the friendly atmosphere and asking the right questions is very important. There are different ways to respond to students' writings proposed by other researchers (Harper, 1997 & O'Shaughnessy, 2001 & Sowers, 1982); one focused on the content to elicit information from writers, another balanced the portion of the content and surface parts, and still another let students decide what to respond, teacher's intervention only when needed. To motivate students to revise their writings, Brady (1998) pointed out that teachers should encourage and give students the reasons for doing more revisions; furthermore, teachers should set the priorities when commenting and only address a few items to revise at a time. Based on the experiments that former scholars did, the researcher found that teaching revision pertains more to the response content rather than the teaching mode, and to provide appropriate feedback on writings takes time to practice for both teachers and students.

Recently, with the availability of the computer and Internet, the way people write and communicate has changed. Due to the convenience of the computer, people use word processing programs to compose instead of paper and pencil. This phenomenon has had great impacts on traditional writing. Ferris (2002) concluded four major influences on writing which included: literacy to orality, writing become more like conversation; linearity to connectivity, the logic of the content become loose; passivity to interactivity, readers have control of the information they received; and traditional quality to value, the value of a Internet published work depends on how many people interested in it rather than how well it is composed. In another piece of research, Warschauer (2004) pointed out that technology changed the purposes of writing, genres of written communication, and nature of audience and author.

Because of the wide use of new technology in public daily lives and its great impacts on

writing, the educators had tried to incorporate the technology with language instruction. In the field of teaching writing, there are two new concepts provide teachers as well as students new experiences of teaching and learning to write. One is computer-mediated communication (CMC), which is using computer and the Internet as media to carry out the communication among teachers and students or students and students. The other concept is computer assisted language learning (CALL), which includes broad uses of the computer that range from simply using the computer to compose a work to using computer software to improve writing ability (Anson, 1999). In recent years, it is more and more common to apply CMC and CALL into writing instruction, but as other new experiments, there are both positive and negative opinions of their efficacy in the writing instruction.

Owing to the obvious changes in writing, writing teachers worry about the influences of technology on students' writing ability, especially the "chat room" style writing. Abdullah (2003) indicated that students are more accustomed to write in the conversational style, and their productions illustrated poorly structured and articulated skill, and even accompanied with additive writing style, which proved teachers' concerns. As for the conveniences technology provides, for instance, the auto-correction function in word processor, Abdullah (2003) found students would over rely on auto-correction function, which results in the unwillingness of revisiting the words. Anson (1999) also doubted the necessity of using a computer in composing and using software to learn, because in the end, the outcome still emerged on paper, and students still need to attend to class.

However, Anson (1999) also indicated that the cut-and-paste function of the word processors enable students to manipulate texts in an easier way, therefore, increasing their motivation to write. Grejda and Hannafin (1992) and Cunningham (2000) found similar results that word processors can actually assist students with revision, and improve students' attitude toward writing. Unfortunately, the results showed that these revisions do not influence the overall quality of students' writings, because they only add more texts. The word processor is simply a tool that saves time on recopying the text.

As for the use of CMC, in terms of writing instruction, the researches showed that it is quite beneficial for both students and teachers. It promotes students' cooperative relationship and more complicated language use, increases students' interaction with others and the motivation to learn, and it reduces the communication anxieties as well (Warschauer, 2004). The features of CMC allow people to get the information at their disposal and have more opportunities to gain the help they need. (Riding and Rayner, 1995 & Coogan, 1994 & Shedletsky, 1996).

Because of the accessibility technology provides, distance learning has become possible and benefit both students and teachers. Students obtain all kinds of resources they need worldwide, and teachers can exchange their ideas with other teachers from all over the world (Shedletsky, 1996; Riding and Rayner, 1995; Anson, 1999). Technology has opened doors for writing instruction, which brings about both positive and negative consequences. How to use it wisely would be a challenge for both students and teachers.

METHODOLOGY

This research is a case study, since it focused on a specific group of people under a certain circumstance (Rossman, 1998), therefore, to understand pre-service teachers' teaching strategies and perceptions, the researcher used case study for close examination. The methods the researcher used include interviews and content analysis. The researcher chose to interview with tutors, because there are some information that is hard to tell from their performance. Besides, through interviewing with tutors they may disclose other valuable information which can assist the researcher to answer the research questions. The researcher used content analysis is trying to interpret the tutors' teaching practice precisely, because the

data showed what they did. Moreover, there are many aspects in a piece of writing that tutors can respond to, and the content analysis can help the researcher to categorize tutors' responses more accurately to find out what aspect they focus on. With the combination of the results with the findings in the interview, the researcher can understand how the teacher training impact on tutors' practice. Though the researcher could not interview with the students who joined the online project in the fall semester 2006, through analyzing the content of revisions collected from blog, the results still provide the information about tutors' focus while responding to students' writings.

In this research, the interview conducted through semi-structured questions, with the questions being prepared in advance by the researcher to inquire about their criteria used in deciding how to prioritize their responses and their perception of the revision process, but they are not fixed. All of the interviews would be implemented face-to-face and the conversations were recorded. The willingness of participation would be determined through e-mails. The interviewees were informed in advance about the interview procedure, and also notified that the conversation would be recorded. The content analysis referred to the collection of responses that MSU students posted on the blog, and the researcher would analyze the content of responses by categorizing them into different revision aspects. The criteria of categorization is based on the other scholars' researches (Hillocks, 1986 & Graves, 1983), Ms. Bell's instruction, and actual situation, therefore, the categories would be clear to show tutors' focus. After categorizing, the researcher would count occurrences of response to find out the major types tutors made. Thereafter, the researcher would combine the results and analyze the interview, to discover their reasons of choosing certain type of responses.

Participants

In this research, the participants are MSU students, who are mostly English major students and future English teachers. The students are enrolled in Ms. Bell's Teaching Writing in Secondary Schools class, and the total enrollment is 11 students including the researcher. Every participant would participate in the online project in 2007 spring semester. In this class, some of the students had participated in the online project in 2006 fall semester; the portion of the experienced students to new students would be approximately fifty-fifty. The experienced student means students who had participated in the online project in 2006 fall semester.

Procedure

In the beginning of the research, the researcher collected all of the responses MSU students posted on the Lyon County blog in fall semester 2006. This data enriches the information about types of responses that pre-service teachers would make. After the online project started again, the researcher collected all of the responses MSU students posted on the blog during spring semester 2007. Although Ms. Bell informed tutors that tutees may repost their writings on the blog, as the project went through, none of the Lyon County students has done that. Therefore, the researcher did not have the chance to see how MSU students respond to those revised writings. The overall data collection finished at the end of the spring semester 2007.

Interviews were conducted when every MSU student posted their responses for Lyon County students' writings in spring 2007. Before the interview, the researcher asked every MSU student for permission to schedule interview; either through sending e-mails or asking in person. In addition, the researcher informed every participant that the interview would be recorded by MP3 player, and the content would be confidential, only used for the research paper. Also, if there is the need for indication of the interviewees, the names would be pseudo names to protect their privacy.

After collecting the data from the blog and interviews, the researcher examined the responses MSU students made on Lyon County students' writings and categorized them into

different revision aspects to see tutors' focus in responding. The researcher set the criteria according to the researches done by Hillocks (1986) and Graves (1983), since the situation of Lyon County project is very different from theirs, the researcher thought it is necessary to modify the classifications.

Because of the state law in Kentucky (Kentucky Department of Education, 2006), all of the writing teachers are restricted to give responses in the form of questions or hints. Therefore, the researcher used the broad categories to classify responses according to the formation, which are question and statement. The statement in the research means any ways to provide students hints or suggestions to make them notice the parts that need to be improved. In addition, based on Harper (1997) mentioned that students' problem of misunderstanding the meaning of revising and editing, therefore, to see MSU students' focus in their revisions, the researcher subdivided under the two main categories into form and meaning to show their focus.

To be more specific on the types of problems that pre-service English teachers would focus on, the researcher listed the areas that MSU students had mentioned in their revisions. The researcher decided the title of each area in accordance with the actual situation, and the following are explanations for the areas that may cause confusions. First, the lists under the Form, the word choice, which with broader explanation that it includes choose the more specific words to express oneself and the right way to write the word. Sentence structure includes all the problems on sentence level. While citation means students forget to cite their information.

Second, the lists under the Meaning, the information means that tutors ask for more details about facts. The explanation means tutors ask for more details about writers' perception or interpretation of the things/issues mentioned in the writing. The source means tutors ask students to provide the information or examples to support their own arguments. The idea development means the tutors try to elicit new direction of thinking for students. The selection of title means students' title and content are not consistent. The repeating is the condition which students repeat or rephrase the same thing throughout the writing. Last, there are some areas which do not belong to either form or meaning, and that would be article construction and Compliment. As for article construction, the researcher gave it a broad meaning; it includes the flow and logic of the article, and other problems bigger than sentence level. The compliment indicates any sentences that tutors wrote for encouraging students or pointing out good parts in the writings. The following table 1 shows the detailed categories with examples for each one.

Table 1. *The categories and examples of responses.*

Question			
Form		Meaning	
Article Format	Do you need a title?	Information	Where is it?
Word Choice	Should you spell out the number?	Explanation	Why do you think so?
Contraction	Should you use contractions in formal writing?	Source	Where does this information come from?
Sentence structure	Could you write this sentence without the word "there"?	Idea development	Is this relevant to your piece?
Grammar	Should this be a question or a statement?	Selection of title	Is this title attention grabbing?
Spelling	Are you sure about the	Repeating	Did you say that

	spelling?		before?
Capitalization	Why is IN capitalized?	Voice	Could you use active voice here?
Punctuation	Could you use a colon or semicolon here?		
Citation	Can you tell me how you came up with 92%?		
Transition	Can you make a transition between these two sentences?		
Article Construction		Could this paragraph go some place else?	
Statement			
Form		Meaning	
Article Format	I think you need an introduction paragraph.	Information	I'd like to know more about the place.
Word Choice	Watch your choice of language.	Explanation	I don't understand what you mean.
Contraction	Contraction.	Source	It may be necessary to briefly state your source and page number.
Sentence structure	This sounds redundant, try to present your statistic efficiently.	Idea development	Think about what students' duties are.
Grammar	Consider grammar here.	Selection of title	I think you can write a better title.
Spelling	Rethink the spelling here.	Repeating	Yes, but you have already said that.
Capitalization	Check capitalization.	Voice	Try to use active voice.
Citation	I think you need to cite the information.		
Punctuation	Think about a possible punctuation here.		
Transition	Try to make a connection with the former paragraph.		
Article Construction		Try to differentiate them or combine them.	
Compliment		Great thesis!	

Note: The examples listed in the table are randomly choose from the data or invented by the researcher which may not exist in the data.

The researcher categorized all of the responses into specific categories. By discovering the proportion of main categories, the researcher obtained the answer of types and focus of responses that pre-service English teachers would make in the online circumstance.

In the second part of the content analysis, the researcher summarized all of the interview content and transcribed some examples of interviewees' replies; the example was selected according to its usefulness of answering the research questions. The interview contents were

summarized into different classifications mostly according to the questions the researcher designed, such as perception of teaching revision. Because the interview questions are not fixed, there are some other classifications the researcher created while summarizing the recordings, for instance, process of giving feedback. Some of the classifications only showed up in few interviewees. When finishing summarizing, the researcher could understand pre-service English teachers' standards and reasons for determining their priorities in giving feedback to students' writings.

The last part of analysis the researcher compared the results of the responses tutors made, which represent the real situation while pre-service teachers teaching, to the content of the interview, which shows the participants' explanations for their performance, to see whether the teacher training influence tutors' teaching practice or there are some other reasons.

FINDINGS

To investigate pre-service English teachers' performance in giving responses to students' writings, their perception of revising, and the impact of the teacher training on the teaching practice, the researcher collected the revisions MSU students posted from the blog, and interviewed with six voluntary participants. To understand the information the researcher collected so far, first, the researcher dealt with the revisions collected from fall 2006 to spring 2007. The researcher categorized all the responses into different categories, to show the proportion of each part which assisted interpreting the results. Second, the researcher would summarize and transcribe the interview recordings from the spring 2007. While summarizing the content, the researcher would base on the interviewees' responses, classified them into different parts, and these parts would be the elements to answer the research questions.

Revisions

In the first part of findings, the researcher collected the revisions MSU students posted on the Lyon County blog during the 2007 spring semester. Because the revisions that MSU students posted during the 2006 fall semester remained on the blog at the time the researcher started collecting data, these revisions are also collected to provide larger sample of data to see pre-service English teachers' performance in giving responses.

The total revisions the researcher collected are 67, and there are 38 revisions from 2006 fall semester, and the other 29 revisions from the 2007 spring semester. In each revision, there are many responses, the researcher counted each response by the sentence tutors wrote, which means each sentence represents each response. The total responses are 928, and the average responses of each revision are about 13.85.

The finding showed that there are 64% (591) of responses are provided in question form and 36% (337) are in the statement form, which indicated that MSU students slightly preferred to give feedback by asking questions. In terms of the categories of focus, the finding showed that there are 57% (527) of responses on the form and 33% (303) on the meaning, but there is an uncategorized part, the researcher called it the overlap, because the features of this area of response that would overlap in both areas, and there are 10% (97) of responses in this classification. The findings showed that MSU students gave more feedback in terms of form, however, among all kinds of areas the tutors focus on eliciting more information from students, while looking at the proportion in each semester, the results remain the same, and this show that MSU students consistently put their focus on eliciting information. There are two areas of focus having the highest portion in 2007, one is information (13%), the other is compliment (13%), when compare to the results in 2006, the compliment (4%) obviously increased, which display the different courses focus. As for comparing the main focus of each semester, although form (60%, 2006; 52%, 2007) is the main focus in both semester, but in 2007, the portion has been decreased.

Interviews

In order to understand tutors' criteria of determining the priorities in responding to students' works, and influences of teacher training on tutors' performances, the researcher had interviewed six out of eleven tutors. To obtain the answer of the tutors' standards of deciding the priorities in giving feedback, the researcher classified the answers into three parts, which are the perception of teaching revision, the focus of giving response, and the purpose of giving response. As for the answer of influences of teacher training on tutor's performances, the researcher found that it is difficult to get the answer from certain question, because it required seeing tutors' perception and their actual behavior, and it is hard to pin down any questions for that. Therefore, the researcher did not set a certain section here, but it will be discussed in the conclusion.

Perception of teaching revision Based on the recordings, the researcher found that every interviewee thought revision is very important and necessary in writing instruction, but their reasons are various. The researcher categorized them into four reasons, which are revision can change the form of writing, revision can improve students' ideas, revision can let writers gain the response from readers, and revision provides students opportunities to practice writing.

During the interview, the researcher had asked interviewees their philosophy of teaching revision, and the researcher obtained a similar answer from all of the interviewees that is no one has a certain guideline to teach revision, as the interviewee C said, "It's really hard to have certain philosophy, because I think as a teacher you have to be a flexible person, because every student, every situation is different." While the researcher asked this question, some of the interviewees stated how they would teach their students, and the researcher found two methods are pretty useful, that are do not do revision everyday and revise from part to part. The researcher also found that these two methods match the suggestions in the Kentucky Writing Handbook, but from the recordings when interviewees talked about their teaching strategies, they sounded like they came up with ideas from their own beliefs, besides, there are so many examples in the Kentucky Writing Handbook. The researcher thought it is pretty interesting to know that tutors' perceptions match to the regulation.

Focus of giving response The researcher found that four out of six interviewees said they would focus on the meaning of the writing first; whereas, the other two interviewees said they would focus on the form first. In addition to simply listening to interviewees' responses, the researcher also looked at the revisions that interviewees did, to see whether the interviewees acted as they said. Among the four interviewees A, B, D, and F, who said they would focus on the meaning of the paper, the researcher found that from interviewee F's revision, which seemed like it focused on the form. The researcher has inquired interviewee F for reasons, and the explanation is that the content of that writing must be pretty good, so she gave the further suggestions to the next step which is starting to editing the writing. Because of all of the revisions the researcher had downloaded from the blog, there is only one revision that the interviewee F did; therefore, it is hard for to tell what the interviewee F would really do. The interviewee B, he didn't post any revisions on the Internet, for the reason that he did not know how to post them on the blog. The researcher still interviewed him to gain more points of view from pre-service English teachers, and by revising student's writing during the interview, the researcher was still be able to know how the interviewee responded to Lyon County students.

As for interviewees C and E, both of them mentioned that they focus on the form of a writing first, and they have a similar explanation, for instance, the excerpt from interviewee E, "The first I will hit like the grammar errors, misspelling, punctuations, wrong words like that one, because I think it is really hard to actually go in and work on the content, when there are so many surface errors". Moreover, the researcher found that both of them did not deny that

the content of the writing is the most important. Because it is harder to deal with content problems, they thought students should fix surface problems first, and the perception to solve the problems from easy to hard, which responds to Graves' (1983) research finding that the development of revision is from the surface level to the text level.

Purpose of giving responses To fully understand how the interviewees set a certain area as the priority in doing revision, the researcher inquired the reasons of giving certain responses. Based on the interview recordings, the researcher discovered four different purposes, one is about learning method, and the other three are concerning writing itself. The interviewee B thought the purpose of giving responses is to assist students to learn how to write well through hands-on experiences, therefore, he always provides a lot of suggestions for students to work on. The interviewee A thought her purpose of giving response is to try to understand students' writings, and then she would know how to help them to improve. The interviewees C, E, and F thought their purpose is to make the writing easier to read, but here are two approaches to achieve this goal. One is to fix surface problems make the writing easier to read in terms of the form, and the other is to fix content problems make the concept clear, so the writing would be easier to be understood. The Interviewee D thought her purpose is to encourage students and make them think more in-depth of what they write to develop their ideas.

CONCLUSION

Recently, the phenomenon of students' writing ability deteriorating caught the attention from educators as well as business leaders. Since students' writing ability is not sufficient for them to achieve in higher education and work places, to improve students' writing skill, the educators started reforms from many aspects, such as, writing instruction and assessment (Report of the National Commission on Writing, 2003). For example, the Scholastic Assessment Tests (SAT) has changed the test content to add the writing section in it, to cater to the college professors' concerns.

This decline of students' writing ability may result from many reasons, and one of them is the use of technology, especially the Internet. Although the use of the Internet in language teaching area is controversial, it is impossible to neglect it, since the modern public cannot live without the Internet and computers. In this research, the online writing assistance writing project the MSU students did is using the Internet to do distance teaching and learning, and this kind of utilization of the Internet in writing instruction is still quite new, so it is hard to know for sure its impacts on the participants. Furthermore, the tutors in this project are still under teacher training, and their performance on responding to students' writings would be worth observing to see the efficacy of teacher training, and from the feedbacks tutors made, the researcher can understand what they perceive to be important while teaching revision.

According to the content analysis, the researcher found that tutors focused on the form while they giving feedback on students' writings. However, the findings of the interviews tells a different story: it showed that every tutor considered the content of writing as the most important. The explanations for this discrepancy may be because of MSU students' previous learning experiences that made them unconsciously focus on the form, or as they said in the interviews that based on students' ability, they thought problems on the form should be dealt with first. This finding also illustrated that it is hard for tutors to act as they think. To further investigate, the researcher found that both semesters tutors focused on the form. But when the researcher looked more closely into the result, the proportion of the form decreased and the overlap part increased in 2007 spring semester. The most prominent increment is the compliment, which clearly revealed the influence of Ms. Bell's instruction, because one of the emphatic points in 2007 course is to find the good part in students' writing and encourage them. Based on the findings of the interviews, the researcher concluded that tutors' criteria of

determining the priority of giving responses would be students' writing skill level, and because this would be vary according to students, therefore, there is no fix responses that tutors would make to each writings. Moreover, tutors' answers about their perception, focus, and purpose of doing revision, showed that most of their answers match to the instruction they received in class.

During the interviews, the researcher found there are factors other than teacher training would influence pre-service English teacher on making instructional decision. The researcher discovered two main factors, one is tutors' personal experiences, and the other is their living environments. First, tutors' personal experiences. Based on interviewee E told the researcher, she has got used to be corrected every single problem in her writing since she was a high school student, therefore, it is okay for her to get the paper with all marks in red, and she did not feel her self-esteem has been hurt, which is different from scholars' assumption. Hence, this is the reason why she thought correct the surface problems for students is acceptable. As for the influence of living environment, the interviewee C told the researcher that, in the city she lives there are many students who have very limited English proficiency in high school she observed. Because of the environment the interviewee C lives, which makes her set up her top priority on the form, because in the circumstance she described the understandability indeed is the most important thing for students to achieve in school.

This research showed that the pre-service teachers still focus on the form while responding to students' writings, which may explain why students' writing ability would not get better because they are still instructed in the way many scholars thought as inappropriate for developing a proficient writer. However, this doesn't mean the teacher training has no influence on pre-service teachers, as the finding illustrated most of tutors believe what they have been taught. Although teacher training provides all the principles and knowledge to tutors, there might be some conflicts that occur with tutors' personal beliefs. During the interviews, tutors clearly showed that they are not simply accept whatever the instructor taught them, they also develop their own teaching philosophies with the effects of other factors. Yet, these tutors did not always practice as they think, and this may because of their previous experiences, and old habit is hard to change. Therefore, to reform the writing instruction, the researcher thinks to assist pre-service teachers become more aware of what they do would be helpful for them in the teaching practice.

While conducting the interview, the researcher had asked interviewees a question about using the Internet and computer in this project, and surprisingly most of the tutors are not in favor of this online project. They prefer to have the chance to meet face-to-face with students. Moreover, tutors think the interaction with writers is very important so that they could have a better idea of how to help them. Furthermore, some MSU students also indicated that the convenience of technology is not necessary a good thing, which the researcher thinks is quite interesting. This research found both positive and negative results of using technology in writing instruction and the difference is on the usage, therefore, the researcher thinks there is no right or wrong to use technology in teaching, as long as the users can find the most beneficial way for themselves. Teaching writing is never an easy job for an English teacher since every student needs different kinds of help, but with the use of technology, the researcher thinks it may provide a chance for every student to get the assistance they need in achieving in writing.

Limitations of this Research

While doing this research, there are a few limitations, first, the defects in the on-line project; there is no good coordination between Lyon County Middle School and MSU, which led to the bad operating procedure and result in some tutors never receiving the writings from tutees. Moreover, tutees never reposted their revised writings, so tutors have no chance to find out if their advice is helpful. Second, is the problem with the technology. Because this is

a new project for both schools, and neither sides are familiar with the use of the blog, some tutors and tutees could not post their writings, therefore, no one benefits from the project. Third, participation was an issue. Because this on-line project is not compulsory for students in Lyon County Middle School, MSU students did not know the exact number of students, which causes some tutors not to find any writings from their tutees. Fourth, there is no control of students' writing abilities, so their proficiency levels might be very different. Last, the inexperience of conducting interview and lack of the interview skills was problematic because this is the first time the researcher has done such interviews. During the process the researcher was unable to keep interviewees talking, therefore, there might be some information that was not acquired from interviewees. The researcher thought after overcoming these limitations, the findings might be more accurate and helpful.

REFERENCES

- Abdullah, M. H. (2003). *The Impact of Electronic Communication on Writing*. Retrieved June 12, 2007, from http://www.eric.ed.gov/ERICDocs/data/ericdocs2/content_storage_01/0000000b/80/2a/39/48.pdf National Society for the Study of Education.
- Anson, C. M. (1999, January). Distant Voices: Teaching and Writing in a Culture of Technology. *College English*, 61 (3), 261-279.
- Applebee, A. N. (1986). Problems in process approaches: Toward a reconceptualization of process instruction. In A. Petrosky & D. Bartholomae (Eds.), *The teaching of writing: Eighty-fifth yearbook of the National Society for the Study of Education, Part II*.
- Brady, L. (1998). *Responding to Writing*. Retrieved January 18, 2007, from <http://www.as.wvu.edu/%7Elbrady/response.html>
- Cunningham, K. (2000, May). *Integrating CALL into the Writing Curriculum*. Retrieved February 10, 2007, from <http://iteslj.org/Articles/Cunningham-CALLWriting/>
- Coogan, D. (1994, March). *Towards a Rhetoric of On-line Tutoring*. Paper presented at the Annual Meeting of the Conference on College Composition and Communication, Nashville, TN.
- Ferris, S. P. (2002, August). *The Effects of Computers on Traditional Writing*. Retrieved February 10, 2007, from <http://www.press.umich.edu/jep/08-01/ferris.html>
- Graves, D. H. (1983). *Writing: Teachers and children at work*. Portsmouth, NH: Heinemann.
- Graves, D. H. (1977, April). Language Arts Textbooks: A Writing Process Evaluation. *Language Arts*, 54, 817-823.
- Grejda, G. F. & Hannafin, M. J. (1992). Effects of Word Processing on Sixth Graders' Holistic Writing and Revisions. *Journal of Educational Research*, 85 (3), 144-149.
- Harper, L. (1997, March). The Writer's Toolbox: Five tools for Active Revision Instruction. *Language Arts*, 74, 215-222.
- Hillocks, G. Jr. (1986). *Research on Written Composition: New Directions for Teaching*. Illinois: ERIC Clearinghouse on Reading and Communication Skills.
- Ioannou-Georgiou, S. (2006, October). Online Forum Report: The future of CALL. *ELT Journal*, 60 (4), 382-384.
- Kentucky Department of Education: Kentucky Writing Handbook. (n.d.). Retrieved June 12, 2007, from <http://education.ky.gov/KDE/Instructional+Resources/High+School/English+Language+Arts/Writing/Kentucky+Writing+Handbook.htm>
- Minot, W. S. (1994). *Composition and Rhetoric: A Natural Alliance?* Paper presented at the Annual Meeting of the National Council of Teachers of English Conference on College Composition and Communication, Nashville, TN.
- Murray, D. M. (1989). The maker's eye, revising your own manuscripts. In S. Gillespie, R.

- Singleton, & R. Becker (Eds.), *The Writer's Craft*, (pp. 88-91). Glenview, IL: Scott, Foresman and Company.
- O'Shaughnessy, K. (2001). *Everything I know About Teaching Language Arts I Learned at the Office Supply Store*. Retrieved January 18, 2007, from <http://www.writingproject.org/cs/nwpp/lpt/nwpr/147>
- Osterhaus, K. & Milner, J. (1997, December). *Writing Instruction Approaches in the High School English Classroom and Students' Writing*. Paper presented at the meeting of the Annual Research Forum, Winston-Salem, NC.
- Riding, R. J. & Rayner, S. (1995, December). The Information Superhighway and Individualized Learning. *Educational Psychology*, 15 (4), 365-378.
- Rossman, G. B. & Rallis, S. F. (1998). *Learning in the field*. Thousand Oaks, CA: SAGE.
- Scardamalia, M. & Bereiter, C. (1983). The development of Evaluative, Diagnostic, and Remedial Capabilities in Children's Composing. In M. Martlew (Ed.), *The Psychology of Written Language: A Developmental Approach*, (pp. 67-95). London: John Wiley & Sons.
- Shedletsky, L. (1996). *Teaching with Computer-Mediated Communication*. Retrieved February 10, 2007, from http://www.eric.ed.gov/ERICDocs/data/ericdocs2/content_storage_01/0000000b/80/10/c8/2e.pdf
- Smith, C. B. (2000). *Writing Instruction: Changing Views over the Years*. Retrieved June 12, 2007, from <http://www.vtaide.com/png/ERIC/Writing-Instruction-Views.htm>
- Sowers, S. (1982). Reflect, Expand, Select: Three Responses in the Writing Conference. In T. Newkirk & N. Atwell (Eds.), *Understanding Writing: Ways of Observing, Learning & Teaching* (pp. 76-89). Chelmsford, MA: The Northwest Regional Exchange.
- Sommers, N. I. (1979). Revision in the Composing Process: A Case Study of College Freshmen and Experienced Adult Writers. *DAI*, 39.
- Stallard, C. K., Jr. (1974). An analysis of the Writing Behavior of Good Student Writers. *RTE*, 8, 206-218.
- Tompkins, G. E. (2004). *Teaching Writing: Balancing Process and Product*. (4th.). New Jersey: Pearson Education, Inc.
- Varble, M. E. (1990). Analysis of Writing Samples of Students Taught by Teachers Using Whole Language and Traditional Approaches. *Journal of Educational Research*, 83 (5), 254-251.
- Warschauer, M. (2004). Technology and writing. In C. Davison and J. Cummins (Eds.), *Handbook of English Language Teaching*. Kluwer: Dordrecht, Netherlands.
- White, R. & Arndt, V. (1991). *Process Writing*. (1st.). UK: Longman Group UK Limited.
- Yoder, S. L. (1993). Teaching Writing Revision: Attitudes and Copy Changes. *Journalism Educator*, 47 (4), 41-47.
- The National Commission on Writing: The Neglect "R": The Need for a Writing Revolution*. (n.d.). Retrieved June 12, 2007, from http://www.writingcommission.org/prod_downloads/writingcom/neglectedr.pdf

Titles of Research Articles: Syntactic Structures and Communicative Effectiveness

Chih-wei Kuo(郭治偉), Chih-hua Kuo(郭志華)

National Chiao Tung University
Minghsin University of Science and Technology

louie.tesol96g@g2.nctu.edu.tw
chk@must.edu.tw

A research article depends much upon its title to designate its distinctiveness of content from many other studies on a similar or even the same research topic. However, little research has paid attention to the titles of research articles. In this study, we examined titles from a genre perspective; specifically, the communicative effectiveness of various types of titles was analyzed in relation to syntactic structures. Furthermore, disciplinary variation as reflected from the use of different types of titles was explored. A corpus of 454 titles in the fields of applied linguistics, industrial engineering, and sports science was compiled. The results show that nominal titles are the most frequently used type of titles across the three disciplines. In addition, it is interesting to find that the patterns of “effect of A on B,” “influence of A on B,” and “impact of A on B,” are recurrent features in titles, suggesting the research trend or focus. A title is regarded as communicatively effective when it contains the information content that is the focus of research; for example, a study emphasizing a self-designed research method designates the method in its title. On the other hand, compound titles occur most frequently in applied linguistics; further analysis shows topic-scope, topic-source, and topic-method are the most common combinations of the two parts of titles. There are some occurrences of declarative full-sentence titles in sports science, but this construction rarely occurs in the other two fields, marking disciplinary variation. The results of this exploratory study on titles clearly show the communicative value of titles.

INTRODUCTION

The importance of a research article title, though a small part of the text, cannot be overemphasized. For readers, when they attempt to review relevant literature, either through keyword searches in databases or through reference scanning in journal articles, books, or conference papers, titles are often the first contact before they really read the article. The first impression is often decisive. As Day (1998) indicated, titles are read by thousands of people, but few, if any, will read the entire paper. In addition, a successful research article is often known by its title. An effective title can designate the distinctive content of the article. For writers, therefore, titles should be constructed with great care. However, writing a good title is not as easy as expected. Novice writers need to know the requirements and structures of good titles.

In recent years, research articles (RAs) have received much attention. Studies on RAs are conducted along two major lines. One line of research has analyzed the macro-structures of various sections in research articles (e.g., Swales, 1990; Lim, 2006; Yang & Allison, 2003). The other line has examined the linguistic features of this particular genre, such as voice, tense, reporting verbs, or modals (e.g., Hawes & Thomas, 1997; Malcolm, 1987; Swales, 1990). Titles, however, have been largely neglected in the flood of RA studies. As a matter of fact, most of the studies on titles have been done by information scientists (e.g., Hartley, 2005,

2007a, b; Kilgour, 2004; Lewison & Hartley, 2005; Yitzhaki, 1994). They have focused on the length and classification of titles. In contrast, title studies from the perspectives of applied linguistics or genre have been relatively under-represented.

Academic writing style manuals or textbooks have proposed various ideas for good RA titles. For example, Swales and Feak (1994) indicated three requirements for good research paper titles: “(1) The title should indicate the topic of the study; (2) The title should indicate the scope of the study; (3) The title should be self-explanatory to readers in the chosen area” (p. 278). Davis (1997) proposed that a good title should use: “(1) the most precise words possible, (2) words that indicate the main point of the paper, and (3) words that lend themselves to indexing subject” (p. 111). Day (1998) defined a good title as “the fewest possible words that adequately describe the contents of the paper” (p. 15). Despite some minor differences among these authors, by and large, there are two characteristics of good RA titles, namely, informativity and economy (Soler, 2007).

Although the above studies have pinpointed the requirements or criteria for good titles, we do not know the structural constructions of good RA titles based on empirical analysis, in particular, titles of RAs that appear in prestigious journals. The frequently used syntactic structures of these RA titles can show the academic conventions of constructing titles. Moreover, it is of interest to find how researchers in different disciplines may construct RA titles differently as a result of the different nature of their research.

Therefore, the present study is an attempt to analyze the syntactic structures of RA titles empirically and explore disciplinary variations in the use of various types of titles as well. A corpus-based approach is taken in our empirical and computational analysis of the titles. Two research questions are addressed in this study.

1. What are the syntactic structures of RA titles?
2. Do researchers in three relatively heterogeneous disciplines, namely, applied linguistics, industrial engineering and sport science, use different constructions for their RA titles?

Structural Constructions of Titles

Analysis of the structural constructions of titles could be found in a number of recent studies (Haggan, 2004; Soler, 2007; Wang & Bai, 2007). Haggan’s (2004) analyzed 751 titles from linguistics, literature, and science, classifying the titles into three categories: full-sentence titles, compound titles, and the remaining title structures. The category of the remaining title structures encompassed three subcategories: noun phrases, prepositional phrases, and participial phrases. The results demonstrated that in literature RAs, compound titles predominate (60%), while in science and linguistics RA titles, approximately two-thirds use nominal construction. In addition, there is a much higher percentage (8.5%) of full-sentence titles in science papers than in the other two disciplines.

Soler (2007) examined titles across both disciplines and genres. Two genres — research papers (RP) and review papers (RVP) in two contrasting fields—biological sciences and social sciences were analyzed in terms of four structural constructions of titles: nominal, question, compound, and full-sentence constructions. Social sciences include three subfields, namely, linguistics, anthropology, and psychology, whereas biological sciences encompass another three subfields—biology, medicine, and biochemistry. It was found that full-sentence construction was a peculiarity of RP titles in biology, medicine, and biochemistry, but no occurrences of this construction were found in the RVP titles. In addition, compound construction occurred more frequently in RP titles than in RVP titles.

A third study by Wang and Bai (2007) focused only on one discipline. They analyzed the syntactic structures of 417 English titles in medical research papers and found that a very high percentage of nominal constructions (99%) were used in this discipline. Three types of nominal groups, including uni-head, bi-head, and multi-head nominal groups were identified.

In addition, the functions of all the nominal groups were also analyzed in terms of both pre-modifiers and post-modifiers. For instance, present participles used as post-modifiers in nominal groups aim to inform readers of the most important point of the study, such as “*IGF-I Receptor Mutations Resulting in Intrauterine and Postnatal Growth Retardation.*”

Compound Titles

Among the various types of titles that have been identified, colonic titles, or compound titles, attract our attention since they seem to be a special syntactic device meeting the criteria of both informativity and economy. Titles consisting of two parts separated by a colon are usually called colonic titles (e.g., Hartley, 2005). They are also given other names, such as hanging titles, series titles, or compound titles (Day, 1998). In this study, we use the term “compound titles” instead of “colonic titles” since punctuations other than colon can be used to form a two-part title. Compound titles appear to be a special feature of RA titles in social sciences. Dillon (1981) indicated “titular colonicity” was a prevailing feature of scholarly publication. Hartley (2007a) investigated colonic titles in seventeen disciplines and found that they are used throughout all disciplines but the percentage increases from natural sciences to social sciences. Swales and Feak (2004) indicated that one of the functions of the colon is to separate related ideas into two parts. They identified four possible combinations for colonic titles: Problem : Solution, General : Specific, Topic : Method, and Major : Minor. Adapting Swales and Feak’s categories, Anthony (2001) found other combinations in computer science RA titles, such as Name: Description, Topic : Scope, Topic : Method, Description : Name, and Topic : Description. The first three types of combinations were the top three frequently used colonic titles in computer science. As readers usually expect to know from the title not only the general topic of a study but also its specific content, which distinguishes this specific study from other studies, the two parts of a compound title provide a clever format for writers to include both in a single title. We are thus motivated to further examine possible relationships between the two parts of compound titles in our corpus.

METHOD

The Title Corpus

In this study, we first compiled a corpus consisting of 454 RA titles from three disciplines, applied linguistics, industrial engineering, and sport science. All of the titles were selected according to the following criteria. First, in each discipline, we chose two prestigious journals; both belong to Social Science Citation Index (SSCI) or Science Citation Index (SCI) journals. Second, in each journal, there may be various types of articles, such as research articles, review articles, and correspondences. We selected only titles of research articles. Since these RAs all have undergone a strict peer-reviewed and ratified process, we may assume that they tend to have effective titles, too. Table 1 shows the basic information of the title corpus, including the selected journals, publication type, time span, and number of titles collected.

Table 1. *Basic Information of Six Selected Journals across Three Disciplines*

Journal Title (Discipline)	Publication Type	Month/Year	Vol./No.	No. of titles
TESOL Quarterly (Applied Linguistics)	Quarterly	Mar. 2004 to Dec. 2007	38(1) to 41(4)	75
Modern Language Journal (Applied Linguistics)	Quarterly	Mar. 2004 to Dec. 2007	88(1) to 91(4)	75
Applied Ergonomics (Industrial Engineering)	Bimonthly	Jan. 2004 to Mar. 2005	35(1) to 36(2)	78
International Journal of Industrial Ergonomics (Industrial Engineering)	Monthly	Jan. 2004 to Jan. 2005	33(1) to 35(1)	78
Human Movement Science (Sport Sciences)	Bimonthly	Jun. 2004 to Dec. 2005	23(1) to 24(6)	75
Journal of Motor Behavior (Sport Sciences)	Quarterly Bimonthly	Mar. 2004 to Dec. 2004 Jan. 2005 to Nov. 2005	36(1) to 37(6)	73
Total				454

Data Analysis

After the compilation of the title corpus, a two-phase analysis process was implemented. In the first phase, all of the titles were analyzed according to their syntactic structures, falling into four mutually exclusive categories: (1) nominal construction in which a noun is used as a head, which can be preceded by pre-modifiers or followed by post-modifiers (e.g., *Hemispheric Competition in Left-Handers on Bimanual Reaction Time Tasks*); (2) compound construction in which two parts are linked together by a colon, dash, question mark, or period (e.g., *Pointing Out Frequent Phrasal Verbs: A Corpus-Based Analysis*); (3) full-sentence construction, which is a subject-verb-object structure, including both declarative structures (e.g., *Dual-Task Training Reduces Impact of Cognitive Task on Postural Sway*) and interrogative structures (e.g., *Is there an Academic Vocabulary?*); (4) other constructions, which refer to constructions that do not belong to the previous three categories, such as prepositional phrase structures (e.g., *Toward an Understanding of the Role of Applied Linguists in Foreign Language Departments*), gerund structures (e.g., *Evaluating the coordination dynamics of handwriting*), and others (e.g., *Grip force when grasping moving cylinders*). As the first three categories constitute the majority of titles (94%), we focus on the three main categories in the next analysis.

The second phase was a semantic analysis of the compound titles in the discipline of applied linguistics. Both researchers (a professor specializing in EAP research and a graduate student majoring in TESOL) were involved in analyzing the two parts of each title. Initially, the two researchers classified the titles independently. Then, in the weekly meeting, discussion was carried out over disagreed analysis results until consensus was reached. We started our analysis using Anthony's (2001) classification of compound titles in computer science. His categorization encompassed five types – Name : Description, Description : Name, Topic : Description, Topic : Scope, and Topic : Method. Such a categorization, however, was not fully applicable to RA titles in applied linguistics. We thus deleted some categories and added some other categories which were more appropriate for compound titles in our corpus. A new set of categories adapted from Anthony's work was developed, consisting of eleven types of combinations of the two parts of compound titles. Definitions and instances of each category are given below:

- (1) Topic : Scope: The first part indicates a general topic statement, followed by information restricting the scope of the topic, e.g., *The Development of Practices for Action in Classroom Dyadic Interaction: Focus on Task Openings*.
- (2) Topic : Source: The first part is a general topic statement, followed by information indicating the source of data, e.g., *Syntactic Transfer: Evidence from the Interlanguage of Hong Kong Chinese ESL Learners*.
- (3) Topic : Method: A general topic statement is used in the first part, followed by information indicating the research design or research method adopted, e.g., *Effects of Study-Abroad Experiences on EFL Writers: A Multiple-Data Analysis*.
- (4) Topic : Method + Source: The first part is a general topic statement, followed by information depicting both the method of the study and the source of data, e.g., *Pronunciation Issues and EIL Pedagogy in the Periphery: A Survey of Greek State School Teachers' Beliefs*.
- (5) Metaphor : Topic: A metaphorical expression is used in the first part; then, the true topic is unveiled in the second part, e.g. *Slicing the Onion Ethnographically: Layers and Spaces in Multilingual Language Education Policy and Practice*.
- (6) Topic : Metaphor: A general topic statement is presented first, followed by a metaphorical expression, e.g., *Learner Views of Using Authentic Audio to Aid Pronunciation: "You Can Just Grab Some Feelings."*
- (7) Topic : Description: A general topic statement is used in the first part and then a further explanation is given in the second part, e.g., *Differences in Language Skills: Heritage Language Learner Subgroups and Foreign Language Learners*.
- (8) Method : Topic: The research design or method is specified first, followed by a topic statement, e.g., *Cooperative Strategy Training and Oral Interaction: Enhancing Small Group Communication in the Language Classroom*.
- (9) Source : Topic: The first part indicates the source of data, followed by a topic statement, e.g., *Native Speakers of Arabic and ESL Texts: Evidence for the Transfer of Written Word Identification Processes*.
- (10) Topic : Question: A general topic statement is given in the first part, followed by an interrogative statement, e.g., *Second Language Listening: Listening Ability or Language Proficiency?*
- (11) Name : Method: An acronym is given first, followed by information indicating the research method or design of the study, e.g., *CALL: A Survey of K-12 ESOL Teacher Uses and Preferences*.

RESULTS AND DISCUSSION

Syntactic Structures of Titles

Analysis of the syntactic structures of RA titles in the corpus shows that nominal titles constitute the highest percentage of all types of title structures, followed by compound titles, full-sentence titles and titles with other structures (see Table 2). Comparing nominal titles in the three disciplines, we find that a lot of industrial engineering RA titles use this structure (71.8%), as shown in Table 3, although it also forms more than half of sport science RA titles (61.5%) and slightly over one-third of applied linguistics RA titles.

Table 2. *Syntactic Structures of All Titles*

	Occurrences	Percentage (%)
Nominal Titles	261	57.49
Compound Titles	133	29.30
Full-sentence Titles	31	6.83
Other Constructions	29	6.39
Total	454	100.00

Table 3. *Titles with Nominal Construction*

	Total No. of titles	Occurrences	Percentage (%)
AL*	150	58	38.7
IE	156	112	71.8
SS	148	91	61.5

*AL: applied linguistics; IE: industrial engineering; SS: sport science

This result is similar to that in Busch-Lauer's (2000) and Haggan's (2004), which indicated that nominal construction is a structure widely used in research articles regardless of fields. The reason for its wide use is probably its grammatical flexibility to condense information in an economical way. This is achieved via the use of pre-modifiers (e.g., adjectives) or post-modifiers (e.g., prepositional phrases). Some examples are given below.

1. *Ergonomic intervention in carpet mending operation* (Industrial Engineering)
2. *Gender Differences in Leg Stiffness and Stiffness Recruitment Strategy during Two-Legged Hopping* (Sport Science)
3. *Students' and Teachers' Assessments of the Need for Accuracy in the Oral Production of German as a Foreign Language* (Applied Linguistics)
4. *Effects of Input Elaboration on Vocabulary Acquisition Through Reading by Korean Learners of English as a Foreign Language* (Applied Linguistics)
5. *Influence of tyre inflation pressure on whole-body vibrations transmitted to the operator in a cut-to-length timber harvester* (Industrial Engineering)
6. *Impact of a simulated accident in virtual training on decision-making performance* (Industrial Engineering)

Title 1 is a typical example of the least complicated nominal construction. The type of structure contains a noun as a head, preceded by an adjective and followed by a prepositional phrase. Title 2 presents a slightly complex construction with two prepositional post-modifiers. Title 3 shows a highly complex structure by using both pre-modifiers and post-modifiers. This highly complicated structure offers readers a very explicit and detailed picture of their research. It gives a lot of necessary information to readers at their first sight. Titles 4, 5 and 6 are recurrent patterns in our corpus: "effect of A on B," "influence of A on B," and "impact of A on B." However, it is worth noting that when the head noun is in a singular form (e.g., *influence* or *impact*), some researchers tend to omit the article *the*. This reflects the principle of economy in title writing (Day, 1998; Haggan, 2004; Soler, 2007).

Compound structure is also a frequently used construction, as indicated earlier. Compound titles are bi-structured titles in which two parts are juxtaposed on either side of a colon, a dash, or a full stop. Table 4 presents the percentages of compound titles in the three fields. As can be seen, this structure occurs most frequently in AL (52.7%). Thus, we can

conclude that academics in industrial engineering and sport science tend to use nominal titles most often, while compound titles are favored by researchers in applied linguistics.

Table 4. *Titles with Compound Construction*

	Total No. of titles	Occurrences	Percentage (%)
AL	150	79	52.7
IE	156	26	16.7
SS	148	28	18.9

This result seems to reflect disciplinary variation in structuring titles. When constructing titles, AL researchers tend to first indicate a general research topic with which the study is concerned in one part of the compound title, and then narrow down the topic to the focus or the distinctive feature of the study in the other part; for example, in Title 7, “*From Receptive to Productive*” presents the general topic, while “*Improving ESL Learners ...Task*” unveils the focus of the study on using postreading composition task to improve ESL learners’ vocabulary. When reading such titles, readers can quickly and clearly capture the essence of the study.

In contrast, academics in hard sciences, such as industrial engineering and sport science, tend to present their research topic in a more straightforward way, using nominal construction, so as to attain concision and informativity. For instance, Title 8 indicates the topic in the base noun phrase while highlighting the focus or feature of research in the pre-modifiers.

7. *From Receptive to Productive: Improving ESL Learners' Use of Vocabulary in a Postreading Composition Task* (Applied Linguistics)
8. *Biomechanical evaluation of the comfort of automobile clutch pedal operation* (Industrial Engineering)

Furthermore, we find that patterns of “effect of A on B,” “influence of A on B,” and “impact of A on B” are recurrent in both nominal titles and compound titles. In compound titles, these patterns can be arranged in either the first part (Title9) or the second part (Title10). The other part can then be arranged with extra information, thereby providing more clues to readers. However, writers sometimes decompose these patterns, producing two types of recurrent structures: (a) A: effect on B”; (b) B: effect of A. For example, Title 11 uses the pattern of “A: effect on B” while Title 12 uses the pattern of “B: impact of A.” In both titles, A is the topic (the cause of the effect/impact) while B is the focus of the study (the effect/impact).

9. *Effects of correct and transformed visual feedback on rhythmic visuo-motor tracking: Tracking performance and visual search behavior* (Sport science)
10. *EMG trapezius muscle activity pattern in string players: Influences of basic body awareness therapy on the violin playing technique* (Industrial Engineering)
11. *Left-handed versus right-handed computer mouse use: effect on upper-extremity posture* (Industrial Engineering)
12. *Acquisition of Requests and Apologies in Spanish and French: Impact of Study Abroad and Strategy-Building Intervention* (applied linguistics)

As compound titles are widely used in applied linguistics, we further analyzed the semantic relationships of the two juxtaposed elements. Table 5 shows the frequencies of the eleven types of compound titles we have identified in applied linguistics, with a total of 79

occurrences. The top three categories are Topic : Scope, Topic : Source and Topic : Method relationships, amounting to 70.89 % out of all occurrences.

Table 5. *Semantic Relationships of Compound Titles in Applied Linguistics*

	Occurrences	Percentage (%)	Rank
Topic : Scope	25	31.65	1
Topic : Source	19	24.05	2
Topic : Method	12	15.19	3
Topic : Method + Source	6	7.59	4
Metaphor : Topic	6	7.59	4
Topic : Description	4	5.06	6
Topic : Metaphor	2	2.53	7
Topic : Question	2	2.53	7
Method : Topic	1	1.27	9
Source : Topic	1	1.27	9
Name : Method	1	1.27	9
Total	79	100.00	

In Anthony's (2001) research, the top three types of compound titles in computer science are Name : Description, Topic : Scope and Topic : Method combinations, constituting 92.69% of all compound titles. The Description : Name and Topic : Description combinations in Anthony's corpus constitute the remaining 7.31%. Comparing our results with Anthony's (2001) shows that there seem more variations in the semantic combinations of compound titles in applied linguistics than in computer science. Also, some combinations may be discipline-specific. For instance, the Name : Description relationship (e.g., *ADOME: An Advanced Object Modeling Environment*) is the top one category in computer science but in our corpus we did not find any title using this combination. In computer science, research often proposes an innovative technology, a new model, a new device, etc., which is often coined with a specific name. Constructing titles with the Name : Description title can clearly tell readers what their invention is. However, in social sciences, such as applied linguistics, this occurs less often. In contrast, the Topic: Source title is a more prevailing category in applied linguistics (see Titles 15 and 16). As it is important for readers to understand the source of data in this field, writers often inform readers, in the titles, of the research site, context or the participants from whom the data are collected. Some instances of compound titles in applied linguistics are listed below.

13. *Constructing Gender in an English Dominant Kindergarten: Implications for Second Language Learners* (Topic : Scope)
14. *Speaking Foreign Languages in the United States: Correlates, Trends, and Possible Consequences* (Topic : Scope)
15. *Applying Conceptual Grammar to Advanced-Level Language Teaching: The Case of Two Completive Constructions in Korean* (Topic : Source)
16. *Speech Rhythm in World Englishes: The Case of Hong Kong* (Topic : Source)
17. *Effects of Study-Abroad Experiences on EFL Writers: A Multiple-Data Analysis* (Topic : Method)
18. *"I'm Tired. You Clean and Cook." : Shifting Gender Identities and Second Language*

Socialization (Metaphor : Topic)

19. *What Level of English Proficiency Do Elementary School Teachers Need to Attain to Teach EFL? Case Studies from Korea, Taiwan, and Japan (Topic : Method + Source)*

Titles 13 and 14 are compound titles with the Topic : Scope relationship. In Title 13, the topic of the study is first indicated, and then it is narrowed down to pedagogical implications of research, focusing on this specific aspect of research. In addition, if a study has several highlights, the researchers tend to summarize them in the title. Title 14 is a good instance of this type of research. Title 17 is a compound title showing Topic : Method relationship emphasizing the unique research method designed or adopted in the study. Of various research methods that occur in the compound titles of our corpus, *case study* is the most frequently mentioned. Sometimes the researchers intend to emphasize not only their research method but also their research context or participants, then the Topic : Method + Source combination will be used. For instance, Title 19 indicates not only the topic but also the research method and source of data as well. The Metaphor : Topic compound titles are an interesting category which occurs only in applied linguistics titles. This type of titles tries to attract readers by using a metaphor in one part and revealing the real topic of their research in the other part. For instance, when readers read the first element in Title 18, they may be puzzled, but attracted, by the metaphorical expression of “*I’m tired. You Clean and Cook.*” As they read the other part of the compound title, *Shifting Gender Identities and Second Language Socialization*, which reveals the research topic, they realize what is implied in the metaphor. The use of this construction can arouse readers’ curiosity or even leave a strong impression on readers.

A third type of title structure is declarative full-sentence construction. As shown in Table 6, this construction is characteristic of sport science titles, constituting 14.9% of all titles, in contrast with 0% and 1.28% in the other two disciplines. This finding confirms one of Soler’s (2007) findings that full-sentence titles are not only a feature of biology papers but also a peculiarity of branch fields of biology, such as medicine and biochemistry.

Table 6. *Titles with Declarative Full-sentence Construction*

	Total No. of titles	Occurrences	Percentage (%)
AL	150	0	0.00
IE	156	2	1.28
SS	148	22	14.90

Development in biology-related fields has been very rapid in recent years. To help readers grasp the fleeting trends in these fields, coining titles with full-sentence construction provides a good structure which can directly and completely reveal the findings of research to readers. Two instances of SS titles having this construction are presented below.

20. *Advancing age progressively affects obstacle avoidance skills in the elderly (Sport Science)*
21. *Limiting the recruitment of degrees of freedom reduces the stability of perception–action patterns (Sport Science)*

In contrast to titles with declarative construction are titles with interrogative full-sentence construction. Day (1998) suggested that titles in a question form should be avoided; however, our analysis reveals this construction is indeed used, as shown in Table 7, although it constitutes only a very small percentage (1.54% out of all titles).

Table 7. *Titles with Interrogative Full-sentence Construction*

	Total No. of titles	Occurrences	Percentage (%)
AL	150	3	2.0
IE	156	3	1.9
SS	148	1	0.7

This title structure makes retrieval process difficult when readers use the indexing system (Soler, 2007). Nevertheless, it still bears pragmatic value. Titles in a question form can arouse readers' curiosity and form a certain degree of expectation in readers. They tend to read the article itself in order to find the answer to the question raised in the title. Such interaction between writers and readers is similar to the adjacency pair—questions and answers—in our daily conversation where the question performs the function of raising hearers' expectations (Haggan, 2004). Three instances of titles in question form in our corpus are shown below.

22. *Is There an "Academic Vocabulary"?* (Applied Linguistics)
23. *Are the postures adopted according to requested linguistic categories similar to those classified by the recording protocols?* (Industrial Engineering)
24. *Is the learning of goal-directed displacement effector-independent?* (Sport Science)

Goodman et al., (2001) advised that “a title include not only the topic, but also study design and methods and when space allows, results and conclusions” (P. 78). However, such an all-embracing title is too difficult to be constructed. First, although titles are miniatures of RAs, due to the space limit, it is impossible to put all major ideas of an article in the title. Moreover, the information in titles should be only the most revealing. For example, it is not necessary for every title to indicate the research method. The inclusion of the method in the title is essential only when the study emphasizes a self-designed or specific research method. If data source is critical in investigating a given research question, then an effective title should designate where these data come from.

On the other hand, the structures discussed above have shown different syntactic advantages in reporting information. Declarative full-sentence structures, for instance, can tell readers their major results due to their S-V-O structure. Nominal construction, in contrast, mainly aims to present the topic of research. Compound titles offer a flexible form for writers to combine any two main elements of research, such as the topic, scope, source of data, or method.

CONCLUSION

A title is a miniature of the whole paper. A successful article can be easily recalled by its title. In this study, we compiled a corpus consisting of RA titles from three disciplines—applied linguistics, industrial engineering, and sport science, with a total of 454 titles, aiming to investigate how academics in the three fields structure their RA titles. The results show nominal titles predominate in the corpus, constituting 57% of all titles. Compound titles rank second (29.3%). However, they are favored by applied linguistics researchers, exceeding half of applied linguistics RA titles (52.7%). Declarative full-sentence construction titles are found a peculiarity in sport science while rarely occur in the other two disciplines.

Furthermore, the analysis of the two parts of compound titles in applied linguistics, reveals that the top three frequent combinations are *Topic : Scope*, *Topic : Source*, and *Topic : Method*, constituting about 71% of all applied linguistics titles. A comparison of this result

with Anthony's (2001) analysis of colonic titles in computer science also shows disciplinary variation. Overall, of various title structures found in our corpus, titles in industrial engineering and sport science tend to be more informative and concise. Titles in applied linguistics, in contrast, are found relatively unclear, abstract or even abstruse. The reason is probably because academics in applied linguistics are expert scholars having good command of language, they thus tend to play words in titles so as to make titles brilliant and extraordinary.

The pedagogical implications of this study are three-fold. First, since different syntactic structures offer writers different ways to organize and report research information, RA learners should be informed of their respective features. Also, although nominal titles and compound titles are more prevailing than other structures in the three disciplines we studied, disciplinary variation should be noted. Moreover, compound titles are a feature in academic writing; however, the combinations of the two parts may differ from discipline to discipline. EAP instructors should inform students of the most frequently used categories in students' disciplines.

As to the limitation of our research, as this study is a small-scale study, future research with a larger corpus consisting of titles from a wider range of journals would yield more convincing results. Moreover, we only analyzed applied linguistics compound titles, a more comprehensive study can be conducted to unveil how title structures may reflect research trends in a specific field.

REFERENCES

- Anthony, L. (2001). Characteristic features of research article titles in computer science., *IEEE Transactions on Professional Communication*, 44(3), 187-194.
- Anthony, L. (2009). AntConc (Version 3.2.1w) [Computer software] Retrieved February 25, 2009, from <http://www.antlab.sci.waseda.ac.jp/software.html>
- Busch-Lauer, I. A. (2000). Titles of English and German research papers in Medicine and Linguistics. In A. Trosborg (Ed.), *Analysing Professional genre* (pp. 77-94). Amsterdam: John Benjamins.
- Davis, M. (1997). *Scientific Papers and Presentation*. San Diego: Academic Press.
- Day, R. A. (1998). *How to write and publish a scientific paper*. Cambridge, UK: Cambridge University Press.
- Dillon, J. T. (1981). The Emergence of the Colons: An Empirical Correlate of Scholarship, *American Psychologist*, 36(8), 879-884.
- Goodman, R. A., Thacker, S. B., & Siegel, P. Z. (2001). A descriptive study of article titles in peer-reviewed medical journals. *Science Editor*, 24(3), 75-78.
- Haggan, M. (2004). Research paper titles in literature, linguistics and science: dimensions of attraction. *Journal of Pragmatics*, 36(2), 293-317.
- Hartley, J. (2005). To attract or to inform: What are titles for? *Journal of Technical Writing and Communication*, 35(2), 203-213.
- Hartley, J. (2007a). Planning that title: Practices and preferences for titles with colons in academic articles. *Library & Information Science Research*, 29(4), 553-568.
- Hartley, J. (2007b). There's more to the title than meets the eye: Exploring the possibilities. *Journal of Technical Writing and Communication*, 37(1), 95-101.
- Hawes, T., & Thomas, S. (1997). Tense choices in citations. *Research in the Teaching of English*, 31(3), 393-414.
- Kilgour, F. G. (2004). An experiment using coordinate title word searches. *Journal of the American Society for Information Science and Technology*, 55(1), 74-80.
- Lewis, G., & Hartley, J. (2005). What's in a title? Numbers of words and the presence of colons. *Scientometrics*, 63(2), 341-356.

- Lim, J. M. H. (2006). Method sections of management research articles: A pedagogically motivated qualitative study. *English for Specific Purposes*, 25(3), 282-309.
- Malcolm, L. (1987). What rules govern tense usage in scientific articles? *English for Specific Purposes*, 6(1), 31-43.
- Soler, V. (2007). Writing Titles in Science: An Exploratory Study. *English for Specific Purposes*, 26(1), 90-102.
- Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. New York: Cambridge University Press.
- Swales, J. M., & Feak, C. B. (1994). *Academic writing for graduate students essential tasks and skills* (2nd ed.). Ann Arbor, MI: Univ. Michigan Press
- Wang, Y., & Bai, Y. (2007). A corpus-based syntactic study of medical research article titles. *System*, 35(3), 388-399.
- Yang, R. & Allison, D. (2003). Research articles in applied linguistics: moving from results to conclusions. *English for Specific Purposes*, 22, 365-385.
- Yitzhaki, M. (1994). Relation of title length of journal articles to number of authors. *Scientometrics*, 30(1), 321-332.

A Comparison of Effects of BLOG-based and Paper-based Portfolios on Writing Attitudes and Performances

Mei-hua Lan (藍美華)
Tzu-Hui Institute of Technology
Emmalan2000@yahoo.com.tw

This study compared the effects of the BLOG-based and paper-based writing portfolios on junior college students' attitudes toward English writing, writing achievements and holistic writing performances. Totally 46 subjects, 4th graders, from two classes of the Applied Foreign Language Department at two junior colleges in southern Taiwan were involved in this study. The two classes were randomly assigned to be the experimental group (EG) and the control group (CG). Pre-treatment General English Proficiency Test (Elementary Level) was implemented, and the t-test analysis of the GEPT writing scores indicated no significant differences between groups. The results indicated that both groups had compatible English writing proficiency before treatments. The study was implemented for 16 weeks, with totally 32 class hours. Both groups were provided with similar writing activities and instructions, and were asked to compose for the same topics, and have multiple revisions. However, the EG employed BLOG to conduct electronic English writing portfolios, while the CG compiled traditional paper-based writing portfolios. Also, the Questionnaire on the Attitudes toward English Writing (with 20 five-scale items), and the Achievement Test on the Target Grammar and Descriptive Writing (with 10 sentence-combination items and a picture-guided descriptive writing) were distributed before and after the treatments. The attitude questionnaire was designed by the researcher, and it had acceptable reliability (Cronbach $\alpha = .82$) and validity (KMO = .83). The results of the t-test analyses on group comparisons of the pre-study and the post-study questionnaires indicated that the BLOG-based writing portfolios did not lead to the subjects' better writing motivation, affections and less reluctance to writings. Rather, the paper-based writing portfolios facilitated the CG subjects' more positive writing attitudes by enhancing better writing values and beliefs. Furthermore, the t-test analyses of the writing achievement tests showed that the BLOG-based writing portfolio project did not help the subjects have better writing achievements than the paper-based one. In addition, the t-test analyses of the holistic writing performances of the first drafts and the final drafts on different topics by both groups indicated that paper-based writing portfolios had more positive impacts on the subjects' holistic writing performances on a variety of topics. The findings could provide important insights for the applications of portfolios for English writing instructions at technical colleges.

INTRODUCTION

The ability to write and communicate in English is getting more important in the global community. The advancements in transportation and technology make the interactions among people from different nations and cultures possible, and writing is an important means for intercultural communication. To express thoughts and needs effectively and efficiently, one has to know and learn how to speak or write clearly and intelligibly. Recognizing the significant roles of English writing in the intercultural communications, the explicit English writing instructions in Taiwan are re-emphasized, and English writing is included as a part of the University Entrance exam and other language proficiency certification exams. Recently, teachers and researchers have been studying and practicing innovative English writing pedagogy and assessment in order to teach EFL students practical writing strategies, and to motivate them to like English writing better. To motivate EFL learners in Taiwan to learn English writing and to help them have successful intercultural communications, explicit

English writing instructions and meaningful tasks as well as assessments are essential.

With the reform movements in English writing education currently, many educators and researchers have innovated their pedagogies. In the English writing reform in Taiwan recently, English teachers are encouraged to empower students in their writing, and to promote their autonomy (Chen, 2000, 2005; Chiang, 2002; Sun, 2003; Wu, 2003). Student writers no longer play the roles of passive writers but active writers in choosing something meaningful to write (Fosnot, 1996; Harste & Lowe, 1991), planning or revising for their writings, and reflecting on their writings throughout the entire writing process (Carlos, 1995). Unlike the traditional approaches with an emphasis on grammar and accuracy, recent writing instructions highlight the significance of students' thoughts, personal expressions and active construction processes.

The changes of perspectives in English writing theories and approaches imply that the assessment methods should be modified and improved in order to effectively evaluate learners' writing progresses, and to help them achieve their writing goals better. Alternative assessments, such as portfolios, have been proposed, for their capacities to provide more holistic, meaningful and process-related information regarding students' writing performances. In a word, current innovations in English writing instruction and assessment emphasize learners' active participations and learner-centered activities. Hence, portfolios are widely used for language learning, especially for reading and writing, as an alternative for traditional instructions and assessments.

The study aims to investigate the effects of the electronic portfolio project on English writing attitudes and performance of EFL learners at technical colleges. To investigate the effects of the electronic portfolio project on the subjects' English writing performances, English descriptive writings by the experimental groups (EG) and control groups (CG) were analyzed and compared. The English writing performances were analyzed in terms of the holistic assessment of subjects' writings. The purpose of the analysis of holistic writing performances is to get a general idea about how the portfolio project affects the subjects' overall performances in descriptive writings. The related studies on portfolios in the United States and in other L1 contexts have indicated that portfolios have positive impacts on learners' attitudes toward literacy learning (e.g. Paris & Ayres, 1999; Tiernery, Carter, & Desai, 1991). This study thus intends to figure out if portfolios can have similar positive effects on nurturing EFL learners' attitudes toward English writing. The comparisons of the subjects' attitudes toward English writing before and after treatments provide with information regarding the effects of portfolios on writing attitudes. Electronic portfolios have been suggested beneficial for enhancing students' better writing performances and nurturing positive attitudes toward writing (Frazler, 1995; Phinney, 1991). Therefore, employing BLOG, a commonly used freeware by teenagers, this study hopes to help EFL learners at technical colleges in Taiwan establish their own electronic portfolios, to inspire learners' interests in English writing, and to enhance better performances and attitudes toward English writing.

LITERATURE REVIEW

Portfolios have been used for different domains of education. Most studies are conducted in L1 settings, and recently the use of portfolios is expanded to the ESL/EFL settings. The majority of related studies concerns of the effects of portfolios on literacy development (e.g. reading and writing), learning of writing, and learning motivations or attitudes. In Taiwan, portfolios have been used in different domains and levels with the emphasis on qualitative analysis of the effects of portfolios on learning of English, literacy development and culture learning.

Portfolio assessments have been employed in different levels for a variety of purposes in Taiwan, and relevant studies were summarized in Table 1. According to Table 1,

portfolio assessments have been used for English learning, literature studies, learning of English reading and writing as well as culture learning in elementary schools, junior or senior high schools, and even colleges. As shown in Table 1, the EFL studies on English writing portfolios in elementary schools found that portfolio assessments helped students improve their grammar, writing organizations, contents and styles (Tzou & Kan, 1999). English portfolio assessments were also perceived as a valid tool for improving EFL students' learning of English, yet they were concerned of being time-consuming (Hsieh, 2000). Although portfolios have been claimed to be beneficial for English learning, they are concerned due to their demands of great time and efforts (Hsieh, Lu & Yeh, 2000). In brief, the findings of these EFL studies on portfolio assessments in elementary schools generally support the advantages of portfolio assessments for learning of English and English writing despite of some problems, such as demands of greater time and efforts.

Table 1. *Portfolio Assessments in EFL Studies in Taiwan*

	Researchers	Year	Experiment Durations	Types of Portfolio Assessments	Findings
Elementary Schools	Tzou & Kan	1999	a semester	Writing portfolio assessment	1. Writing improvements in grammar and organization, contents and styles.
	Hsieh	2000	a semester	English learning Portfolio assessment	1. Positive perceptions of portfolio assessments. 2. Complaints about being time-consuming
	Hsieh, Lu & Yeh	2000	a semester	English learning Portfolio assessment	1. Improvements in English learning. 2. Positive attitudes toward portfolio assessments.
High Schools	Chao	2000	a semester	Portfolio assessments in whole language teaching	1. Improvements in writing productions. 2. Problems due to lacks of clear criteria for writing assessments.
	Song	2002	16 weeks	English writing portfolio assessment	1. Writing improvements in dictions, expressions and organizations. 2. Positive responses to portfolio assessments.
	Chang	2003	8 weeks	Portfolio assessment for language problem diagnoses	1. Benefits of portfolio assessments for developments of English reading and writing abilities. 2. Students' more interests in learning of English reading and writing

	Researchers	Year	Experiment Durations	Types of Portfolio Assessments	Findings
High Schools	Chang & Chang	2003	a semester	Portfolio assessment with Multiple Intelligence (MI) Teaching	<ol style="list-style-type: none"> 1. Better performances in achievement tests for MI portfolio assessments than traditional assessments. 2. Better class climate and learning motivation
	Huang	2003	7 weeks	Festival portfolio assessment	<ol style="list-style-type: none"> 1. Positive responses to four types of portfolio assessments. 2. Teacher-assessments' being most beneficial. 3. Self-assessments' being least beneficial
	Sun	2003	16 weeks	Children literature portfolio assessment	<ol style="list-style-type: none"> 1. Improvements in reading comprehensions and strategies 2. Positive attitudes toward portfolio assessments.
Colleges	Chen	1999	a year	English writing portfolio assessment	<ol style="list-style-type: none"> 1. Benefits of portfolio assessment for recognizing students' writing developments 2. Complaints about being time-consuming and lacks of computer resources.
	Chen	2000	a year	English writing portfolio assessment	<ol style="list-style-type: none"> 1. Improvements in reading and writing skills. 2. Improvements in reflective/critical thinking
	Chen	2005	a year	Portfolio-based English writing curriculum	<ol style="list-style-type: none"> 1. Benefits of the portfolio-based curriculum for learner autonomy and collaborative learning 2. Benefits of portfolio assessments for developing confidence in English learning and critical thinking 3. Problems of portfolio assessments (e.g. time-consuming, bias in peer-assessment, limited effects for test performances.)

Several EFL studies on portfolio assessments in Taiwan concerned the uses and effects of portfolio assessments in high school English learning, and mostly positive results were drawn. According to Table 1, portfolio assessments were employed with other language teaching approaches, such as the whole language teaching approach (Chao, 2000) and multi-intelligence teaching approach (Chang & Chang, 2003). The results of Chao's (2000) study indicated that the students became more productive writers, while significant improvements were found in student achievement tests, learning motivation, and class climate (Chang & Chang, 2003). Some studies on portfolio assessments employed in high

schools claimed the benefits of portfolio-assessments for language developments (Chang, 2003), writing developments (Song, 2003), literature learning and reading developments (Sun, 2003) and culture learning (Huang, 2003). Some portfolio assessment studies in high schools also indicated the positive results of portfolio assessments for motivating for learning of reading and writing and cultures (Sun, 2003; Huang). In short, these studies on portfolio assessments in high schools lent support for the positive results of portfolio assessments on literacy development, literature and culture learning.

The studies on portfolio assessments employed in colleges focused on investigating effects of English writing portfolios, and EFL students' responses to portfolio-based curriculum. The portfolio-based tasks were found to make the EFL college students better learners, readers and writers (Chen, 1999), and they were conducive to students' personal growth and learning reflection (Chen, 2000). The EFL college students also indicated that the portfolio-based writing was advantageous for examining self-learning, promoting mutual learning between peers, increasing confidence and motivation, enhancing multidimensional development and being fun and joyful. It was also beneficial in developing English confidence, learning ownership, versatile talents and critical thinking. Problems, such as time-consuming, bias in peer evaluation, limited effects on test performance, difficulties in storage and portability and the effect of language proficiency on portfolio quality, were suggested. These studies confirm the findings of portfolio assessments in L1 studies regarding the advantages and limitations of portfolio assessments.

The above EFL studies on portfolio assessments in English instructions in Taiwan have shown positive effects of portfolio assessments on learning of English, especially on reading and writing developments. However, some of the studies were conducted for a very short period of time (e.g. less than one semester, or a few weeks). Comparatively fewer studies concerned the long-term effects of portfolio assessments and applications of portfolio assessments for language learning on the college level. In addition, few studies analyzed how students' English writings improved throughout the English writing processes, or compared the effects of electronic portfolios with paper-based ones. Concerning some limitations of previous studies, the present study hence attempts to analyze the effects of electronic and paper-based portfolios on EFL students' English writing performances and attitudes from a variety of perspectives. In addition to an emphasis on qualitative data in previous studies, quantitative data is incorporated in the present study in order to provide readers and researchers with more evidence for the effects of portfolio assessments on EFL students' learning and development of English writing.

dy and by an expert for the concern of validity.

METHODOLOGY

Subjects

Totally 46 subjects from the Applied Foreign Language Departments at Tzu-Hui Institute of Technology and Su-Zen College of Medicine and Management were involved in this study. The participants were fourth graders, aged 17 to 19, and most of them had learnt English for at least six years. None of them had experienced in composing their portfolios for English subjects before. The class at Tzu-Hui Institute of Technology, 28 participants, was chosen to be the experimental group, while the class at Su-Zen College of Medicine and Management, 18 participants, served as the control group. None of the participants passed the Intermediate General English Proficiency Test (GEPT), and less than one third of students in both groups passed the Basic GEPT exams. The two groups were chosen because they had the same majors, and compatible language proficiency as measured by GEPT. Besides, after three years of English training at junior colleges, the fourth graders were more likely to express themselves better in writing, comparing to the lower graders, whose English

proficiency might not be good enough to conduct paragraph or essay writing. Also, **both** Tzu-Hui and Su-Zen are medical junior colleges, and share a lot of similarities.

Instruments

To conduct this study and collect data for this study, the researcher adopted 6 instruments as follows:

Retired Elementary General English Proficiency Test (GEPT). The retired GEPT (Elementary-2) was published by the Language Teaching and Testing Center (LTTC, 2002), and it had been used for identifying the subjects' proficiency levels. The retired GEPT used to be a formal test with acceptable reliability. In the present study, only the writing scores of the retired GEPT were used in this study. Before the treatments, the Elementary GEPT retired test was distributed to the subjects in both groups. The average score for the EG was 65.20 while that of the CG was 60.49. The t-test statistical analysis indicated no significant differences among the two groups ($p = 0.47 > .05$.) The result suggested that both groups had compatible writing proficiency before treatments as measured by the GEPT.

Teaching Materials and Worksheets. The teaching materials for English writings referred to two textbooks, including *Composing through Pictures* by Heaton (2005), and *College Writing-From Paragraph to Essay* by Zemach and Rumisek (2003). Worksheets were designed based on the related topics to help the students plan, write and enrich the contents of their compositions. They also aimed to assist the subjects to choose appropriate topics for their writings, and to be aware of the organizations of English writings. The differences in the organizations of English and Chinese writings were emphasized and discussed, too.

A Questionnaire on Attitudes toward English Writing. The questionnaire was developed referring to Chen (1999), Gungle and Taylor (1989) and Rose (1984). The questionnaire was originally written in Chinese, and it included two parts with twenty-seven questions. The first part of the writing attitude questionnaire in this study contained twenty five-point-scale items in terms of writing motivation, affection, values, beliefs and blocking. The second section had two open-ended questions. The questions asked the subjects to response to the factors that made them like, and dislike English writing. To ensure the reliability of the Questionnaire on Attitudes toward English Writing, a pre-test for the questionnaire ($n = 112$) was conducted. The data of the pre-test questionnaires were computed and analyzed by Reliability Statistics in Scale and Kaiser-Meyer-Olken (KMO) and Bartlett's Test of Factor Analysis of SPSS 10.0. According to Bryman and Cramer (1997), Gay (1992) and Wu (2003), if the value of Cronbach's α of a questionnaire is higher than .80, the questionnaire is considered as reliable. The reliability coefficient ($\alpha = .82$) of the Questionnaire on Attitudes toward English Writing in this study was considered pretty reliable since the Cronbach's coefficient alpha (α) of the questionnaire was higher than the acceptable values ($\alpha = .80$) as the related studies suggest (Bryman & Cramer, 1997; Gay, 1992; Wu, 2003). The KMO Measure of Sampling Adequacy was .83 (higher than .80). Hence, the five-point-scale items of the Questionnaire on Attitudes toward English Writing were statistically reliable and valid.

The Pretest and the Posttest of Target Grammar and Descriptive Writing. To evaluate the post-treatment writing achievements of the subjects in this study, The Pretests and Posttests on Target Grammar and Descriptive Writing were designed by the researcher. Both pretests and posttests included two parts. The first part contained ten sentence-combination items for testing the target grammar of the writing instructions in this study, including relative-clauses, conjunctions for stating reasons and some prepositions. The second part of the tests required the test-takers to write a descriptive paragraph or a story based on the picture cues. The scoring rubric and calculation formula were attached to the achievement tests. For the sentence-combination items, the maximum score for each item was 2 for those error-free utterances. The minimum score was 0 for those unintelligible or

not-answered questions. For those sentences with minor errors but correct sentence structures, score 1 was given. For the descriptive paragraph, holistic scores in terms of 1 point to 5 points were given to the subjects' writings based on the provided rubrics that specified the characteristics of each scoring. The first part accounted for 50% of the total score, and the second part for another 50%. The maximum score for the achievement test was 100.

The test items in the two parts of achievement tests were designed on the basis of the contents and activities of the writing instructions involved in this study. Chen (2000) claimed that validity could be enhanced if the assessment items corresponded to the goals and contents of instructions. Before implementation of the formal pretests and posttests, a pilot study was conducted to ensure the reliability of the tests, and the test items were reviewed by the teachers involved in the study and by an expert for the concern of validity.

Self-/Peer- Assessment Forms. Two self-assessment forms and two peer-assessment forms will be used in this study, including the Self-Assessment Form for English Writing, the Self-Assessment Form for English Writing Portfolio, the Peer-Assessment Form for English Writing and Peer-Assessment Form for English Writing Portfolio. The forms were developed by referring to the relevant studies (e.g. Chang, 2003; Chen, 1999; Huang, 2003; Lefkowitz, 1985; Pollari, 2000; Sun, 2003; Tedick, & Klee, 1998) for helping the subjects reflect on their writings and portfolios.

Teacher Assessment Forms and Criteria. Two teacher-assessment forms were employed in the study, including the Teacher-Assessment Form and Criteria for English Descriptive Writing, and the Teacher-Assessment Form for English Writing Portfolio. The researcher of this study designed this Teacher-Assessment Form and Criteria for English Writing by referring to a couple of related studies, including Subrick (2003), Tedick and Klee (1998), and Valdez Pierce and O'Malley (1992). The raters of the Teacher-Assessment Form evaluated the overall performances of the subjects' writings based on the attached holistic scoring rubrics for descriptive writings, and they provided the subjects with holistic scores (e.g. 1 to 5) for their writings. For designing the Assessment Form and Criteria for English Writing Portfolio, the researcher of this study referred to a couple of related studies (e.g. Dogger, Moy & Nogami (2003); Pierce & O'Malley, 1992). The raters of the form filled out a checklist for assessment items regarding the appearances, basic requirements, contents/organizations and grammar/mechanics of the writing portfolios. Detailed weights of each sub-section scores and computing formula descriptions were attached to the assessment forms.

A Feedback Questionnaire to the Electronic Portfolio Project. This Feedback Questionnaire on the Electronic Portfolio Project was designed by the researcher to know the subjects' feedbacks to the electronic portfolio project. It had two sections. The first part included questions concerning the subjects' responses to the electronic portfolio project, the effects of electronic portfolios on their English writing progress, and the impacts of self- and peer-assessments, holistic writing assessments on the improvements of their English writings. The second section contained open-ended questions regarding why the subjects liked or disliked the electronic portfolio projects, what problems they encountered during the writing process, and what benefits they gained from the portfolio projects. The information from this feedback form provided important information for the qualitative analysis of the effects of the electronic portfolio project on the subjects' English writing.

Procedures

The present study was implemented in the two selected classes for 16 weeks, with totally 32 instruction hours. The study was conducted in three stages. In the first stage, the preparation stage, the instruments for the study were developed, including the teaching materials, worksheets, assessment forms and criteria, and questionnaires. The teaching and assessment procedures

were discussed and determined by the teachers involved in this study. In addition, the design for the electronic portfolios was developed. In the second stage of this study, the selected subjects were randomly chosen to be the experimental (EG) and the control group (CG). Both groups had similar writing instructions, but slight different writing and assessment processes. Before the writing instructions, the GEPT was given to both groups of subjects to control the variable of writing proficiency of the subjects in this study. Also, the *Questionnaire on Attitudes toward English Writing* and the *Pretest of Target Grammar and Descriptive Writing* were distributed to both groups before treatments to shed a light on subjects' attitudes toward English writing and writing performances before treatments. Both groups were provided with similar instructions on descriptive writings, target grammar and vocabulary based on the worksheets for each topic during the instructions. Likewise, both groups were asked to draft, revise and edit their writings. Multiple drafts were required for each topic by all subjects in both groups. Both groups underwent identical writing portfolio processes, and experienced many group discussions, self-assessments, peer assessments, and development of personal writing portfolios. Both groups conducted self-assessments and peer-assessments after the first drafts for each topic. Then, they revised for the second drafts based on their self-assessments and peers' comments and suggestions. The raters assessed and corrected the second drafts, and provided the subjects with suggestions and comments in the *Teacher Assessment Form for English Writing*. After that, the students revised and edited for their final drafts. The major difference of the EG from the CG was that during the writing for each topic, the EG subjects were required to conduct many group discussions and assessments online via BLOG. At the last stage of the electronic portfolio project, the EG subjects were asked to choose the best three or the most favorite three compositions, and composed their personal electronic writing portfolios based on the provided guidelines. The design as discussed above was to control the possible confounding variables. In the last stage of the study, the subjects in both groups were asked to take the *Post-study Questionnaire on Attitudes toward English Writing* and the *Post-test of Target Grammar and Descriptive Writing* after treatments while the EG subjects took the *Feedback Questionnaire on the Electronic Portfolio Project*. Then, the collected data were analyzed and discussed.

RESULTS AND DISCUSSIONS

This study aimed to investigate the effects of the English writing portfolio project on the writing achievements and attitudes of the EFL students at technical colleges, and the results were described in the following tables. Table 2 to Table 4 help illustrate the effects of the writing portfolio project on the subjects' writing achievements as measured by the writing achievement tests before and after treatments (i.e. pretests and posttests). Table 5 to Table 7 shows the effects of the portfolio project on the subjects' attitudes toward English writing in terms of writing motivation, affection, value, belief and blocking. Table 8 to Table 10 compare the effects of the BLOG-based and paper-based portfolio projects on the EG and the CG subjects' holistic writing performances for each writing topic. The findings were discussed as follows.

The Effects of the English Writing Portfolio Projects on Writing Achievements

Table 2 to Table 4 show the t-test analyses of the pretests and posttests by the EG and the CG. The results help understand the effects of different portfolio projects on junior college students' writing performances.

Table 2. *Group Comparisons of the Writing Achievement Tests*

Test	Group	No	Mean	SD	<i>t</i>	<i>p</i>
Pre-test	EG	28	58.39	16.38	-.94	.35
	CG	18	62.92	15.35		
Post-test	EG	28	60.80	17.35	-4.17	.00*
	CG	18	79.72	11.01		

*Maximum score=100; $p < .05$

Table 3. *Paired t-test Analysis of the Pre-test and the Post-test by the EG*

Group	No	Pre-test	Post-test	SD	<i>t</i>	<i>p</i>
EG	28	58.39	60.80	16.27	-.78	.44

*Maximum score=100 ; $p < .05$

Table 4. *Paired t-test Analysis of the Pre-test and the Post-test by the CG*

Group	No	Pre-test	Post-test	SD	<i>t</i>	<i>p</i>
CG	18	62.92	79.72	13.08	-5.45	.00*

*Maximum score=100 ; $p < .05$

According to Table 2, the Mean scores of the pretest (58.39) and posttests (62.92) of the EG is lower than that those of the CG (62.92 and 79.72). Table 2 also indicates a significant difference between the two groups in the posttests. Table 3 and Table 4 show that there is a significant difference between the pretest and the posttest by the CG, but no significant difference was found in that of the EG. The results imply that the CG outperformed the EG in the English writing achievement tests, and the paper-based English writing portfolio project helped the subjects have more improvements in writing performances as measured by English writing achievement tests than the BLOG-based English writing portfolio project did. However, it should be noted that both groups showed certain degrees of progress in posttests, and the result suggests that portfolio-based writing projects helped the subjects in this study have better writing achievements, especially the paper-based writing portfolios.

The Effects of the English Writing Portfolio Projects on English Writing Attitudes

Table 5 to Table 7 show the frequency and the t-test analyses of the *Pre-study and the Post-study Questionnaires on Attitudes toward English Writing* by the EG and the CG. According to Table 5, most subjects in both groups (more than 40% to 60%) agreed to have interests in learning English writing knowledge and a variety of writing skills (about 50% to 80%), and to be

Table 5 Frequency Analysis of the Pre-study and the Post-study Questionnaires on Attitudes toward English Writing by the EG and the CG

Questions	Test	EG (%)					CG (%)					
		5	4	3	2	1	5	4	3	2	1	
M O T I V A T I O N	1. 我對英文寫作很有興趣	Pre	4	21	64	11	0	11	33	17	33	6
		Post	4	36	21	39	0	6	28	61	6	0
	2. 我喜歡學習有關英文寫作的新知識	Pre	0	54	39	7	0	17	61	11	6	6
		Post	4	57	18	21	0	33	11	50	6	0
T	3. 我想學習並嘗試不同的英文寫作方法	Pre	4	50	43	4	0	28	56	6	6	6
		Post	11	54	14	21	0	22	39	33	6	0
I	4. 我會想用英文寫出我的想法和感受	Pre	7	61	29	4	0	39	50	6	6	0
		Post	11	64	14	11	0	17	67	17	0	0
A F F E C T I O N	5. 我覺得用英文寫作很有趣	Pre	0	36	57	7	0	17	39	33	6	6
		Post	0	39	29	32	0	11	28	61	0	0
F	6. 我喜歡上英文寫作課	Pre	4	21	68	7	0	6	28	50	11	6
		Post	0	32	43	25	0	6	22	51	17	0
C	7. 當我學到新的英文作文技巧時，我覺得很高興	Pre	7	57	36	0	0	50	44	6	0	0
		Post	14	79	4	4	0	39	44	17	0	0
T I O N	8. 我覺得用英文寫作令我有成就感	Pre	14	46	39	0	0	39	33	17	11	0
		Post	18	57	14	11	0	39	44	17	0	0
V A L U E S	9. 我覺得自己的英文作文寫得很好	Pre	0	0	18	61	21	0	0	11	67	22
		Post	0	44	64	25	0	0	6	44	39	11
A L L	10. 我覺得英文寫作很重要	Pre	21	46	32	0	0	33	61	0	6	0
		Post	29	57	11	4	0	6	78	17	56	11
U E S	11. 我喜歡自己寫的每篇英文作文	Pre	4	11	43	39	4	0	17	17	56	11
		Post	4	7	32	54	4	6	11	56	28	0
S	12. 我認為別人會喜歡我的英文作文	Pre	0	0	43	46	11	6	28	56	11	0
		Post	0	4	18	75	4	6	61	28	6	0
B E L I E F S	13. 我認為學習英文寫作，對其他學科的學習很有幫助	Pre	11	46	43	0	0	28	44	22	0	6
		Post	14	61	14	11	0	17	39	39	6	0
L I E F S	14. 我覺得英文寫作只會增加自己的負擔	Pre	4	4	46	39	7	6	72	17	0	6
		Post	4	14	25	54	4	6	50-	33	11	0
F	15. 我覺得英文寫作很困難	Pre	14	54	29	4	0	0	17	17	50	17
		Post	14	61	18	7	0	0	17	33	44	6
S	16. 我不相信自己能克服英文寫作困難，寫好英文作文	Pre	14	39	43	4	0	6	44	17	28	6
		Post	4	29	14	54	0	6	28	56	11	0
B L O C K I N G	17. 我寫英文作文時，會一直專注地寫，直到完成為止	Pre	4	43	36	18	0	6	44	6	39	6
		Post	0	54	11	36	0	22	39	33	6	0
C K I N G	18. 我會盡可能地克服寫作問題，寫好英文作文	Pre	7	50	39	4	0	6	78	6	11	0
		Post	4	57	14	25	0	17	50	28	6	0
I N G	19. 完成英文作文後，我會耐心且多次地修改	Pre	0	32	57	11	0	17	28	22	28	6
		Post	4	36	18	43	0	17	28	39	17	0
S	20. 用英文下筆寫作時，我會覺得無趣而想中斷	Pre	7	29	32	32	0	6	44	22	17	11
		Post	11	43	21	25	0	0	44	44	11	0

*No of EG =28; No of CG=18; 5= Agree a lot; 4=Agree; 3=No opinions; 2=Disagree; 1=Disagree a lot

willing to express thoughts in English (60% to 80%) in both of the pre-study and the post-study questionnaires, and the findings indicate that both groups have high motivation for English writing before and after treatments. Similarly, about 40% of subjects in both groups agreed to like English writing and most subjects (about 60% to 80%) felt happy and a sense of achievement while acquiring new writing skills in both of the pre-study and the post-study questionnaires. The results imply that both groups have high affections toward English writing before and after the treatments.

Furthermore, Table 5 shows a slight increase of percentages in both of the writing motivation and affection in post-study questionnaires by both groups (especially for the EG). This indicates that the treatments have some degrees of impacts on the subjects' writing motivation and affection in both groups although few significant group differences were found. Based on the data in Table 5, both groups held low values toward their own English writings (less than 10% of the subjects agreed to write well, while 10% to 30% of the subjects liked their own writings and thought others would like their writings before treatments. However, it is interesting to find that the CG indicated higher values toward their own writing after treatments (as the results shown in Item 9 and Item 12) and seemed to have more confidence on their own writing. Furthermore, most subjects (about 70% to 90% of the subjects) agreed that English writing was important. More than half of the subjects in both groups believed that English writing would be beneficial for academic learning in both pre-study and post-study questionnaires. However, as shown in Item 14 and Item 15 of Table 5, the subjects in CG thought English writing led to more academic burdens before and after treatments. In addition, most subjects in EG thought English writing was difficult (more than 70%), most of the subject (more than 40%) in both groups didn't believe they could overcome the writing problems. The findings regarding the subjects' beliefs of English writing indicated that most subjects held more negative attitudes toward English writing and thought it difficult although it was beneficial for academic learning.

As for the writing blockings, the data in Table 5 indicated positive results. Accordingly, more than half of the subjects in both groups agreed to be concentrated in their writings (more than 50%), and to try hard to write well. (60% to 80%). However, the percentages of the subjects in both groups who agreed and disagreed to have multiple revisions in pre-study questionnaire were compatible (about one third). To the researcher's surprises, in the post-study questionnaire, more subjects in the EG disagreed to have multiple revisions (about 40%), while more subjects in the CG agreed to have multiple revisions (about 45%). According to the EG subjects' responses to the writing portfolio project, the possible reason for the increase of percentages for disagreeing to have multiple revisions might be that the subjects in the EG lost patience to conduct multiple revisions in the BLOG portfolio project. In addition, it should be noted that a lot of subjects in both groups chose no opinions in the Item 19, and thus speculations of this item would be limited. Finally, in both of the pre-study and post-study questionnaires, most subjects in EG and in CG (about 50%) agreed that in the initial stage of writing, they would stop writing if they felt bored. The findings above suggested that most subjects in both groups would try hard to write and solve related problems, but certain degrees of resistances to writing were found.

According to Table 6, no significant differences in most items regarding writing motivation, affection, and blocking between the EG and the CG were found. The mean scores of both groups in writing motivation and affection as shown in Table 6 indicate no group differences, except Item 4, and the results suggest that both groups had high motivation and affections for English writings before and after treatments, and the portfolio projects, either paper-based or BLOG-based, did not lead to better motivation or affections toward English writing. Likewise, the means scores in writing blocking show no group differences in most items, except Item 17 regarding focusing on writing. Although a significant difference was found in Item 17 indicating the CG could better focus on English writing, the

Table 6. Group Comparison of the Pre-study and Post-study Questionnaires on Attitudes toward English Writing

	Questions		Mean		<i>t</i>	<i>p</i>
			EG	CG		
M O T I V A T I O N	1. 我對英文寫作很有興趣	Pre	3.18	3.11	.25	.81
		Post	3.04	3.33	-1.14	.26
	2. 我喜歡學習有關英文寫作的新知識	Pre	3.46	3.78	-1.30	.20
		Post	3.43	3.72	-1.04	.30
3. 我想學習並嘗試不同的英文寫作方法	Pre	3.54	3.94	-1.64	.11	
	Post	3.54	3.78	-.86	.39	
4. 我會想用英文寫出我的想法和感受	Pre	3.71	4.22	-2.33	.39	
	Post	3.75	4.00	-1.14	.02*	
A F F E C T I O N	5. 我覺得用英文寫作很有趣	Pre	3.29	3.56	-1.12	.26
		Post	3.07	3.50	-1.77	.27
	6. 我喜歡上英文寫作課	Pre	3.21	3.17	.21	.08
		Post	3.07	3.17	-.41	.84
7. 當我學到新的英文作文技巧時，我覺得很高興	Pre	3.71	4.39	-3.31	.00*	
	Post	4.04	4.22	-.96	.34	
8. 我覺得用英文寫作令我有成就感	Pre	3.75	4.00	-.98	.33	
	Post	3.82	4.22	-1.63	.11	
V A L U E S	9. 我覺得自己的英文作文寫得很好	Pre	1.96	1.89	.41	.69
		Post	1.89	2.44	-2.52	.02*
	10. 我覺得英文寫作很重要	Pre	3.89	4.22	-1.48	.15
		Post	4.11	3.89	1.12	.27
11. 我喜歡自己寫的每篇英文作文	Pre	2.71	2.39	1.23	.23	
	Post	2.54	2.94	-1.64	.11	
12. 我認為別人會喜歡我的英文作文	Pre	2.32	2.28	.21	.84	
	Post	2.21	2.67	-2.43	.02*	
B E L I E F S	13. 我認為學習英文寫作，對其他學科的學習很有幫助	Pre	3.68	3.89	-.85	.40
		Post	3.79	3.67	.47	.64
	14. 我覺得英文寫作只會增加自己的負擔	Pre	2.57	3.72	-4.58	.00*
		Post	2.61	3.50	-3.40	.00*
15. 我覺得英文寫作很困難	Pre	3.79	2.33	5.75	.00*	
	Post	3.82	2.61	4.99	.00*	
16. 我不相信自己能克服英文寫作困難，寫好英文作文	Pre	2.64	3.17	-1.89	.07	
	Post	2.82	3.28	-1.68	.10	
B L O C K I N G	17. 我寫英文作文時，會一直專注地寫，直到完成為止	Pre	3.32	3.06	.91	.37
		Post	3.18	3.78	-2.16	.04*
18. 我會盡可能地克服寫作問題，寫好英文作文	Pre	3.61	3.78	-.80	.43	
	Post	3.39	3.78	-1.45	.15	
19. 完成英文作文後，我會耐心且多次地修改	Pre	3.21	3.22	-.03	.98	
	Post	3.00	3.44	-1.50	.14	
G S	20. 用英文下筆寫作時，我會覺得無趣而想中斷	Pre	3.11	3.17	-.19	.85
		Post	3.39	3.33	.22	.83

*No of EG=28; No of CG=18; $p < .05$

frequency analysis in Table 5 indicated both groups showed low blocking in focusing on writing. Hence, the results shown in Table 6 indicate that the portfolio project did not lead to less reluctance or blocking for the writing revisions although low blockings were found in focusing on writing and overcoming related problems.

In addition, the results in Table 6 indicate that both groups held comparatively low values toward their own writings (as shown in Items 9, 11 and 12), but considered English writing important (Item 10). The data in Table 6 also suggest significant differences in higher values toward their own writing after treatments (as the results shown in Item 9 and Item 12) and the CG seemed to have more confidence on their own writing after treatments.

Finally, no conclusive results were found in writing beliefs between the EG and the CG. Accordingly, no significant differences in writing beliefs between both groups were found in Item 13 (i.e. believing English writing beneficial for academic learning) and Item 16 (i.e. believing oneself able to overcome writing problems). However, a significant group difference in Item 14 (i.e. believing English writing as burden) and Item 15 (i.e. believing English writing difficult) of the pre-study and the post-study questionnaires were suggested. According to the frequency analysis in Table 5, the possible explanation for the differences is that most subjects in the EG believed English writing was difficult, and the portfolio project did not have significant effects to change the belief. On the other hand, a certain amount of the subjects in the EG and the CG (about 50%) chose no opinions, and this contribute to the inconclusive results in the subjects' beliefs toward English writing. Nevertheless, both groups believed English writing beneficial but difficult, and the results indicated that the portfolio project had no limited effects to change the belief held by the subjects in both groups.

In a word, both groups had high motivations and affections toward English writing, but low values toward writing before treatments, and the portfolio project had limited effects on writing motivation, affections, values and blocking of the subjects in this study.

The t-test analyses in Table 7 indicate no significant effects of the portfolio project on all aspects of writing attitudes for the EG, but few significant differences were found in writing values and blocking for the CG. Significant differences were found in Item 7 (i.e. feel happy while acquiring new writing skills) between the pre-study and the post-study questionnaires by the EG.

However, no significant differences were found in other aspects of writing attitude. In addition, Table 7 shows no significant differences between the pre-study and the post-study questionnaires by the CG in most aspects of writing attitudes, except writing values. Confirming the findings in the previous studies (e.g. Chang, 2003), the findings in this study indicate that the paper-based portfolio project had some impacts on helping the subjects for developing confidence in English writing in the sense that the subjects liked their own works better, and they thought others would appreciate their works as well. However, the paper-based portfolio project had no impacts on writing motivation, affections, beliefs, and blocking. In other words, unlike the previous studies suggested, the results of Table 7 indicated that portfolios did not motivate the subjects to like English writing better, to change their beliefs toward English writing, or to help them less reluctant to English writing.

In summary, the findings of this study confirm the positive effects of the portfolio project on the writing achievements as suggested by the previous studies (e.g. Chang & Chang, 2003; Easterwood, 1997), but the portfolio project had limited effects on the writing attitudes of the subjects. The study found a significant difference between the pretest and the posttest by the CG, but no significant difference in that of the EG. A significant difference in writing achievement tests between the EG and the CG was also found. The results imply that the CG outperformed the EG in the English writing achievement tests, and the paper-based English writing portfolio project helped the subjects in this study had better writing performances in English writing achievement tests than the BLOG-based English writing portfolio project did.

Table 7. Paired *t*-test Analysis of the Pre-study and Post-study Questionnaire on Attitudes toward English Writing by the EG and by the CG

Questions	EG (Mean)					CG (Mean)					
	Pre	Post	SD	<i>t</i>	<i>p</i>	Pre	Post	SD	<i>t</i>	<i>p</i>	
M O T I V A T I O N	1. 我對英文寫作很有興趣	3.18	3.04	.71	1.07	.29	3.11	3.33	1.00	-.84	.39
	2. 我喜歡學習有關英文寫作的新知識	3.46	3.43	.88	.21	.83	3.78	3.72	1.11	.21	.83
	3. 我想學習並嘗試不同的英文寫作方法	3.54	3.54	.94	.00	1.00	3.94	3.78	.79	.90	.38
	4. 我會想用英文寫出我的想法和感受	3.71	3.75	.79	-.24	.81	4.22	4.00	1.06	.89	.39
A F F E C T I O N	5. 我覺得用英文寫作很有趣	3.29	3.07	.96	1.19	.25	3.56	3.50	1.26	.19	.85
	6. 我喜歡上英文寫作課	3.21	3.07	.65	1.16	.26	3.17	3.17	1.03	.00	1.00
	7. 當我學到新的英文作文技巧時，我覺得很高興	3.71	4.04	.77	-2.20	.04*	4.39	4.22	1.10	.64	.53
	8. 我覺得用英文寫作令我有成就感	3.75	3.85	.81	-.47	.65	4.00	3.44	1.26	-.81	.41
V A L U E S	9. 我覺得自己的英文作文寫得很好	1.96	1.89	.54	.70	.49	1.89	2.44	.78	-3.01	.01*
	10. 我覺得英文寫作很重要	3.89	4.11	.74	-1.54	.14	4.22	3.89	.84	1.68	.11
	11. 我喜歡自己寫的每篇英文作文	2.71	2.54	.98	.96	.35	2.39	2.94	1.10	-2.15	.05*
	12. 我認為別人會喜歡我的英文作文	2.32	2.21	.57	1.00	.33	2.28	2.67	1.04	-1.59	.13
B E L I E F S	13. 我認為學習英文寫作，對其他學科的學習很有幫助	3.68	3.79	.83	-.68	.50	3.89	3.67	1.00	.94	.36
	14. 我覺得英文寫作只會增加自己的負擔	2.57	2.61	.96	-.20	.85	3.72	3.50	.81	1.17	.26
	15. 我覺得英文寫作很困難	3.79	3.82	.58	-.33	.75	2.33	2.61	1.18	-1.00	.33
	16. 我不相信自己能克服英文寫作困難，寫好英文作文	2.64	2.82	1.02	.93	.36	3.17	3.28	1.02	-.46	.65
B L O C K I N G S	17. 我寫英文作文時，會一直專注地寫，直到完成為止	3.32	3.18	1.01	.75	.46	3.06	3.78	1.23	-2.50	.02*
	18. 我會盡可能地克服寫作問題，寫好英文作文	3.61	3.39	.83	1.36	.18	3.78	3.78	.84	.00	1.00
	19. 完成英文作文後，我會耐心且多次地修改	3.21	3.00	.79	1.44	.16	3.22	3.44	1.06	-.89	.39
	20. 用英文下筆寫作時，我會覺得無趣而想中斷	3.11	3.39	.85	-1.77	.09	3.17	3.33	1.42	-.50	.63

*No of EG = 28; No of CG = 18; $p < .05$

However, both groups showed certain degrees of progresses in posttests, and the result suggests that portfolio-based writing projects helped the subjects in this study have better writing achievements, especially the paper-based writing portfolios.

Unlike the findings in some previous studies, which argued for the positive effects of the portfolios on writing attitudes (e.g. Shoberm, 1996), the writing portfolio project in this study had limited effects on learners' writing attitudes. The study findings are consistent with the conclusions by Starck (1999) that suggested no positive effects of portfolios on writing attitudes. Few significant differences in writing motivation, affections, values, beliefs, and blocking were found after treatments, and few significant group differences were suggested although both groups in this study had high motivation and affections toward English writing.

The Effects of the English Writing Portfolio Projects on Holistic Writing Performances

Table 8 to Table 10 show the t-test analyses of the first drafts and the final drafts of different writing topics involved in this study by the EG and the CG. The results help understand how the portfolio projects affect the subjects' writing performances in a variety of writing tasks.

Table 8. *T-test Analysis of the Holistic Writing Performances by the EG and the CG*

Topic	Draft	Group	No	Score	SD	<i>t</i>	<i>p</i>
BEACH	Draft 1	EG	28	2.82	.86	-2.60	.01*
		CG	18	3.56			
	Final	EG	28	3.04	.79	-3.65	.00*
		CG	18	3.94			
CAMPING	Draft 1	EG	28	2.96	.84	-.56	.58
		CG	18	3.11			
	Final	EG	28	3.21	1.03	-.81	.42
		CG	18	3.44			
PICNIC	Draft 1	EG	28	3.00	.77	-1.50	.14
		CG	18	3.39			
	Final	EG	28	3.21	.99	-5.57	.00*
		CG	18	4.67			
SURPRISE	Draft 1	EG	28	2.86	.85	-1.18	.08
		CG	18	3.28			
	Final	EG	28	3.11	.88	-7.52	.00*
		CG	18	4.78			

*Maximum score=5 ; $p < .05$

According to Table 8, the Mean scores of the first drafts and the final drafts of different topics by the EG were lower than those of the CG although both groups showed some progresses. Table 8 also indicates significant differences between the two groups in writing performances of most topics, except the topic "Camping."

Similarly, Table 9 shows that there are significant differences between the first drafts and the final drafts of the topics "Camping" and "Surprise" by the EG. According to Table 10, significant differences between the first drafts and the final drafts are found in all topics written by the CG. The results imply that the CG outperformed the EG in the English writing performances, and the paper-based English writing portfolio project helped the subjects in this study had better English writing performances holistically than the English writing portfolio project with BLOG did. However, both groups showed certain degrees of progress in final drafts of various topics, and the results suggest that portfolio-based writing projects helped the subjects in this study had better holistic writing performances, especially the paper-based writing portfolios.

Table 9. *T-test Analysis of the Holistic Writing Performances by the EG*

Topic	No	Draft 1	Final	SD	T	p
Beach	28	2.82	3.04	.69	-1.65	.11
Camping	28	2.96	3.21	.65	-2.05	.05*
Picnic	28	3.00	3.21	.63	-1.80	.08
Surprise	28	2.86	3.11	.52	-2.55	.02*

*Maximum score=5 ; $p < .05$ Table 10. *T-test Analysis of the Holistic Writing Performances by the CG*

Topic	No	Draft 1	Final	SD	T	p
Beach	18	3.56	3.94	.70	-2.36	.03*
Camping	18	3.11	3.44	.49	-2.92	.01*
Picnic	18	3.39	4.67	.89	-6.06	.00*
Surprise	18	3.28	4.78	.79	-8.10	.00*

*Maximum score=5 ; $p < .05$

CONCLUSIONS AND IMPLICATIONS

In conclusion, the findings of the study indicate that English writing portfolios, both paper-based and BLOG based ones, had positive effects on junior college students' English writing achievements and holistic writing performances, but had limited impacts on attitudes toward English writing. In consistent with the findings of the related literature, the present study found that English writing portfolios had positive effects on English writing performances, yet the results of the study did not draw a conclusive result regarding the impacts of portfolio-based writing instructions on students' attitudes toward English writing. Beyond the researcher's expectations, the use of BLOG-based English writing portfolios did not lead to significantly more improvements in English writing achievements or holistic writing performances. It is speculated that the familiarities with the functions of BLOG and distractions of students' attentions by computer applications might somehow influence students' English writing performances. Further studies hence could focus on investigating the effects of other forms of electronic portfolio project, involve more subjects and prolong the experimental and instruction time. The findings of the present study and future studies could inspire more teachers and researchers to incorporate portfolios into English writing instructions and to develop innovative and motivating English writing instructions and pedagogy.

REFERENCES

- Bryman, A. & Cramer, D. (1997). *Quantitative Data Analysis with SPSS for Windows*. London: Routledge.
- Carlo, J. E. (1995). *Perspectives in whole language*. London: Allyn & Bacon.
- Chang, C. C. (2003). *Diagnosing English learning for junior high school students by portfolio assessment*. Unpublished Master Thesis. Kaohsiung: National Kaohsiung Normal University.
- Chang, Y. R. & Chang, J. Y. (2003). The effect of multiple-intelligence-based teaching and portfolio assessment on English learning achievement, learning motivation and class climate in English class of junior high school. *Journal of Educational Psychology*, 34 (2), 199-220.

- Chao, T. C. (2000). Authentic assessment in the EFL school context: An MI-based model. *The Proceedings of the 17th Conference on English Teaching and Learning* (pp. 486-496). Taipei: The Crane.
- Chen, Y. M. (1999). A portfolio approach to EFL university writing instruction. *Proceedings of the 16th national conference on English teaching and learning in the Republic of China* (pp. 312-332). National Changhua Normal University.
- _____. (2000). Learning writing as writers: A portfolio writing curriculum for EFL university first-year students. *Proceedings of the 17th Conference on English Teaching and Learning in the Republic of China* (pp. 294-309). Taipei: Crane.
- _____. (2005). Using portfolios at EFL junior high classrooms in Taiwan. *Proceedings of 2004 International Conference on English Language Assessment and Instruction* (pp. 99-127). Taipei: Crane.
- Chiang, T. T. (2002). *Collaboration-based instruction and second language writing: How teachers impact the quality of student work in collaborative writing classroom*. Unpublished Doctoral Dissertation: Alliant International University.
- Dogger, B., Moy, J. & Nogami, M. (2003). A quest for the perfect portfolio. In C. A. Coombe and N. J. Hubley (Eds.), *Assessment practices*. Alexandria, Virginia: Teachers of English to Speakers of Other Languages, Inc.
- Fosnot, C. (1996). *Constructivism: Theory, perspectives and practice*. New York: Teachers College Press.
- Gay, L. R. (1992). *Educational Research Competencies for Analysis and Application*.
- Gungle, B. W. & Taylor, V. (1989). Writing apprehension and second language writers. In D. M. Johnson & D. H. Roen (Eds.), *Richness in writing: Empowering ESL students* (pp. 235-248). New York: Longman.
- Harste, C. J. & Lowe, K. S. (1991). Whole language: Getting the act together. *The Education Digest*, 57(4), 45-49.
- Hsieh, Y. F., Lu, C. Y. & Yeh, H. N. (2000). Implementation of portfolio assessment in a sixth grade EFL classroom. *Proceedings of 2000 National Workshop and Conference on the Application of English Teaching* (pp. 112-139). Taipei: Crane Publishing Co., Ltd.
- Huang, Y.H. (2003). *Four types of portfolio assessment in EFL instruction: Alternative festival portfolio assessment*. Unpublished Master Thesis. National Kaohsiung Normal University.
- Paris, S.G. & Ayres, L. R. (1994). *Becoming reflective students and teachers with portfolios and authentic assessment*. Washington, DC: American Psychological Association.
- Phinney, M. (1991). Word processing and writing apprehension in first and second language writers. *Computers and Composition*, 1(1), 65-82.
- Pierce, L. V. & O'Mally, J. M. (1992). Performance and portfolio assessment for language minority students. *NCBE Program Information Guide Series*, 9.
- Pollari, P. (2000). *This is my portfolio—Portfolios in upper secondary school English studies*. Unpublished Ph. D. Dissertation. ED 450415.
- Song, S. Y. (2002). *The effects of a portfolio project on the English learning of senior high school students*. Unpublished Master Thesis. Kaohsiung: National Kaohsiung Normal University.
- Subrick, D. M. (2003). *Effects of writing portfolio instruction on mean scores of 5th-grade students in Team Approach to Mastery (TAM) classroom*. Doctoral Dissertation: Wilmington College.
- Sun, H.C. (2003). *Effects of the western children's story portfolio assessment on English reading and writing development of junior high school students*. Unpublished Master Thesis. Kaohsiung: National Kaohsiung Normal University.
- Tierney, R. J., Carter, M. A., & Desai, L. E. (1991). *Portfolio assessment in the*

- reading-writing classroom*. Norwood, MA: Christopher-Gordon.
- Wu, M. L. (2003). *SPSS statistical Analysis Applications and Practice—Questionnaire Analysis and Applicative Statistics*. Taipei: Acore Digital Technology.
- 鄒慧英、甘孟龍 (1999)。國語科寫作檔案評量與直接寫作檔案評量之發展。國科會計畫研究報告。NSC88-2413-H024-011。
- 謝永芳 (2000)。檔案式評量於小學六年級英語教學上的應用。未出版碩士論文。台北: 國立台灣師範大學。

Integrating Discourse-based Approach into English Writing Instruction

Hui-ling Lang (郎慧玲)

Ming Chuan University

huilinglang@hotmail.com

Since the early 1980s applied linguists and language teachers have shown a great deal of interests in discourse-based approaches to the analysis of written and spoken discourse. Researchers in writing use genres as a tool for analysing the overall organisation and patterns or distinctive linguistic features within a particular genre (Hyland, 2002, 2003 and 2004; Hyon, 1996; Johns, 2002 and 2003; Paltridge, 2001 and Ramanathan and Kaplan, 2000). Such analyses can be used for teaching planning and learning activities. However, most studies, often corpus-based, provide supportive findings about raising students' awareness of generic structures and explicit features; fewer have shown a case of implementing a discourse-based pedagogical approach in a writing context and reveal the effects through a detailed examination of students' written products. Therefore, this experimental and classroom-based study aims to explore the effects of using a discourse-based approach to English writing with the evidence of students' authentic written work. The instructor adopted a discourse-based approach in the writing class by assisting students to read different types of texts of the target genre, 'description' in this study, and guided students to *analyse* the organisational structures and linguistics features in the model texts. Detailed text analysis of students' writing assignments demonstrates that students grasped the basic structures of the target genre and main linguistic features to compose acceptable descriptive texts. This study offers useful insights and teaching activities for English writing instructors working on the first level in an English writing teaching classroom.

INTRODUCTION

How to select appropriate approaches to teaching writing can be acknowledged as a challenging task for many writing teachers. For example, a writing class with a heavy emphasis on grammar can easily become an additional grammar class for students and, as a result, students may see writing as a grammar practice. Using a discourse-based approach to teaching writing have been raised a great deal of interests for scholars (Hyland, 2002, 2003 and 2004; Johns, 2002 and 2003; Paltridge, 2001 and Ramanathan and Kaplan, 2000). A discourse-based approach to teaching means to utilise the findings of discourse analysis to raise students' awareness of the overall organisations, the target linguistic patterns and features within the target text. Findings of these analyses can be used for teaching activities.

Definitions of discourse analysis are various. McCarthy (1991:5) states that 'Discourse analysis is concerned with the study of the relationship between language and the contexts in which it is used. Gee's concept (2000:92) of discourse analysis involves asking questions about how language is used to construct the aspects of the situation network. Johnstone (2002:5) sees discourse analysis to reveal 'how meaning can be signalled via the arrangement of chunks of information across a series of sentences or via the details of how a conversationalist takes up and responds to what has just been said'. Among them, perhaps Nunan's definition (1993:7) can be seen as a more concrete one. He states 'The analysis of discourse involves the analysis of language in use. The aim to discourse analysis is to interpret the relationship between the patters and meanings and purposes through the discourse'. An important concept of discourse analysts to bear in mind is to search for the

meanings (function) which are revealed through the language (form) in the text.

When analysing a discourse, it is essential to be aware of the salient linguistic features appeared in the text and also its function served in the text. These features can be markers (references) or particular lexical words (Nunan, 1993). This kind of analysis is not restricted to micro-level of the text itself but analysts also need to take notice of the macro-level of the text, i.e. organisation. Organisational moves are also important to be noticed in analysing a discourse. When adopting this approach into a classroom, teachers can use their discourse knowledge to raise students' awareness of salient linguistic features and clear organisation patterns so that students can grasp the basic understanding about how this particular discourse consists of and then produce this type of texts. As Hyland notes (1990:66), 'The difficulties faced by EFL/ESL students whom asked to produce a piece of writing are often due to inadequate understanding of how texts are organized. To facilitate effective writing, therefore, teachers have to familiarize students with the rhetorical structures which are an important part of the meanings of texts'.

While many studies on discourse analysis, often corpus-based, provide supportive findings about raising students' awareness of generic structures and explicit features; fewer have shown a case of implementing a discourse-based pedagogical approach in a writing context and reveal the effects through a detailed examination of students' written products. Therefore, this experimental and classroom-based study aims to explore the effects of using a discourse-based approach to English writing with the evidence of students' authentic written work. The instructor adopted a discourse-based approach in the writing class by assisting students to read different types of texts in the description style and raised students' awareness of salient organisation patterns and linguistics features within the model texts through instructions and guidance provided by the writing instructor.

METHODS

Procedure

Twenty-eight English major freshman students in the Basic Writing class were guided to read the target genre 'description' through a discourse-based approach. The model text¹ chosen to demonstrate this type of genre is presented as follows (Adopted from Folse, Muchmore-Vokoun, Solomon, 2004, p140).

A great living room

My living room may be *small*, but it is *tidy* and *well-organized*. On the right, there is a *wooden* bookcase with *four shelves*. On top of the bookcase is a *small* lamp with a *dark* base and a *matching* lampshade. The first and third shelves are filled with *carefully arranged* books. On the second shelf, there is an *antique* clock with *faded* numbers on its face. The bottom shelf has *a few* newspapers. On the opposite side of the room is an *old* television set with *nothing on top of it*. Between the television and the bookcase is a *large* sofa. *A fat, striped* cat with *long whiskers* is curled up in a ball on the right side of the sofa. Lying to the left of my cat is a *single* sock that *the cat probably brought from another room*. Directly in front of the sofa, there is a *long* coffee table with *short legs*. On the right side lie two magazines. They are *stacked one on top of the other*. Perhaps the most *striking* item in the room is the *beautiful* beach painting above the sofa. This *extraordinary* painting shows a *peaceful* beach scene with *a sailboat on the right, far from the beach*. Although it is a *small*, everything in my living room is its place.

After reading this text together, students were asked to answer a series of questions based on the main features used in the model text. The purpose for this task was to raise students'

¹ The original text is not marked with italics and without underlining.

awareness of the overall structure and linguistic features in the text and students could act as discourse analysts themselves. Questions are listed in the following.

1. What subject does the writer want to describe (a place, a person or an object)? Can you see a picture coming out from the text?
2. How is this text organised? What is the special order the writer follows to describe his/her room? Does the writer describe the room from the bottom to the top or vice versa OR from the right to the left or the opposite?
3. Could you underline all the prepositional phrases to indicate locations?
4. What kind of details does the writer use to describe individual objects in the room?
5. Could you identify and circle all the adjectives used in this text? After doing so, could you categorise them into the five sensory groups, they are, 'sight', 'smell', 'taste', 'hearing' and 'touch'?
6. Does the writer use any similes or metaphors or even exaggeration to describe his/her room?
7. What impression does the writer aim to create for readers?

After this analysis, students were then guided to read a passage given by the instructor in order to compliment some theoretical background for a descriptive text. It is as follows.

Description provides the reader with a picture consisted with words and the writer's task is to create this picture of a specific person, a particular place or the look of a particular object. In order to help readers to visualise this picture, the writer selects key details to develop this description. The specific descriptive words the writer chooses depend on the particular impression which the writer aims at building. This kind of impression is conveyed by the essential qualities of the subject which the writer chooses to describe and also the significance of this subject for the writer. In other words, a descriptive text is not a bland descriptive text but it needs to provide details to demonstrate the meanings, e.g. fondness of this subject for the writer. Of course, descriptive words and phrases are important to effective writing as they make an object concrete for the reader by describing how it looks, sounds, tastes, smells or feels. These sensory details can create a distinct impression or image of that which is described and thus help the reader to visualise the writer's ideas. In terms of organisation, the writer must organise the details according to a logical arrangement and this arrangement is spatial. The writer can organise the text from top to down, back to front, left to right or right to left and he/she then can demonstrate the special relationship of items in space and readers can follow them through the writer's guidance on space. (Adopted from Collin, 2002 and Ruetten, 2003)

Students were then informed to write an assignment in a descriptive style and they could choose any topics as they like.

Texts

Three assignments written by Ann, Cindy and Dolly (anonymous names) described a place, a person and an object respectively were used to demonstrate the positive effects of adopting a discourse-based writing approach.

Analysis

Each sentence in the text was marked in numbers individually for the purpose of analysing the text with convenience. The rationale for this analysis is based on the genetic nature of descriptions (Collin, 2002, Ruetten, 2003 and Folse, Muchmore-Vokoun, Solomon, 2004)

RESULTS AND DISCUSSION

Each student's text is presented first and words in italics indicate adjectives and underlined phrases refer to prepositional phrases.

Text 1 (Description of an object)

My favourite pencil

1. There is a pencil that has been put in my pencil box since I was a junior high school student. 2. It is my favourite, and it is made in Japan. 3. Even though I have bought many other pencils, I still use this pencil when writing. 4. It is a very *normal* pencil. 5. It is covered with *bright yellow* colour on its cover and with *few blurred red* words on it. 6. On the top of it sits a doll and he is a *little* boy. 7. He is wearing a *yellow* hat with a feather tied on it. 8. He has *shiny dark* hair and a pair of *bright, blue,* and *big* eyes. 9. His nose is very different from others because his nose is twice longer than other ordinary people. 10. Under his nose, it is his mouth full of energy, always with a big smile on it. 11. He is wearing a *red* overalls and a *short yellow* T-shirt with a *big, blue* bowknot on it. 12. It looks really fun and naughty. 13. He is wearing a pair of *white* gloves on his *little* hands. 14. What's more interesting is that his *red* shoes on his feet are as big as his head. 15. In addition, he is sitting at the floor, open his legs widely. 16. His appearance is so unique that I can't throw it away. 17. That is why I cherish it so much.

Ann chose to describe a subject, 'her favourite pencil', in her assignment and line 1 immediately signals the history and background of this pencil and her emotional attachment with it in lines 2 and 3. Beginning with a brief introduction to this pencil, Ann highlighted the significance of this pencil to her and provides some qualities about this pencil. Before she initiated to describe this pencil in details, she especially emphasised the ordinariness of this pencil with the use of the first adjective *normal*, which is also the impression she wants to convey to readers. She tried to avoid reader's imagination to visualise it as something peculiar. The first three sentences create a particular impression which Ann aimed at building and she successfully transmitted this impression to her readers. In terms of the spatial order, she started to describe this pencil from its outside features with the use of adjectives such as *bright, yellow, few, blurred* and *red* so that readers can visualise this pencil vividly. She then moved her focus to describe an even more specific subject of this pencil, a little boy on the top of the pencil.

Lines 7 to 15 were used to describe this little boy and she again had a clear structure to focus on where she aimed at emphasising. Her organisation to describe this little boy is from the top to bottom. In line 7, she described the colour of this little boy's hat as *yellow* and this hat is *with a feather* which is also another way to describe a subject without using adjectives. She then described the boy's hair and eyes in line 8 and used more adjectives such as *shiny, dark, bright, blue* and *big*. Comparing with her previous description, she used comparative phrases, 'twice longer than ordinary noses' to highlight the little boy's nose and this use can be seen as a simile. From the prepositional phrase 'under his nose' in line 9, she indicated a clear position for readers to follow her description from the nose to the boy's mouth. She not only used the adjective *big* to describe the mouth but also a metaphor 'full of energy' to show her good skills in using descriptive devices. After she finished describing facial features of the little boy, she switched her focus to the clothing the boy wears. We can see that she had a clear mind about where her focus is; from the natural appearance to the outside wearing. In line 11, she used adjectives to show the size and

colour of the clothes the little boy wears. And in line 12, the adjectives *funny* and *naughty* she used strengthens her emotions towards this particular subject. In line 13, she used *white* to indicate the colour of the gloves and used *little* to signal the boy's hands. In contrast with the adjective *little*, we see *big* when she depicts the boy's shoes and she used the phrase 'as big as' to show her good skills in description. She is not good at using adjectives in colour and sizes skilfully and also adverbs as we can see this use in line 15 'widely' to describe the boy's feet. Lines 16 and 17 serving as concluding sentences, showing Ann's strong attachment with this pencil and these sentences echo the first three sentences in this text. The impression and picture of this pencil is consistent through the text.

In general, we can see in Ann's text that she used numerous adjectives in colour and size to describe this rather small and particular subject 'pencil' and readers can visualise this pencil following her description and advance in creating a picture of this pencil in their minds. All of her descriptive devices are visual and her use of metaphors and similes conveys the vividness of this specific subject and is used as an alternative way to describe subjects without using adjectives.

Text 2 (Description of a person)

Joyce

1. Joyce is pretty and kind. 2. She is one of my friends. 3. She has *coal-black* hair, so she would not like to dye it. 4. On top of her head, she always binds a bowknot. 5. Her eyes are *black* and *bright*. 6. Therefore, when anyone sees her at any time, she looks like as if she is full of energy. 7. Around her neck, it hangs a *crossed* necklace, because she is a Christian. 8. If she encounters any difficulties, she always touches her necklace so that she relieves her mood. 9. Her skin is *smooth*. 10. Sometimes we can not help to touch her skin. 11. She likes wearing perfume, so when she passes me, I often know that the one is her. 12. Her voice is *sweet*. 13. Thus, we like to ask her to sing songs for us. 14. On her fingers, she wears a *heart-shaped* ring. 15. On her wrist, she has a *shiny* watch. 16. When she undoes it, she feels unease. 17. She always smiles. 18. Thus, everyone thinks that she is *kind*. 19. Indeed, when we ask her to help, she always tries her best to solve our problems happily. 20. On her ears, she wears *lovely* earrings. 21. She likes to wear *trendy* clothes. 22. So Joyce is not only *beautiful*, but also a *nice* person to us.

Cindy chose to describe a person 'Joyce' in her assignment. Similar to Ann, Cindy provides the impression she aimed at building in line 1, she used *pretty* and *kind* to create a gentle impression about Joyce and these two adjectives also show Joyce's appearance and character in a general way. Through the text, we can see that Cindy depicted Joyce in details and an interesting technique she adopted to add the vividness is to build an interactive effect. This effect is fulfilled when every time after she described one part of Joyce, she indicated how she and her other friends reacted to Joyce's look or personality. We see the evidence of the use of the pronoun 'we' every time after she finished her description about Joyce and this technique indicates persuasiveness about her description.

In line 3, she emphasised the darkness of Joyce's hair with the use *coal-black* and the fact that Joyce would not dye her hair provided by Cindy strengthens a vivid image of its darkness. Cindy had a clear organisation to describe her friend from top to bottom as she firstly attracted readers' attention to Joyce's head where she binds a bowknot. Then she moved to describe Joyce's eyes as *black* and *bright* and she used a metaphor 'looks like if she is full of energy' to emphasise this feature. She omitted to mention Joyce's nose and mouth but moved to highlight a distinct object on Joyce's neck, a crossed necklace, and this signals Joyce's religion and by saying when Joyce touches this necklace in line 8, Cindy successfully demonstrated the meaning/function of this necklace for Joyce and shaped Joyce's image as

someone religious. In lines 9 and 10, we can see Cindy described how smooth Joyce's skin is and also described her and her friends' reaction to it when they see it. This is the instance where Cindy tried to describe Joyce about the sense of 'touch'.

Lines 11 and 12 show Joyce's habit to wear perfume and Cindy tried to describe Joyce with regard to her smell to readers. We see from line 12 that Cindy described Joyce's voice as *sweet* and this is where Cindy tried to draw on the sound of Joyce's utterances. By saying friends' request to ask her to sing songs in line 13, we can imagine Joyce's voice would be beautiful. Cindy then moved her focus to describe Joyce's hands in line 14, 15 and 16 and chose to describe the small part 'fingers' first and used the adjective *heart-shaped* to describe the ring Joyce wears and then moved to describe Joyce's twist in which she wear a *shiny* watch. Line 16 shows the importance of the watch for Joyce and this phenomenon carefully highlights the important position this watch means for Joyce and perhaps suggests Joyce's punctuality. Lines 17 to 19 are sentences to describe Joyce's character and we can imagine a happy and kind figure in our minds. Although Cindy seems to have a good skill in terms of organisation, an inappropriate special order occurs in line 20 when Cindy described Joyce's ears. A more appropriate position to put it may be *before* Cindy described Joyce's fingers and twist. This is because all of these features are related to Joyce's appearance so it seems to be sensible to group these details together. Line 21 demonstrates that Joyce is a fashionable person as she likes to wear *trendy* clothes. The concluding sentence in line 22 shows Joyce's outside appearance and gentle character and it is the image that Cindy created in line 1 as well.

In general, we can see that Cindy attempted to describe Joyce in two ways- her outside look and her inner personality. Her focus is not only on the appearance but also on building a figure with nice characters and qualities. In term of her strategy to describe Joyce's look, we find that she used many adjectives and prepositional phrases so we can locate and follow her emphasis when placing our attention to the one part of Joyce's body. Through the text, we see indications to what Cindy aimed at creating with sensory details of sight, smell, touch and hearing. Joyce's image can be imprinted in readers' minds.

Text 3 (Description of a place)

My lovely room

1. I have a *small* but *lovely* room. 2. When you first enter my room, you will notice there is a very *big* **mirror on your right side**. 3. **One the back of the mirror**, there is a *wooden striped desk in the middle of my room*. 4. **On my desk**, there are plenty of *colourful* school supplies, and also some *small* but *cute* decorations, such as *small* animal dolls or some *cute* cards from my friends. 5. **Beside my desk**, there is a *high* **closet** that has the same style as my **desk**. 6. **Opposite my closet** is my **bed**. 7. My sheets are *in pink* and have many cartoon characters on them. 8. And my pillows also have the same design as my sheets; the only difference of my bed is my quilt. 9. I also have a pink quilt but there are not any cartoon characters on it. 10. However, it matches with my sheets. 11. **On the right of my bed**, there is a **bookshelf** on the wall and the bookshelf has three shelves. 12. I put all my school textbooks on the first shelf, storybooks on the second and books for references on the third. 13. I have a very *small* and *simple* room, so that is all about my *lovely* room.

Dolly selected to describe a place-her room. Dolly firstly opened this text with two adjectives *small* and *lovely* to indicate the size and atmosphere of her room and we can quickly obtain a basic impression about her room. Through the text, we can notice Dolly's good skills in using one of the cohesive devices, repetition (Nunan, 1993:29), in which she used repeated items to connect the subject which she tried to describe, e.g. mirror, desk,

closet in bold in lines 2,3,4,5,6,11. The repetition of these subjects creates a neat order for readers to follow each object that Dolly described and it builds a smooth flow of the text. In line 2, she began to position readers' attention when they first entered her room and the mirror is the first object which she described as *big*. She also clearly pointed out the place where she put her mirror and in so doing the special order is explicit for readers. In line 3, with the use of prepositional phrase 'on the back of the mirror', we can follow to see the next object easily, a wooden and striped 'desk' and she also made it easier for readers to find the desk with the use of another prepositional phrase 'in the middle of the room'. As her focus becomes the object 'desk', she described these miscellaneous items she put on the desk and she used adjectives *small* and *cute* to describe them. The repetition of her use of small and cute in line 4 suggests that Dolly may not have a wider range of vocabulary in describing objects. In line 5, she used the desk as a point to introduce the next object which she was going to describe, a *high* closet matching with the style of the desk.

Aforementioned, Dolly was good at using cohesive devices and prepositional phrases in the text, so we can find the location of her bed through the prepositional phrase 'opposite the my closet' in line 6. However, it should be noted that there seems to be disconnection weakening the flow of the text when Dolly described the items she put on the bed such as sheets, pillows and quilt, she did not indicate the place in which those subjects were placed clearly. A phrase 'on the bed' seems to be appropriately used here so line 7 would be 'On my bed, my sheets...' and the flow of the text would be then continued. She described her sheets, pillows and quilt from lines 7 to 10. In lines 8 to 10, she used the technique of comparison and contrast to convey the picture of those objects in terms of colour. In line 11, she introduced the bookshelf in her room and used the prepositional phrase 'on the right side of my bed' to make it explicit for readers to *see* where the bookshelf is and also indicated more specifically about where the bookshelf is with the use of the other prepositional phrase 'on the wall'. She created a neat organisation in line 12 as we can see very clearly about what objects she put in each shelf. The concluding sentence in line 13 returns back to the impression she revealed in line 1 and closes a well-structured text with details.

CONCLUSION

This study demonstrates the effect of using a discourse-based approach to teaching writing in the freshman English writing class through the analysis of students' texts. Students produced well-structured text through the instructor's explicit guidance and awareness-raising activity about the genetic nature within the target genre. Explicit instructions to the organisation patterns and linguistic features in the text play a significant role when teachers request students to produce a particular genre. Furnishing students with explicit knowledge of target language discourse forms is necessary. The more explicit we can make out knowledge of particular genres, the clearer we can be when asking out students to write. Writing should be a teachable skill and an indispensable area of this teaching is to increase the visibility of what is to be learnt.

REFERENCES

- Colin, M. (2002) *Patterns plus*, Seventh edition, Houghton Mifflin Company,
 Folse, K., Muchmore-Vokoun, A, and Solmon, E (2004) *Great paragraphs*, second edition,
 Houghton Mifflin Company, Boston, New York
 Gee, J. (1999) *An introduction to discourse analysis*, Routledge, London and New York
 Halliday, M. (1985) *An introduction to functional grammar*, Edward Arnold, London
 Hyland, K. (1990) A genre description of the argumentative essay, *RELC Journal*, 21,1,
 66-78.
 Hyland, K (2002) Genre, language, context and literacy, *Annual Review of Applied*

Linguistics, 22, 113-135.

- Hyland, K (2003) Genre-based pedagogies: a social response to process, *Journal of Second Language Writing*, 12, 17-19
- Hoey, M. (2001) *Textual interaction, an introduction to written discourse analysis*, Routledge, London and New York
- Johns, A. M. (2002) *Genre in the classroom: multiple perceptive*, Mahawah, NJ: Lawrence, Erlbaum
- Johns, A. M. (2003) *Genre and ESL/EFL composition instruction*. In B. Kroll (Ed) Exploring the dynamics of second language writing, 195-217, Cambridge University Press
- Johnstone, B. (2002) *Discourse analysis*, Blackwell Publishers
- McCarthy, M. (1999) *Discourse analysis for language teachers*, Cambridge university Press
- Nunan, D. (1993) *Introducing discourse analysis*, Penguin Books
- Partridge, B (2001) *Genre and the language learning classroom*, Ann Arbor, MI: The University of Michigan Press
- Ramanathan, V. and Kaplan, R.B. (2000) Genres, authors, discourse communities: theory and application for L1 and L2 writing instructors, *Journal of Second Language Writing*, 9, 171-191
- Ruetten, M. (2003) *Developing composition skills*, Thomson and Heinle
- Swales, J. (1981). *Aspects of articles introductions*. Birmingham, The University of Aston, Language Studies Unit.
- Swales, J. (1990). *Genre analysis*. Cambridge, Cambridge University Press.

The Advantages of Task-Based Learning Pedagogy in Macao

Ines Lau (劉曼玲)

Macao University of Science and Technology
mllau@must.edu.mo

The goal of teaching is to foster effective learning. In order to achieve this goal and to bridge the gap between teaching and learning in the EFL classroom, SLA researchers in the past decades have become increasingly interested in pedagogical conceptions, such as Learner-centeredness, Communicative Language Teaching (CLT), and Task-based Learning (TBL). In particular, research into task-based learning pedagogy has recently had an important influence on the field of English language teaching in recent years in the West because this pedagogy is based on a basic principle of communicative approaches and helps learners to develop integrated competence in reading, writing, speaking and listening through interactive tasks. However, there have been few studies into task-based learning in the EFL classroom in Macao. Thus, this exploratory study to be reported in this paper examined whether or not task-based learning pedagogy could be beneficial to tertiary learners in terms of their motivation for English learning. Based on the philosophy of constructivism, the study aimed to investigate their motivation in English language learning through an action research carried out before, during and after the implementation of a specially designed program of task-based English teaching. 24 undergraduate learners were investigated by quantitative and qualitative instruments before, during and after the 15-week task-based EFL program. Both quantitative and qualitative findings reflected that the learners' motivation, particularly intrinsic motivation for learning English, were increased by the task-based EFL program, thus convincingly demonstrating that task-based learning pedagogy is beneficial to tertiary learners of English.

INTRODUCTION

The study to be reported in this paper was inspired by Allwright's (1984) question 'Why don't learners learn what teachers teach?'. In Macao, teacher-centered approaches have often been used in traditional EFL classroom. Consequently, a big gap between teaching and learning always exists. One of the main problems in learning English in Macao is that students seem very passive and sluggish in learning.

In order to solve this problem, a new pedagogy based on the concepts of learner-centeredness and communicative language teaching (CLT) was used in this context. The aim of the study was to examine whether this pedagogy can motivate learners in English language learning. Therefore, it focused on the investigation into the correlation between learners' motivation and a program designed by this pedagogy.

BACKGROUND

Many problems occur in English language teaching and learning in Macao. Teachers comment on students' reluctance to learn English, while students complain about the boredom of learning English. English learning, in this sense, is not satisfactory. Students are very passive in learning and unable to communicate with others in English in their daily life, even though they have been taught English as a compulsory subject in schools for more than ten years. Traditionally, from a teacher's perspective, students' sluggishness could be diagnosed as a symptom of low achievement, low competence and low motivation in English learning. However, rather than murmuring 'why don't learners learn what teachers teach?', why do English teachers not adopt a different teaching method to help students in English

learning?

Recent research makes an attempt to narrow the gap between teaching and learning English by advocating a new pedagogical concept – communicative language teaching (CLT) into EFL classroom (Littlewood, 1981; Nunan, 1988, 1999; Richards and Rodgers, 2001; Liao, 2004). From the literature (e.g., Ellis, 2003; Nunan, 2004; Van den Branden, 2006; Willis and Willis 2007), I realized that Task-based Learning (TBL), which reinforces the concept of CLT, would be beneficial to my students. Since it is considered to be a new teaching concept to Macao, where the EFL teaching environments have mainly focused on traditional teaching methods, I believe some reexamination of effective English pedagogies in Macao is necessary in order to enhance English learning.

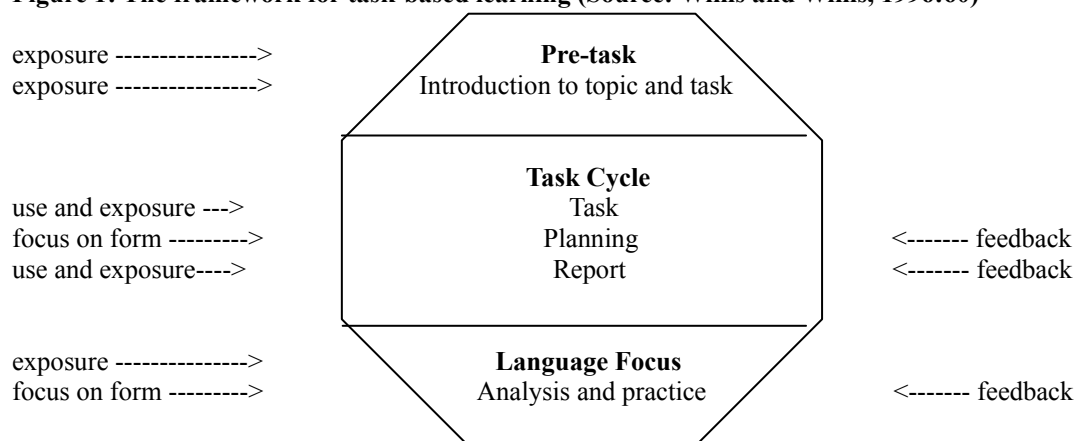
LITERATURE REVIEW

Task-based Learning (TBL)

A task-based learning approach is the strong version of CLT that starts from a theory of language as communication. It aims to stimulate the development of language through use (i.e., using English to learn it). Therefore, it involves the use of tasks to focus learners on constructing and internalizing the knowledge that develops the ability to use language to get things done in a real-life context. As Willis (1996:1) states, ‘task-based learning combines the best insights from communicative language teaching with an organized focus on language form’ (Willis, 1996).

The framework of the task-based learning

In contrast to the traditional teaching approach that focuses on a particular form, TBL provides a holistic language experience where learners naturally integrate four skills through tasks (Willis, 1996). In a framework for task-based learning, Willis (1996) proposes with rationale, the basic procedures of three phases: Pre-task, Task Cycle and Language Focus, as shown below:

Figure 1: The framework for task-based learning (Source: Willis and Willis, 1996:60)

At the Pre-task phase, the teacher introduces and defines the topic to the class, highlights useful topic-related words and phrases but would not pre-teach new structures, and ensures the learners understand task instructions. This initial stage gives useful exposure to help the learners to recall relevant words and phrases and to recognize new ones.

Then, the Task Cycle, which consists of three components: task, planning and report, is central to the framework. First, at the task stage, the learners are encouraged to do the task in pairs or small groups in order to use whatever language they can recall to fulfill the task outcome. Based on the principle that 'output' (i.e., use of language) is likely to help stimulate 'intake' (i.e., acquisition of new forms) (Willis 1996), this stage which emphasizes on spontaneous, exploratory talk and confidence building is vital to language acquisition for the learners. Second, at the planning stage, the learners are encouraged to prepare a report to the whole class (orally or in writing) how they did the task, and what they decided or discovered. It is at the stage that emphasizes on clarity, organization and accuracy, thus a focus on form is natural and teacher advice and correction is likely to be of most use (Willis and Willis, 1996). Third, at the report stage, the learners listen to others doing the same task and compare the ways in which they did the task themselves. This stage gives a natural stimulus to improve their language. It encourages them to think about the form and meaning; accuracy and fluency when they present their reports in the target language. It also provides them with useful exposure in which they hear or read what other learners have done in the same task (Willis, 1996; Willis and Willis, 1996).

The last phase in the framework, Language Focus, encourages the learners to focus their attention to forms of the language which they have already processed for meaning during the task cycle. It consists of two components: analysis and practice. At the stage of analysis, the learners do consciousness-raising activities to identify and process specific language features from the task text or transcript. By the process of identification and generalization, the learners' language knowledge is constructed and the specific language features become part of their internalized language system. After that, at the practice stage, practice activities that are based on the language analysis work are conducted. This stage builds learners' confidence and helps them to systematize their knowledge and broaden their understanding, so that they can carry on learning outside the classroom and after their language course (Willis, 1996; Willis and Willis, 1996).

The rationale of task-based learning

According to Willis (1996:11), 'all good language learners take full advantage of their exposure to the target language in use'. In order for anyone to learn a language with reasonable efficiency, four conditions (three are essential and one is desirable) are concerned to be met. The three essential conditions are exposure, use, and motivation; while the

desirable condition is instruction.

In the framework of TBL, exposure and input consciously and subconsciously come from the pre-task and the analysis stages. This leads the learners to pay attention to useful language features in what they hear and read, and to process the input more analytically. Opportunities are given to the learners to experiment to make use of the target language, to test hypotheses, what they want to say and express what they feel or think through interaction for a task outcome. The planning stage in a TBL framework gives the learners chances to recall the useful language features and the support they need before they perform in public. Then, the report stage in the framework offers them the challenge to refine their performance for a wider audience. Success in achieving the goals of task stimulates the learners' motivation. When they are satisfied with what they have achieved, through their own individual effort, they are more likely to participate the next time. This satisfaction leads them to seek out opportunities for exposure to and use of the target language outside the classroom. In order to improve language accuracy, the learners need chances to reflect on language and to try to systematize what they know. Hence, there is the fourth condition: instruction, which focuses on language form, and is highly desirable. In the TBL framework, there is a natural focus on form when the learners rehearse oral reports or organize written reports for a public presentation. Also, the consciousness-raising activities at the analysis stage provide them chances to identify and systematize specific language features and strive for language accuracy (Willis, 1996; Willis and Willis, 1996).

Motivation theories in L2

No one denies that learner behavior is complex, and motivation is one of the key factors that lead to learning success. However, it is not easy to fully explain the multifaceted nature of motivation. In spite of the wide conceptual differences, Dörnyei (2001a) claims that all the motivation theories reach a consensus that motivation determines three aspects of human behavior: the *choice* of a particular action, the *persistence* with it, and the *effort* expended on it. In other words, motivation explains *why* people decide to do something, *how hard* they are going to pursue it and *how long* they are willing to sustain the activity (Dörnyei, 2001a).

Dörnyei's framework of L2 motivation

Dörnyei's framework of L2 motivation (1994), which specifically focused on motivation from a classroom perspective, attempted to synthesize various lines of motivation research. It categorized an extensive list of motivational components into three main dimensions: (1) language level; (2) learner level; and (3) learning situation level.

The most elaborate part of the framework which is related to this study is the third level – the learning situation level. This level is associated with situation-specific motives, such as intrinsic and extrinsic motives and motivational conditions of L2 learning within a classroom setting. It contains course-specific, teacher-specific, and group-specific motivational components.

Course-specific motivational components related to the syllabus, the teaching materials, the teaching methods and the learning tasks. According to Keller's (1983) education-oriented theory of motivation, the four *Course-specific motivational components* are: interest, relevance, expectancy, and satisfaction (Dörnyei, 2001b; Julkunen, 2001). *Teacher-specific motivational components* concern the motivational impact of the teacher's personality, behavior and teaching style. Finally, *group-specific motivational components* are related to the various characteristics of the learner group such as goal-orientedness, group norms, cohesiveness, and classroom goal structure (competitive, cooperative or individualistic) (Dörnyei, 1994, 1998, 2001a, 2001b, 2003).

Williams and Burden's social constructivist model

Another elaborate framework of motivational components which related to this study is Williams and Burden's (1997) social constructivist model. With the perspective of social

constructivism, Williams and Burden (1997) assert that L2 motivation is internally or externally affected through social interaction. As Williams and Burden (1997:120) mention,

A constructivist view of motivation centers around the premise that each individual is motivated differently. People will make their own sense of the various external influences that surround them in ways that are personal to them, and they will act on their internal disposition and use the personal attributes in unique ways.... However, an individual's motivation is also subject to social and contextual influences. These will include the whole culture and context and the social situation, as well as significant other people and the individual's interactions with these people (Williams and Burden, 1997).

Williams and Burden (1997) claim that why learners are motivated to learn L2 is because their cognitive and emotional arousal is triggered by two main influences: internal or external causes. Internal causes come from inside the learner, such as interest in L2 or a wish to succeed. On the other hand, external causes come from the influences of other people or social context, such as teachers, peers or societal expectations.

Dörnyei and Ottó's process model of L2 motivation

With a dynamic view of L2 motivation, Dörnyei and Ottó (1998) try to account for the changes of L2 motivation over time by breaking down several discrete temporal segments in motivational process. They claim that motivation undergoes a cycle that has at least three main stages: *pre-actional stage* (choice motivation), *action stage* (executive motivation) and *post-actional stage* (motivational retrospection).

The *pre-actional stage* is corresponding to 'choice motivation' that learners are involved in the learning process in that their motivation is generated from their initial wishes to action. In this stage, learners transform their wishes, hopes and desires into goals, then into intentions, and eventually lead to action (Dörnyei and Ottó, 1998).

The *action stage* is corresponding to 'executive motivation'. Once learners launch their actions, their generated motivation needs to be actively maintained and protected. In this stage, whether learners choose to carry on or terminate their particular action is depended on their action control which is affected by many motivational factors, such as quality of learning experience, peers influence, and goal structure (e.g., competitive, individualistic, or cooperative) (Dörnyei and Ottó, 1998).

The third stage in Dörnyei and Ottó's (1998) process model is *post-actional stage*, which is corresponding to motivational retrospection. After completing or terminating the task, learners critically evaluate the process and form causal attributions in their retrospection. In this case, the post-actional motivation process evolves into a pre-actional stage and the cycle begins in a new way again (Dörnyei and Ottó, 1998).

The motivation models for this study

All in all, L2 motivation is a complex and multi-dimensional construct. Since this study focused on investigating the correlation between learners' motivation and a task-based learning approach, the third level of Dörnyei's framework of L2 motivation (1994), which specifically focuses on course-specific, teacher-specific and group-specific motivational components from a classroom perspective, are best related to this study. Based on this model, this study took a precise look at L2 motivations with the perspective of social constructivism. Williams and Burden's (1997) social constructivist model which focuses on intrinsic/extrinsic motivations through social interaction is the best model as the research instrument of L2 motivation for the study. As discussed above that learners' L2 motivation is not stable and constant but rather dynamic and with ongoing changes over a period of time, Dörnyei and Ottó's (1998) model with three main stages (i.e., *pre-actional stage*, *action stage*

and *post-actional stage*) soundly provides the best conditions to examine the students' ongoing changes of L2 motivation through the task-based EFL programme.

The objectives of the study

Having known the ineffectiveness of English learning in traditional EFL classrooms, I intended to find out what teachers can do to help students develop their motivation for English learning. For this reason, I combined theories of L2 motivation and TBL with practice in an action research, investigating the correlation between a task-based learning pedagogy and learners' motivation for English learning. This study focused on these questions:

1. What changes in tertiary learners' motivation through a task-based learning program are apparent and why?
2. How does TBL positively affect tertiary learners' motivation in English language learning?

THE STUDY

Research setting

The research setting of this study was Macao University of Science and Technology (Universidade de Ciência e Tecnologia de Macau). Established in March 2000, it is one of the six private institutions of higher education in Macao. All the students in the university are ethnic Chinese but mainly divided into two groups: local and non-local. In the academic year 2006/2007 which the research was conducted, a total of 8334 students were studying in the university. Of these, nearly one-third were local and two-thirds were non-local students from various regions of China.

With this special background, it is a good opportunity to undertake a research in this setting because there has not been any research related to the correlation between TBL and Chinese tertiary learners' motivation for English language learning.

The program under study

A task-based EFL program which was set up as an elective program was designed for this study. It was a program which was weekly taught for 100 minutes during a semester of 15 weeks. The program was based on Willis's (1996) framework of task-based learning, consisting three phases: pre-task, task-cycle and language focus. It was a general English program with a topic-based syllabus, which integrated the four skills. Various tasks and authentic materials which related to the weekly topics were used. It encouraged collaborative work and enhanced learners' interpersonal skills by being involved in the process of the tasks in different kinds of groupings: individuals, pairs, small groups and a whole class.

Subjects

The 24 participants in the research were Chinese undergraduate students who attended the elective program that was based on a task-based learning pedagogy. The first 24 students who had registered onto the program on line were directly given an offer to enroll in the program and chosen for this research, subject to their agreement. They all were young adults, ranging in the age from 18 to 21, from different faculties. Only one of them was local and the others were from different regions of mainland China. They had learned English for at least eight years in traditional EFL classrooms and their English was at an intermediate level.

Procedure of the study

The study was undertaken during a 15-week semester in the academic year 2006/2007. In the first week, the 24 registered students were given an introduction to the task-based EFL program and an explanation about the aims of the research and their requirements in the research. After they understood their rights and involvements, they were given a booklet for

data recording. After that, each of them was given a Motivation Questionnaire as pre-treatment investigation. In the second week, they were required to record their high or low points in the ‘motivation’ graph with reasons and keep weekly diaries with prompts in the booklet during the research. By the end of the program, the same Motivation Questionnaire was used again as post-treatment. Then, each of the students was individually arranged to have a semi-structured interview in the final week. In addition, from the beginning to the end of the program, the teacher reflected on their interaction in tasks and keep records in field notes.

Research methods

With specific reference to the research into learners’ motivation in the past decades, researchers have largely used quantitative approaches to collect and analyze data with standardized instruments (Dörnyei, 2001b). One of the advantages is that it is relatively easy to categorize the data into a set of findings. However, the outcomes derived from this method are too general, as Dörnyei (2001b:193) argues that ‘it is impossible to do justice to the subjective variety of an individual life’ (Dörnyei, 2001b). Therefore, Dörnyei (2001a) asserts that qualitative research should be used more because it is more sensitive than quantitative research in covering the motivational dynamics involved in the L2 learning process (Dörnyei, 2001a).

For this reason, even though this action research made use of both quantitative and qualitative methods, in order to provide broad and in-depth insights into the nature of the changes in learners’ motivation, how and why they occurred, only qualitative methods are presented in this paper. Following Grotjahn’s (1987) research paradigms and Nunan’s (1992) distinctions among pre-experiments, quasi-experiments, and true experiments (Figure 3), this research was pre-experimental-qualitative-interpretive. It utilized a pre-experiment but yields qualitative data, which were analyzed interpretively.

Figure 3: Contrasting pre-experiments, quasi-experiments, and true experiments: (Source: Nunan 1992: 41)

Type	Characteristics
Pre-experiment	May have pre- and post-treatment tests, but lacks a control group.
Quasi-experiment	Has both pre- and post-tests and experimental and control groups, but no random assignment of subjects.
True experiment	Has both pre- and post-tests, experimental and control groups, and random assignment of subjects.

Since this research was pre-experimental, it was not like traditional experimental research which set up a control group and an experimental group and generalizes the data with statistics. It focuses on controlling dependent variables (i.e., learner personal constructs such as motivation) pre- and post- an independent variable (i.e., a task-based EFL program) in order to study the causality of the dependent variables and the independent variable. Therefore, it measures the dependent variables with quantitative method such as questionnaires in order to survey the differences of the dependent variable before and after the program. Besides, it investigated these dependent variables during the program with qualitative methods such as weekly records of the change of the dependent variables (i.e., ‘motivation’ graphs and notes), field-notes from observation, and learner diaries without attempting to interfere.

A 'motivation' graph and notes

A weekly record of students' high or low points of 'motivation' (i.e., how they generally felt about their English learning at specific time and their reflection on these) towards the tasks and their reflective comments on TBL can provide significant evidence as to whether and how the students' motivation was affected during the task-based EFL program. This 'motivation' graph was adapted from Ellis and Sinclair (1989). Students were required to mark their points of 'motivation' on a graph in the booklet from the range of -5 to +5 and explain the reasons either in Chinese or in English during the program.

Since the 'motivation' graph notes were completed by students without any prompts, the data provide further interesting insights into how and when the students changed their motivation throughout the program. Accordingly, the data collected from the 'motivation' graph notes can be analyzed and triangulated with the data derived from their weekly diaries and follow-up interviews. That makes the research more valid and reliable.

Weekly diaries

Diary entries are beneficial to this research because they are reflectively recorded learning experiences over a period of time from learners' first-person observations. In this research, they were encouraged to weekly record their thoughts and feelings about the task-based EFL program with the structured diary prompts rather than with free-form in the booklet. Also, they were encouraged to respond to the prompts freely in either their native language or target language nine times over the period of the program. They had the opportunity to respond both positively and negatively to these prompts.

The prompts in the booklet were related to regular questions in ordinary weeks and specific questions in review weeks. Every two ordinary weeks was followed by one review week. The specific questions in the first review week were used for looking back generally and specifically at what the students had thought about the tasks and the program. The specific questions in the second review week were used for looking forwards to what the students wanted to do in class. The specific questions in the third review week were used for looking back at general things. Some of the specific questions in the first and the second review weeks referred to task motivation, and were adapted from Cohen & Dörnyei (2001). Other specific questions about students' opinions on using tasks in English learning were asked again as triangulation in the follow-up interviews.

Field-notes

Another significant source of data collected for later qualitative analysis during the task-based EFL program was the field-notes from my observation. The field-notes which were immediately kept after class mainly focused on general comments, the students' motivation as made evident in the class, their interaction in tasks, the performance of an individual student, ideas or problems, and things that worked or did not work well from the class. The data collected from field-notes were analyzed and triangulated with the data derived from the students' weekly diaries and follow-up interviews. That makes the research more valid and reliable. This kind of information recorded from the researcher's observation and reflection on professional practice is helpful for the study and for professional development.

Follow-up interviews

At the end of the task-based EFL program, each of the students was individually given a semi-structured interview. The main advantage of a semi-structured interview is that it is more flexible. Due to its flexibility, it has been popularly used in an interpretive research tradition (Nunan, 1992). Accordingly, an interview schedule was designed with prompts to investigate students' motivation towards tasks in the program. The interview mainly focused on these areas: (1) their attitudes towards English learning after the task-based EFL program; (2) their comments on task-based language learning; (3) whether and how their

motivation in English learning was affected by the program; (4) the reasons why or why not the program affected their English learning motivation; (5) what factors affected their motivation most over the period of the program; (6) whether the program improved their English learning proficiency; (7) what skills had or had not been improved and the reasons; (8) their task preferences and the reasons; (9) whether or not the program satisfied their initial expectations of English learning; and (10) suggestions for improving the program.

Data analysis

The data were analyzed with quantitative and qualitative methods. The quantitative data, collected from the Motivation Questionnaires, were analyzed with the statistical package SPSS for Windows. In order to explore how and why the changes occurred during the task-based EFL program, the qualitative data collected during and after the program using instruments such as the 'motivation' graphs and notes, learners' diaries, field-notes and follow-up interviews were analyzed with NVIVO.

RESULTS AND DISCUSSION

Quantitative data provided evidence of the apparent changes of the students' motivation between pre- and post- the program as base-line data; whereas, qualitative data provide broad and in-depth insights into the nature of the changes. Accordingly, the section 6.1 presents the insights into the nature of changes from while-program data; while the section 6.2 presents the insights from the post-program data, respectively.

Evidence of changes from while-program qualitative data

The qualitative data on the changes of the student's motivation were collected during the program through the 'motivation' graph notes, weekly learners' diaries, and researcher's field-notes, as described in 5.5.1 to 5.5.3. The students had the opportunity to log both negative and positive comments, but the data in the paper refer to positive mentions. In order to explore what changes occurred during the program, the qualitative data were scrutinized for relevant evidence.

Change 1: Motivation for learning English was generally increased.

With regard to motivation in English learning, the results from the qualitative data indicate that, during the program, the students became generally more motivated in English learning. A total of 163 items (mentioned by all the 24 students) relating positively to motivation were found to exist in the qualitative data. Of these, 122 items were found in their weekly diaries, 31 in their 'motivation' graph notes and 10 in the researcher's field notes, as categorized below:

- Intrinsic motivation (A total of 136 items: 101 in diaries; 26 in 'motivation' graph notes; and 9 in field notes)
- Extrinsic motivation (A total of 16 items: 13 in diaries; and 3 in 'motivation' graph notes)
- Instrumental motivation (A total of 11 items: 8 in diaries; 2 in 'motivation' graph notes; and 1 in field notes)

The results provide strong evidence that the task-based EFL program increased the students' overall motivation. The qualitative data findings provide in-depth insights that the students' motivation was complex and multi-faceted.

Change 2: They tended to agree more that the process of learning English was fun and interesting.

The results from the qualitative data show that, during the program, the students felt that learning English through tasks was fun and interesting experience. A total of 418 items (mentioned by all the 24 students) relating to the pleasantness of English learning experience

in the program. Of these, 309 items were found to exist in the diaries, 96 items were found in 'motivation' graph notes and 13 items in field notes. The high frequency of mentions provides strong evidence that the task-based program positively affected the students' intrinsic motivation in English learning. For example:

Learner Diaries\Week 11-Q3\Case 07

Q3: Did the tasks motivate you in your English learning? Why?

Yes, the tasks did motivate me in my English learning and the reason was quite simple, because those tasks were really useful and I had fun with it. (original)

'Motivation' graph notes\Week 14\Case 05

I really enjoyed the task! Even though I always go out with my friends, it was my first time to speak English with strangers outside the classroom. So it was challenging. And I was eager to test my English ability. In the task, my friends and I pretended to be tourists and we talked English with strangers. That was a new experience and we had fun with it! Actually, I've found that speaking English in real life was not as difficult as I thought. I just spoke it naturally. This experience was very helpful for my English learning. I hope that I can continue to use English in daily life, to make friends with others and to expand my world. Learning English through tasks, I feel great! (translated)

In my field notes, I noted that the students were increasingly motivated in learning English through tasks, as follows:

Field notes\Week 06

After the oral reports, I showed them the correct CV template on board. Some of them actively asked me why the aspects should be arranged in this order. I was very happy to see that they became fully involved and motivated in learning. When I was explaining to them, they naturally and autonomously wrote down the notes. I think their motivation came from their interests in the topic and in the task, their curiosity, and their desire for knowledge. Besides, I think it was also because of the joyful classroom atmosphere and supportive group cooperation. (original)

Change 3: They were motivated because the tasks were relevant to them.

The results from the qualitative data indicate that the students' motivation was affected by the tasks. When the students found that the tasks were relevant to their needs, interests, and life experience, they became motivated in their English learning. A total of 346 items (mentioned by 24 students) positively relating to the relevance of tasks were found to exist in the while-program data. Of these, 255 items were found in the diaries, 85 items were found in 'motivation' graph notes, and 6 items in field notes. Two of their mentions are shown as examples:

Learner Diaries\Week 08-Q3\Case 09

Q3: What factors motivated you most in the tasks?

When the topics were practical and relevant to my life, my motivation became strong. When I realized that the topics were important and helpful for my life, I was eager to learn and I willingly learnt harder. And then, after mastering them, I felt happy and satisfied. (original)

'Motivation' graph notes\Week 14\Case 07

The outdoor task is just what I want because we can finally put our English skills into a real-life practice. This is why I chose this class, because I want to interact with people in English in a real-life situation. Although this outdoor task may be a big challenge for some of us, I still think it is worthwhile, no matter how challenging it is. (original)

Also, in my field notes, I noted that the students were very excited when they knew that they were going to do an outdoor task in real-life situation. They seemed very happy to have a chance to practice their English outside the classroom:

Field-notes\Week 12

At the end of the class, I had prepared 10 different outdoor tasks for the students. I explained to them that they were required to work together in pairs, ask for directions of a particular place in a particular tourist centre and do a particular thing by following some instructions. Besides, when they had to do the task, they could only speak English. They need to take pictures and recordings as well. Then, they should prepare everything well with PowerPoint for the oral presentation next week. After the explanation, each pair came to the front to draw a task. I can still remember their laughs and screams when they were drawing the task! Some pairs especially burst into laughter when the task that they drew required them to find a particular restaurant from a particular tourist centre and then order a meal there. I think the reason why they love the outdoor tasks was that, in the past, they had not been given such a chance to speak English outside the classroom; but now, that they were given a chance to practice their English in tasks which are relevant to their life. Therefore, they are eager to put their knowledge into practice in a real-life situation and to have fun in speaking English! (original)

Change 4: They were motivated because of the collaborative nature of task-based learning.

Apart from the relevance of the tasks, the other motivational factor found in the qualitative data was the collaborative nature of task-based learning. A total of 200 items (mentioned by 24 students) positively relating to collaborative learning were found to exist in the while-program data. Of these, 181 items were found to exist in the diaries, 17 items were found in 'motivation' graph notes, and 2 items in field notes. Two of their mentions are extracted below:

Learner Diaries\Week 08-Q2\Case 14

Q2: How much do you like learning / working together with your peer students in the tasks? Why?

I like team work very much. We not only have fun with each other, but also learn from each other. Besides, working together can cultivate our morale. We feel like we are in the same boat so that we would achieve the goals efficiently and effectively. (original)

'Motivation' graph notes\Week 10\Case 05

I am so happy to do the task with my partners! When we were doing the PowerPoint together for the presentation, all of us were very happy to see our beautiful outcome after our endeavor. The feeling was great! Furthermore, in today's class, I am happy that I could also learn a lot from other groups' presentations. That's great! (translated)

Also, in my field notes, I noted that the students were motivated by task collaboration and became actively involved in their learning, as follows:

Field-notes\Week 12

Today we continued the topic 'travel'. At the pre-task – Holiday Maze, they were organized into seven groups of three. They needed to discuss in English which of the three options they compromised to choose for moving to the next step. I noticed that they all were highly involved in their discussion. Sometimes when their partners did not agree with them, they would try their best to convince them with reasons. They had a lot of fun doing it! (original)

Change 5: They tended to set goals for English learning.

The results from the qualitative data indicate that the students had a greater tendency to set their own goals for English learning. A total of 46 items (mentioned by 19 students) relating to goal setting were found to exist in the data. Of these, 41 were found in the diaries, 3 items in 'motivation' graph notes and 1 item in field notes. For example:

Learner Diaries\Week 11-Q3\Case 20

Q3: Did the tasks motivate you in your English learning? Why?

Yep! I was eager to enlarge my vocabulary. Then, speaking English like a native speaker is also my goal. (original)

'Motivation' graph notes\Week 03\Case 14

The atmosphere of the class was lively and active. It was hard not to get involved in learning in the class. So, for this semester, to actively get involved in every class of Interactive English is my goal. I hope I can keep this active attitude in learning English during this semester. (translated)

Field notes\Week 06

I found that they loved and enjoyed this task a lot! They dedicated themselves to the group collaboration. They worked and helped one another in order to complete the task. I found that they spoke a lot in English in discussion than before. Their ambitions were obvious. They wanted to reach this goal. (original)

Change 6: They seemed to agree that they would continue learning English out of the class.

Also, evidence related to learner autonomy were found to exist in the qualitative data. A total of the 29 items (mentioned by all the 24 students) relating positively to continuing to learn English out of the class were found in their diaries. The qualitative results reflect that they would take on their responsibility and keep on learning English out of the class. For example:

Learner Diaries\Week 05-Q3\Case 14

Q3: Did the tasks motivate you in your English learning? Why?

Yes. Sometimes when I could not express my ideas and opinions well in English in class, I would eager to know how to say it properly afterwards. So, when I faced something which I did not understand in class, I would try to learn it after class. (translated)

Learner Diaries\Week 11-Q7\Case 15

Q7: How motivated do you feel to carry on learning English?

I'm sure I will carry on learning English. English is very important in these days.

From this class, I understand that I have a lot of things to learn in the “English world”. That is what I feel now. So I suppose myself keep on learning English. (original)

Change 7: They were motivated to learn English because they wanted to pass exams or find a good job in the future.

With regard to their instrumental motivation in English learning, as mentioned in 6.1.1 above, only 11 items (mentioned by 9 students) relating to instrumental motivation were found to exist in the qualitative data. Of these, 8 items were found in their diaries, 2 items in ‘motivation’ graph notes, and 1 item in field notes. They were mostly related to opportunities associated with English learning, e.g., passing exams or finding a good job in the future. For example:

Learner Diaries\Week 11-Q4\Case 18

Q4: What else motivated you?

Of course, the bigger motivation was for my future needs. In this modern world, English plays an important role in our daily life. And perhaps one day in my job, English will become necessary. So I study English, and I will never stop learning it. (original)

Learner Diaries\Week 11-Q7\Case 01

Q7: How motivated do you feel to carry on learning English?

The score of the class will be greatly affect learning motivation. (original)

In sum, the qualitative data which were collected during the program through the ‘motivation’ graph notes, weekly learners’ diaries, and researcher’s field-notes insightfully reveal that the students’ increased motivation in English learning was internal more than external. The results provide evidence that the students’ desire to learn English was related to feeling good, e.g., the process of learning English was fun and interesting. Also, the relevance of tasks and the collaborative nature of TBL were the motivational factors for English learning. In addition, most of them had been motivated to set goals for their English learning and would autonomously continue learning English out of the class. The external factors related to opportunities associated with English learning, e.g., finding a job in the future or passing exams, although not mentioned much in the qualitative data, also influenced their motivation in English learning to some extent.

Evidence of changes from post-program qualitative data

Since the qualitative data on the changes of the student’s motivation during the program have already been discussed above, it is necessary to turn to the post-program qualitative data for exploring more in depth how and why these changes occurred. The data were collected after the program through follow-up interviews, as described in 5.5.4, when students had the opportunity to reflect back on the whole experience of learning through the task-based EFL program.

Change 1: Motivation for learning English was generally increased.

The post-program interview data indicate that the students’ overall motivation for learning English was increased through the task-based EFL program. A total of 353 items (mentioned by all the 24 students) relating positively to motivation in English learning were found to exist in the interview transcriptions, as categorized below:

- ❖ Intrinsic motivation (A total of 194 items)
 - They were motivated because of the tasks (128 items)

- They were motivated because of the pleasant learning atmosphere (24 items)
- They were motivated because of their own reasons (24 items)
- They were motivated because of their own goals (18 items)

- ❖ Extrinsic motivation (A total of 114 items)
 - They were motivated because of peer students (78 items)
 - They were motivated because of the teacher (30 items)
 - They were motivated because of their parents (6 items)

- ❖ Instrumental motivation (A total of 45 items)
 - They were motivated because they wanted to pass an English proficiency test (like TOEFL and IELTS) (17 items)
 - They were motivated because they wanted to study abroad (13 items)
 - They were motivated because learning English would help them to have good jobs (12 items)
 - They were motivated because learning English would help them to be successful in their studies (3 items)

As shown above, the number of positive mentions for the tasks themselves is by far the greatest (128 of 188 mentions) of the factors listed above. The next highest number of mentions relates to the students' peers, again reflecting the collaborative nature of task-based learning. Clearly, then, the students enjoyed the classes mainly because of tasks and the approach. More findings and discussion about the changes in motivation is presented in the following sections.

Change 2: They tended to agree more that the process of learning English was fun and interesting.

The results from the post-program interview data also provide strong evidence that the students found the process of learning English through tasks was fun and interesting. A total of 127 items (mentioned by all the 24 students) relating to the pleasantness of English learning in the program were found to exist in the interviews. This supports the finding noted in 6.1.2, indicating that the task-based program positively affected the students' intrinsic motivation in English learning. For example:

Interview\Q9\Case 04

The teacher: Did you enjoy your English classes more or less when you did tasks?

The student: Yes. More.

The teacher: Why?

The student: First, the way of learning English in the class was different from other English classes I had experienced. In Mainland China, it was impossible for us to explore English by tasks and to have fun in the learning process. Second, I found that the learning atmosphere in the class was lively and enthusiastic. When I learnt something from the class, I would write it down. I felt relaxed in English learning in this way. Moreover, I have learnt a lot from the class during this semester indeed! (translated)

Change 3: They were motivated because the tasks were relevant to them.

The results from the post-program qualitative data are in line with the results from the qualitative data during the program. A total of 104 items (mentioned by all the 24 students) positively relating to the relevance of tasks were found to exist in the interviews. For

example:

Interviews\Q12\Case 12

The teacher: Through the program, what factors motivated you most in the tasks?

The student: I think there were two factors which motivated me most. One, task collaboration; the other, whether the tasks were relevant to my life.

The teacher: You just said that the relevance of the tasks was one of the motivational factors in your learning. How did it motivate you?

The student: When the tasks were relevant to my life, it means that they would likely be used in practice in the future. Then I had a desire to learn them well because I wanted to know whether I could apply what I had learnt in the class to my real life in practice. So, I became motivated to learn.

The teacher: That motivated you to learn?

The student: Yes! (translated)

Change 4: They were motivated because of the collaborative nature of task-based learning.

The results from the post-program qualitative data indicated that the students were motivated by task collaboration. A total of 78 items (mentioned by all the 24 students) positively relating to task collaboration were found to exist in the interviews. One transcript is extracted as follows:

Interviews\Q9\Case 24

The teacher: Did you enjoy your English classes more or less when you did tasks?

The student: Yes, more.

The teacher: Why?

The student: As I said, I liked the interactive way to learn English. I liked the way of working with partners for the tasks. For example, when we were doing the task about résumé, we had to discuss and negotiate with one another until we reached a consensus. Then we decided who would be a representative to show our 'outcome' to the class in presentation. I loved this way of learning English. I felt that the main reason why I loved this class was because of task collaboration. (translated)

Change 5: They tended to set their own goals for English learning.

Again, in relation to goal-setting for English learning, the results from the post-program qualitative data are in line with the qualitative finding during the program, noted in 6.1.5. The post-program qualitative results indicate that some of the students had set goals for English learning through the task-based program. A total of 22 items (mentioned by 17 students) relating to goal-setting for English learning were found to exist in the interviews. For example:

Interview\Q12\Case 10

The teacher: What factors motivated you most in the tasks? Why?

The student: I think they were first, the topics of the tasks. Second, whether the tasks were of benefit to my goal. Third, whether they were relevant to my life.

The teacher: Anything else?

The student: Perhaps, whether they were challenging and whether the tasks could help me to improve my abilities.

The teacher: So these factors would motivate you in your English learning?

- The student: Yes.
The teacher: OK. What was your goal?
The student: I wanted to improve my abilities, especially English speaking competence. In the past, I had not had much confidence in speaking English with people. So, I wanted to improve my English speaking through the tasks. When I found the tasks could be beneficial to my goal, I was motivated to learn.
(translated)

Change 6: They seemed to agree more that they would continue learning English out of the class.

Again, the post-program qualitative results are in line with those in qualitative data during the program. A total of 38 items (mentioned by 20 students) positively relating to continuing to learn English out of the class were found in the post-program qualitative data. One example is shown as follows:

Interview\Q5\Case 08

- The teacher: You are saying that the tasks motivated you in your English learning? How?
The student: Yes, especially in the outdoor task, I found that English was really close to our daily life. Because, you know, we only had learnt English in the classroom. But when we were given a chance to use English outside the classroom, I found that my horizon was expanded. There were lots of things we could learn in English. So, through the outdoor task, I realized that I had problems in vocabulary in communication. Then, this motivated me to keep learning because I wanted to improve my English. For me, I won't regard learning English as only a school subject. I will keep learning it not only in class but also out of the class.
The teacher: So you will autonomously keep learning English?
The student: Yes, because I wanted to improve my English! (translated)

Change 7: They were motivated to learn English because they wanted to pass exams or find a good job in the future.

As mentioned above in 6.2.1, the students' motivation in English learning was affected by instrumental factors. A total of 45 items (mentioned by 18 students) relating to instrumental motivation was found to exist in the interview. In this aspect, the post-program qualitative results are in line with those in the qualitative data from during the program. The results from the post-program qualitative data mostly indicate that the students' instrumental motivation was related to opportunities associated with English learning, e.g., passing exams or studying abroad in the future. For example:

Interview\Q5\Case 24

- The teacher: You said that the tasks and the CET-4¹ were the most motivational factors during these 15 weeks. How did they affect you?
The student: When I found the tasks were interesting, I would like to surf the Internet, especially English websites, to look for more information about them. And when I realized the exam was approaching, I worked harder to prepare for it. (translated)

¹ College English Test (Band 4) – A national English language proficiency examination for college students in China.

In sum, the results from post-program qualitative were in line with those from qualitative data during the program, insightfully revealing that the students' increased motivation in English learning was internal more than external. The pleasantness of English learning experience through the program, the relevance of tasks, and the collaborative nature of TBL were the main internal factors of motivation for English learning. On the other hand, the external factors were mostly related to opportunities associated with English learning, such as finding a job in the future or passing exams. In addition, most of them had been motivated to set goals for their English learning and their learning autonomy was increased as cited that they would continue learning English out of the class.

CONCLUSIONS

In order to answer the question 'Why don't learners learn why teachers teach?', this study was conducted using a task-based EFL program which is different from traditional English teaching approaches. The purpose was to narrow the gap between English teaching and learning by examining whether a TBL approach could positively affect learners' motivation for learning English. In order to respond to the research questions what changes in motivation are apparent and why, this study used action research as a tool in the process, investigating the correlation between TBL and tertiary learners' motivations in English learning through the programme.

As the results showed above, students' motivation was increased by the task-based EFL program. They were motivated mainly because they found that this way of learning English was fun and interesting. They were also motivated because of task relevance and the collaborative works in the program. In addition, the students became autonomous learners. They tended to set goals for English learning. They were willing to take on their responsibility and keep on learning English out of the classroom. Since the findings in this study provided evidence that the students' motivation for English language learning was enhanced by TBL, it reasonably leads to the conclusion that this approach is beneficial to tertiary learners in English language learning.

With regard to the field of English language instruction in Macao, this exploratory study provides a clue to English teachers to consider a more viable alternative to current practices for teaching English. This study also encourages the Macao government and policy makers to consider a curriculum reform and a long-term educational policy for English language teaching and learning. Not only is TBL worth advocating in Macao, it is worth exploring in other contexts as well.

REFERENCES

- Allwright, R. L. (1984). Why don't learners learn what teachers teach? – The interaction hypothesis. In D. M. Singleton and D. G. Little (Eds.), *Language Learning in Formal and Informal Contexts* (pp. 3-18) IRAAL: Dublin.
- Cohen, A. D., & Dornyei, Z. (2001). Taking my motivational temperature on a language task. <http://carla.acad.umn.edu/profiles/Cohen-profile.html>
- Dornyei, Z. (1994). Motivation and motivating in the foreign language classroom. *Modern Language Journal*, 78, 273-284.
- Dornyei, Z. (1998). Motivation in second and foreign language learning. *Language Teaching*, 31, 117-135.
- Dornyei, Z. (2001 a). *Motivational strategies in the language classroom*. Cambridge: Cambridge University Press.
- Dornyei, Z. (2001 b). *Teaching and researching motivation*. Harlow: Longman.
- Dornyei, Z. (Ed.). (2003). *Attitudes, orientations, and motivations in language learning*. Oxford: Blackwell.

- Dornyei, Z., & Otto, I. (1998). Motivation in action: A process model of L2 motivation. Working Papers in Applied Linguistics (Thames Valley University, London), 4, 43-69.
- Ellis, G., & Sinclair, B. (1989). Learning to learn English. Cambridge: Cambridge University Press.
- Ellis, R. (2003). Task-based language learning and teaching. Oxford: Oxford University Press.
- Grotjahn, R. (1987). On the methodological basis of introspective methods. In C. Faerch & G. e. Kasper (Eds.), Introspection in second language research. Clevedon Avon, England: Multilingual Matters.
- Julkunen, K. (2001). Situation- and task-specific motivation in foreign language learning. In Z. Dornyei & R. Schmidt (Eds.), Motivation and second language learning (pp. 29-42). Honolulu, HI: University of Hawaii Press.
- Keller, J. M. (1983). Motivational design of instruction. In C. M. Reigeluth (Ed.), Instructional design theories and models: An overview of their current status. Hillsdale, NJ: Lawrence Erlbaum.
- Liao, X. (2004). The need for communicative language teaching in China. *ELT Journal*, 58(3), 270-273.
- Littlewood, W. (1981). Communicative language teaching. Cambridge: Cambridge University Press.
- Nunan, D. (1988). The learner-centred curriculum. Cambridge: Cambridge University Press.
- Nunan, D. (1992). Research methods in language learning. Cambridge: Cambridge University Press.
- Nunan, D. (1999). Second language teaching & learning. Boston: Heinle & Heinle Publishers.
- Nunan, D. (2004). Task-based language teaching. Cambridge: Cambridge University Press.
- Richards, J., & Rodgers, T. (2001). Approaches and methods in language teaching. Cambridge: Cambridge University Press.
- Van den Branden, K. (2006). Task-based language education: From theory to practice. Cambridge: Cambridge University Press.
- William, M., & Burden, R. (1997). Psychology for language teachers. Cambridge: Cambridge University Press.
- Willis, D., & Willis, J. (2007). Doing task-based teaching. Oxford: Oxford University Press.
- Willis, D., & Willis, J. (Eds.). (1996). Challenge and change in language teaching. Oxford: Heinemann Macmillan.
- Willis, J. (1996). A framework for task-based learning. Essex: Longman.

Grammar Instruction and Corrective Feedback: Cross-regional Student Perception Differences

Ming-chu Liao (廖明珠), Hung-chun Wang (王宏均)

National Changhua University of Education
Hsin Sheng College of Medical Care and Management

bravoclaire@gmail.com
rogerhchwang@hotmail.com

This follow-up on Liao and Wang's (2008) study of high school teachers' and students' perception of grammar teaching and corrective feedback extends to ascertain whether student perception differences exhibit regional variance. Based on Schulz's (2001) design and Borg's (1998) taxonomy of six teacher behaviors in teaching grammar, a questionnaire composed of thirty items was modified by researchers as a major source of data and distributed among 711 students in ten senior high schools located in the northern, central, southern, eastern regions, plus two off-shore islands. Two high schools were selected from each geographical region for comparison of regional differences in student perceptions. Results indicated students sharing nearly identical views toward grammar instruction in general, but regional variances in beliefs emerged in terms of the role of grammar teaching and corrective feedback. Students in eastern Taiwan held the strongest belief in grammar instruction and corrective feedback among all of the student groups. Based on the findings, it is suggested that teachers in different regions should attend to students' various needs in grammar learning and corrective feedback.

INTRODUCTION

Since the inception of the *Nine-Year Integrated Curriculum* in the 2001 academic year, a great importance has been attached to communicative language teaching (CLT) in English education. This approach has not only dominantly motivated language teachers to highlight development of learners' communicative competence and cultural awareness, but encourage learners' engagement in authentic and functional language use in classroom activities. Two central principles of CLT are that, first, grammar is best learned via learners' involvement in communication. Second, fluency should hold priority over accuracy, meaning teachers should tolerate learner errors if they do not hinder communication. Ever since promotion of CLT in English education, these two ideas of CLT have gradually come to prevail on this island.

Recent years have seen a number of studies have document striking regional differences as regards implementing English education in urban versus rural areas. For instance, a study conducted by King Car Educational Foundation in 2006 indicated the seriousness of urban-rural gap from parents' perspectives. They surveyed parents of primary school students in remote areas of Taitung, Hualien, Kinmen, Matsu, et al. and discovered that about 68.01 per cent of 1463 parents in these rural areas maintained that there were serious urban-rural gap in regional English education. Factors resulting in these differences included the lack of English teachers, learning stimuli, and learning resources. Differences in practice of regional English education can arise in students' achievement. Nationwide research by Chang *et al.* (2004) probed junior high school graduates' performance on the 2002 Basic Competence Test, and showed students in Taipei outperforming their counterparts in other locales, particularly Taitung and off-shore islands like Penghu. Chang *et al.* further attributed these differences to

serious urban-rural differences, parents' socio-economic status, parental involvement, supply of learning resources and qualified teachers, along with students' motivation and interest.

Since not all parts of Taiwan are going forward together with respect to regional English education, this phenomenon impels us to gauge how two CLT concepts, grammar instruction and error correction, have developed in different regions of Taiwan. Effectiveness of explicit grammar instruction and error correction has been a controversial issue in language education. Because "attention to form in the input competes with attention to meaning," they are often considered as conflicting with meaning-based CLT approach (VanPatten, 1990, p. 296). In this way, growing concern for learners' communicative competence has relatively reduced the importance of direct grammar teaching and error correction; instead, greater attention is drawn to how communication can be created to facilitate learning. Also, there are regional differences in permeation of CLT, whose impact may vary between regions islandwide. With this in mind, we assess potential cross-regional disparity in student perspectives on grammar instruction and error correction, our focus on differences and reasons behind them.

LITERATURE REVIEW

SLA researchers have paid a great deal of attention recently to teacher cognition and student perception. The former refers to teacher's understanding of how language is learned and should be taught, which has stark impact on actual instructional practices (Andrew, 2001; Borg, 1999, 2003; Fang, 1996; Johnson & Goettsch, 2000; Kagan, 1990). Importance of teacher beliefs, attitudes, and theoretical knowledge and assumptions in educational research was thus recognized (Burns, 1996; Johnson, 1994). Exploratory studies have unearthed evidence of student perception's strong linkage with learner motivation and strategy (Fox, 1993; Horwitz, 1988; Kern, 1995). A close bond between student beliefs and how (well) they learn has emerged (Bernaus & Gardner, 2008). Efficacy of explicit grammar teaching and corrective feedback in language classrooms has raised concern for foreign or second language pedagogy over the past thirty years. Traditionally, focus of grammar teaching emphasized structural or descriptive grammar, which stressed form over meaning. Recent increased attention to learner's communicative ability has regarded grammar instruction as an impediment since "attention to form in the input competes with attention to meaning" (VanPatten, 1990, p. 296).

Krashen's (1981; 1994) Input Hypothesis argued that comprehensible input and affective filter determine the outcome of second language acquisition: appropriate learning milieu containing sufficiently comprehensible input and low anxiety facilitates language acquisition better than formal instruction. Also, other criticisms of formal grammar teaching have criticized it as "unhelpful" (Prabhu, 1987, p. 2) and "limited" (Krashen, 1982, p. 112). Van Patten, Williams, Rott, and Overstreet (2004) noted that forging connections between form and meaning is a bedrock aspect of language acquisition. Similar to grammar instruction, objection to direct grammar correction has been linked with its ineffectiveness and harmful side effects. Strong opponents, skeptical about the role of direct error correction, included Truscott (1996; 1999), who contended that "oral [grammar] correction does not improve ability to speak grammatically, and Krashen (1985), who claimed that the correction process may raise learner anxiety and hinder performance.

Liao and Wang (2008) evaluated how senior high school English teachers and students nationwide perceived grammar instruction and error correction. Participants hailed from northern, middle, southern, eastern, and off-shore Taiwan. Since economically and socioculturally developed regions could vary in development of English language teaching (Hu, 2003), there arose a crucial call for probing into regional perception differences. This study followed the track and further explored regional differences to uncover any perceptual differences regarding the role of grammar instruction and error correction.

As stated above, due to geographical variation, not all parts of Taiwan have progressed

equally with respect to local English education. Previous studies have also indicated some cross-regional differences in supply of qualified English teachers, learning stimuli, and learning resources. While their findings are intuitively true, little research has been conducted to specifically focus on how learners in different regions perceive grammar instruction and error correction in classroom activities. This paper thus attempts to address this issue in order to find out what are cross-regional perception differences and what causes these differences.

METHOD

Participants

A total of 711 participants were recruited from ten high schools in five geographical regions in Taiwan: Northern, Central, Southern, Eastern, and off-shore. Two schools were selected from each region; all participants were third-year students (Table 1).

Table 1. *Demographic Data of Participants*

Region	North		Central		South		East		Off-shore		Total
School	N1	N2	C1	C2	S1	S2	E1	E2	O1	O2	
N	113	57	83	65	71	72	59	56	60	75	
Total	170		148		143		115		135		711

Instruments

Student Questionnaire. One student questionnaire was developed based upon Schulz's (2001) earlier design. Schulz studied cross-cultural differences in teacher and student perceptions regarding grammar instruction and error correction, and the student questionnaire that he developed included thirteen statements, which lay a solid foundation for development of our student questionnaire. In addition to these 13 statements, our questionnaire further added 17 statements. All 30 statements were rated on a four-point Likert scale from 1 *strongly disagree*, 2 *disagree*, 3 *agree*, and 4 *strongly agree*. Inclusion of these statements was based upon Borg's (1998) taxonomy of teacher behavior in grammar teaching: i.e., error analysis, reference to Students' L1, grammatical terminology, grammar rules, practicing grammar, and grammar and communicative ability. We added one more dimension in student questionnaires: general perception of grammar instruction and error correction, to study their holistic ideas about grammar teaching and error correction. In this vein, these seven dimensions laid a solid groundwork for design of the student questionnaire.

Telephone Interviews. Since the interviewed students came from divergent regions in Taiwan, it is more feasible to conduct telephone interviews instead of traditional face-to face interviews. For data triangulation, though 151 interviewed senior high school students were recruited on a volunteer basis, researchers aimed at selecting those who can reveal reasons that led to those most salient regional differences in terms of two major categories: grammar instruction and corrective feedback. They were informed that 25-minute structured telephone interview would be recorded for further analysis.

Data Collection and Analysis

Researchers first distributed questionnaires to students from ten high schools located in five areas of Taiwan. Students responded to questionnaires under supervision of their English teachers, who helped to distribute questionnaires to their fellow students. After the teachers in each school reported the estimated number of students that could participate in the current study, the researchers mailed student questionnaires to each school. All students responded to questionnaires in class; upon completion, they handed it in to a teacher. Once questionnaires were collected, they were returned to researchers by mail. In addition, data were computed and analyzed by means of Analysis of Variance (ANOVA) and Scheffe post-hoc tests. These discerned any differences in statements on the student questionnaire among students in the

five regions. In addition, upon collection of questionnaires, the first researcher conducted a structured-interview with 151 high school students on a volunteer basis. Those who indicated their interest in joining subsequent interviews on the questionnaire were requested to discuss what they thought of grammar instruction and error correction in great detail on the phone. Data were transcribed and analyzed to discern what factors led to differences in students' perceptions of grammar instruction and error correction.

RESULTS AND DISCUSSION

Questionnaire Data

Results demonstrate students in five regions markedly differing in various aspects of perceived grammar instruction and error correction. Tables 2 to 9 plot their self-perceptions regarding the role of grammar instruction, while Tables 10 to 13 depict various perceptions concerning the necessity of error correction in classroom settings.

Tables 2 and 3 relate to students' perception of the importance of learning grammar. Those in the eastern region seemed to show stronger belief in such learning, to see it as vital to the mastery of English in the sense that it can facilitate their learning achievement.

Table 2. *Importance of Learning Grammar*

#1: I think learning grammar is very important to the mastery of English.					
Region	n	<i>M</i>	<i>SD</i>	F	Post-hoc
North	170	2.92	.63	6.27**	East > North*
Central	148	2.91	.76		East > Central*
South	143	3.01	.62		
East	115	3.27	.72		
Off-shore	135	3.12	.72		
Total	711	3.03	.70		

* $p < .05$ ** $p < .01$

Table 3 concerns student interest in learning English. Again, those in the east indicated the strongest interest in learning grammar. Difference between students in the northern and eastern regions attained a significant level, meaning those in the north showed comparatively lower interest in learning grammar.

Table 3. *Interest in Learning Grammar*

#5: (S) I love learning grammar.					
Region	n	<i>M</i>	<i>SD</i>	F	Post-hoc
North	170	2.05	.82	3.44**	East > North*
Central	148	2.24	.79		
South	143	2.22	.79		
East	115	2.38	.62		
Off-shore	135	2.14	.80		
Total	711	2.19	.78		

* $p < .05$ ** $p < .01$

Table 4 shows students' perception regarding class time needed for teaching grammar rules. Students in the east believed that much class time should be devoted to this phase, a belief far deeper than that of students in the north. This result conforms to findings discussed above: students in the east valued grammar instruction higher than those in other regions and thought it should take much time in class, owing to its importance to mastery of English.

Table 4. *Time Allocation to Teaching Grammar Rules*

#6: (S) English classes should allocate lots of time to teach grammar rules.					
Region	n	<i>M</i>	<i>SD</i>	F	Post-hoc
North	170	2.25	.71	4.11**	East > North*
Central	148	2.44	.84		
South	143	2.48	.68		
East	115	2.58	.65		
Off-shore	135	2.36	.72		
Total	711	2.41	.73		

* $p < .05$ ** $p < .01$

As for the importance of learning terminology, students all regions generally held that it is important to understand grammar terminology (Table 5), yet those in the east showed the strongest belief in it, resulting in significant differences from the other regions.

Table 5. *Importance of Terminology in Grammar Instruction*

#8: (S) Understanding grammar terminologies is very important to the study of grammar.					
Region	n	<i>M</i>	<i>SD</i>	F	Post-hoc
North	170	2.97	.71	5.31**	East > North*
Central	148	2.93	.82		
South	143	2.97	.63		
East	115	3.24	.62		
Off-shore	135	2.87	.57		
Total	711	2.99	.69		

* $p < .05$ ** $p < .01$

Table 6 deals with the need for explicit grammar instruction. The question asked about whether learners prefer to learn grammar by means of various four-skill activities, without direct involvement in explicit grammar instruction. Results indicate students in the east hold the strongest dislike for this statement. From another perspective, their response may further express their emphasis on explicit grammar instruction in English classes. On the contrary, students in the north showed relatively less interest in explicit grammar instruction, believing that they could learn grammar by themselves in non-grammar-focused four-skill activities.

Table 6. *Need for Explicit Grammar Instruction*

#14: (S) I think I can learn grammar in listening/speaking/reading/writing activities by myself, without teacher explicitly teaching grammatical rules.					
Region	n	<i>M</i>	<i>SD</i>	F	Post-hoc
North	170	2.44	.83	6.91**	North > East*
Central	148	2.45	.72		
South	143	2.27	.73		
East	115	2.07	.65		
Off-shore	135	2.52	.88		
Total	711	2.36	.78		

* $p < .05$ ** $p < .01$

Table 7 plots the importance of simulated real-life activities in English class, compared to that of learning and practicing grammar. Results show that students in the central region valued the necessity of simulated real-life activities more than those in other regions. Note that students in the eastern and off-shore regions demonstrated lowest agreement rates to the importance of such activities. This result is in agreement with Table 2 that students in these two regions strongly believed in the importance of learning grammar to mastery of English

and thus paid relatively less attention to simulated real-life learning in classes.

Table 7. *Importance of Simulated Real-life Activities over Learning Grammar*

#16: (S) I think learning English in real situations (e.g., role-play or dialogue practice) is more important than studying grammar and practicing grammatical patterns.

Region	n	M	SD	F	Post-hoc
North	170	3.22	.65	5.12**	Central > East*
Central	148	3.44	.67		Central > Off-shore*
South	143	3.38	.59		
East	115	3.17	.75		
Off-shore	135	3.17	.64		
Total	711	3.28	.67		

* $p < .05$ ** $p < .01$

Table 8 pictures importance of grammar practice in grammar instruction. Students in the south showed the strongest belief in grammar practice in grammar instruction, and their belief resulted in their preference. Interestingly, students in the north showed the least agreement to the importance of grammar practice in general.

Table 8. *Importance of Grammar Practice in Grammar Instruction*

#17: (S) I think it can help students learn grammar better if teachers can provide practicing activities after teaching grammatical rules.

Region	n	M	SD	F	Post-hoc
North	170	3.03	.64	6.69**	South > North*
Central	148	3.22	.58		South > Off-shore*
South	143	3.36	.59		
East	115	3.24	.60		
Off-shore	135	3.11	.58		
Total	711	3.18	.61		

* $p < .05$ ** $p < .01$

Table 9 pertains to first language assistance. Students in general preferred Chinese-only to English-only grammar instruction, with the agreement rate in the east strongest.

Table 9. *First Language Assisting Grammar Instruction*

#22: (S) Teaching grammar in Chinese is more helpful to students than teaching in English.

Region	N	M	SD	F	Post-hoc
North	170	2.63	.74	6.92**	East > North*
Central	148	2.72	.74		East > South*
South	143	2.57	.71		East > Off-shore*
East	115	2.99	.77		
Off-shore	135	2.57	.77		
Total	711	2.68	.76		

* $p < .05$ ** $p < .01$

The second phase of this study aims to discern cross-regional differences in perceptions regarding error correction in English class. First, one striking finding is that students in the eastern region showed the weakest agreement to the statement, 'When students make grammar errors in speaking English, if those errors do not hinder communication, teachers should not correct students.' This shows that they felt the necessity of grammar correction in speaking more than students in the other regions. Moreover, Table 10 is concerned with the need for grammar

correction in speaking. Although students in general perceived the importance of grammar correction in speaking, students in the eastern region indicated the highest agreement.

Table 10. *Attitudes for Grammar Correction in Speaking*

#26: (S) When I make grammar errors in speaking English, I hope teachers can correct my errors.					
Region	n	M	SD	F	Post-hoc
North	170	3.18	.60	4.49**	East > North*
Central	148	3.16	.63		East > Central*
South	143	3.22	.60		
East	115	3.43	.56		
Off-shore	135	3.30	.56		
Total	711	3.25	.60		

* $p < .05$ ** $p < .01$

As can be seen in Table 11, students in eastern region strongly felt necessity of grammar correction in writing, significantly more than students in the other regions.

Table 11. *Attitudes for Grammar Correction in Writing*

#27: (S) When I make grammar errors in writing, I hope teachers can correct my errors.					
Region	n	M	SD	F	Post-hoc
North	170	3.28	.59	6.39**	East > North*
Central	148	3.34	.57		East > Central*
South	143	3.29	.56		East > South*
East	115	3.60	.51		East > Off-shore*
Off-shore	135	3.36	.66		
Total	711	3.36	.59		

* $p < .05$ ** $p < .01$

Tables 12 and 13 depict impact of grammar correction in terms of assisting self-learning and others' learning respectively. Students in all of the five regions generally agreed that they could improve their learning when teachers corrected their or others' grammar errors. Of all students, those in the eastern region still showed the strongest belief in the impact of grammar correction on assisting their own and others' learning.

Table 12. *Impact of Grammar Correction on Assisting Self-Learning*

#29: (S) When teachers correct my grammar errors, it is helpful to my learning.					
Region	n	M	SD	F	Post-hoc
North	170	3.32	.53	3.36**	East > Central*
Central	148	3.29	.53		
South	143	3.37	.55		
East	115	3.51	.54		
Off-shore	135	3.33	.52		
Total	711	3.36	.54		

* $p < .05$ ** $p < .01$

Table 13. *Impact of Grammar Correction on Assisting Others' Learning*

#30: (S) When teachers correct my grammar errors, it is helpful to other classmates' learning.					
Region	n	M	SD	F	Post-hoc
North	170	3.12	.63	9.11**	East > North*
Central	148	3.15	.58		East > Central*
South	143	3.26	.64		East > South*
East	115	3.52	.50		East > Off-shore*
Off-shore	135	3.21	.59		
Total	711	3.24	.61		

* $p < .05$ ** $p < .01$

Interview Data

Since the two main foci of this study lie in grammar instruction and corrective feedback, the telephone interview aimed at eliciting information of selected volunteer participants who manifested perception of differences in these two dimensions. Among regions, questionnaire respondents from northern and eastern Taiwan exhibited greatest statistically significant differences. Therefore, the interviews were conducted with selected volunteered participants from these two regions to ascertain the reason behind the perception differences.

As for grammar instruction, the vast majority of eastern participants (93%) held positive attitudes toward its importance, interest in learning grammar, time allocation to grammar rules, and need for explicit grammar instruction. First of all, participants from Eastern Taiwan presented appreciably stronger belief in the importance of learning grammar than those from the North, and tended to consider good command of grammatical knowledge essential to mastery of English, whereas the latter attributed mastery of English to wider vocabulary.

Grammar is the most important part when it comes to mastering English...Because those who were considered having a good command of English were good at grammar. They tend to score high in English tests of all kinds. (East)

Mastering English involved vocabulary ability other than grammar. Learning grammar is not so important as building larger vocabulary. With a good command of vocabulary, one can guess from the content what the next text is about. (North)

Also, eastern participants not only strongly favored allocating more time to grammar rules but also called for explicit grammar instruction. They regard it helpful to understand the rules well by teachers' demonstration in class, whereas northern participants prefer to learn grammar in listening, speaking, reading, or writing activities.

Grammar is so crucial that the rules need to be fully understood. I like my English teacher to spent about half of the class time in explaining the grammar rules, it makes me feel secure and worthwhile to be in class. (East)

Explaining the grammar rules should not take too much class time, the less, the better. I like to learn grammar through activities in class, be it listening, reading, speaking, or writing. (North)

Regarding error correction, a large portion of eastern participants (80%) also reported positively to their attitudes for grammar correction in speaking and writing as well as impact of grammar correction on learning. Participants from the east demonstrated need and hope for teachers to correct errors in speaking and writing, but those from the north tended to favor not being corrected directly if grammar errors did not hinder communication or comprehension. Those from Eastern Taiwan also expressed that their own and classmates' learning benefit by obtaining the teacher's grammar correction; those from Northern Taiwan disagreed.

I found it impressive for my grammar to be corrected either in speaking or in writing English. I seldom forgot the corrected grammar errors and when I have chance to use them next time speaking or writing English, I would turn more alert to avoid making the same errors again. (East)

If corrected during the process of speaking English, I would feel embarrassed and feel hesitant to speak English or at least avoid using the grammar once being corrected. If my English is communicative, the teachers should not interrupt by correcting my grammar mistake. (North)

CONCLUSION

This study attempted to discern any regional differences in high school students' perception of grammar instruction and error correction in English class. Results of this study have demonstrated several cross-regional differences, mainly associated with students in the eastern region. Speaking of student perceptions of grammar instruction, our findings showed that students in the eastern region held the strongest belief in the importance of learning grammar in the mastery of English. With this belief in mind, they maintained that explicit grammar instruction was highly important in English education and indicated their interest in learning grammar. On the other hand, students in the east also favored error correction more than students in the other regions. They proposed that grammar correction was needed in both speaking and writing. Owing to its corrective merits, grammar correction was also considered to benefit their own learning and other classmates' learning. On the whole, we discovered that, of all five groups, students in the east indicated much stronger belief in efficacy of grammar instruction and error correction.

Senior high students from eastern Taiwan obviously perceived grammar instruction and error correction as more essential and crucial in their English learning, in comparison to those from other regions, especially northern Taiwan. This could be attributed to the difference between regions: e.g., convenience of information in northern Taiwan allowing a variety of access to information on English learning, so that they have been equipped with the recent trend of communicative approach, which neither isolates teaching of grammar rules in class and nor emphasizes accuracy over fluency for communication. Those in the east might be comparatively restricted in information, limited to traditional teacher-centered grammar instruction and error correction, and thus think higher of these aspects, failing to agree that grammar should be learned from listening, speaking, reading, and writing activities.

Furthermore, a pedagogical implication of this cross-regional difference may imply that teachers in various regions should attend to students' different needs of grammar learning and error correction. As Horwitz (1990) and Phillips (1999) stated, mismatch between student and teacher perception of learning may negatively affect instructional performance and learning outcome. Bilateral awareness of each other's perception must be raised, such that mutual compromise can establish teaching-learning correspondence (Brindley, 1989). We suggest that teachers in different regions should notice how grammar instruction and error correction are valued by fellow students, and their understanding of student needs may provide an efficient direction for effective classroom teaching.

This study is limited in the following respects. First, students in each region were not selected on a random basis, so their English achievement might have affected how they perceived grammar instruction and error correction, exerting influence on our cross-regional comparisons. In this regard, future research may consider it necessary to select students with similar English proficiency, in order to ensure the comparability of each student group at the outset of this study. Also, our study focused on students' perceptions of grammar instruction and error correction; no attempt was made at discussing teachers' perceptions in these five regions. Future research may extend the scope of this study to include regional variations in teachers' beliefs in grammar instruction and error correction, and gaps between teachers' and students' perceptions.

REFERENCES

- Andrew, S. (2001). The language awareness of the L2 teacher: Its impact upon pedagogical practice. *Language Awareness*, 6, 136-146.
- Bernaus, M. & Gardner, R. C. (2008). Teacher motivation strategies, student perceptions, student motivation and English achievement. *Modern Language Journal*, 92(3), 387-401.
- Borg, S. (1998). Teachers' pedagogical systems and grammar teaching: A qualitative study. *TESOL Quarterly*, 32(1), 9-38.
- Borg, S. (1999). Studying teacher cognition in second language grammar teaching. *System*, 27, 19-31.
- Brindley, G. (1989). *Assessing achievement in the learner-centred curriculum*. Sydney: National Centre for English Teaching.
- Burns, H. (1996). Starting all over again: From teaching adults to teaching beginners. In D. Freeman & J. C. Richards (Eds.), *Teacher learning in language teaching* (pp. 154-177). Cambridge: Cambridge University Press.
- Fang, Z. (1996). A review of research on teacher beliefs and practices. *Educational Research*, 38, 47-65.
- Fox, C. A. (1993). Communicative competence and beliefs about language among graduate teaching assistants in French. *Modern Language Journal*, 77, 313-324.
- Horwitz, E. K. (1988). The beliefs about language learning of beginning university foreign language students. *Modern Language Journal*, 72, 283-294.
- Horwitz, E. K. (1990). Attending to the affective domain in the foreign language classroom. In S. S. Magnam (Ed.). *Shifting the instructional focus to the learner* (pp. 15-33). Middlebury, VT: Northeast Conference on the Teaching of Foreign Languages.
- Hu, G. W. (2003). English language teaching in China: Regional differences and contributing factors. *Journal of Multilingual and Multicultural Development*, 24(4), 290-318.
- Johnson, B. & Goettsch, K. (2000). In search of the knowledge base of language teaching: Explanations by experienced teachers. *Canadian Modern Language Review*, 56, 437-468.
- Johnson, K. (1994). The emerging beliefs and instructional practices of preservice English as a second language teachers. *Teaching and Teacher Education*, 10, 439-452.
- Kagan, D. M. (1990). Ways of evaluating teacher cognition: Inferences concerning the Goldilocks principle. *Review of Educational Research*, 60, 419-460.
- Kern, R. G. (1995). Students' and teachers' beliefs about language learning. *Foreign Language Annals*, 28, 71-92.
- Krashen, S. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon Press.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.
- Krashen, S. (1985). *The input hypothesis: Issues and implications*. London: Longman.
- Krashen, S. (1994). The input hypothesis and its rivals. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 45-77). London: Academic Press.
- Prabhu, N. S. (1987). *Second language pedagogy*. Oxford: Oxford University Press.
- Schulz, R. (2001). Cultural differences in student and teacher perceptions concerning the role of grammar instruction and corrective feedback: USA-Columbia. *Modern Language Journal*, 85(2), 244-258.
- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46, 327-369.
- Truscott, J. (1999). What's wrong with oral grammar correction. *The Canadian Modern Language Review*, 55(4), 437-456.

- VanPatten, B. (1990). Attending to content and form in the input: An experiment in consciousness. *Studies in Second Language Acquisition*, 12, 287-301.
- VanPatten, B., Williams, J., Rott, S., & Overstreet, M. (2004). *Form-Meaning Connections in SLA*. New Jersey: Lawrence Erlbaum Associates.
- 搶救菜英語 逾 9 成偏遠地區家長贊成 (2006). Retrieved March 10, 2009 from <http://aecer.org/modules/news/article.php?storyid=400>
- 張武昌、周中天、陳純音、葉錫南、林正昌、許月貴 (2004)。國民中學學生基本學力測驗英語雙峰現象暨改進措施。教育部委託專案研究報告。

Learning Alone at a CALL Center: What Frequent Users Tell Us

Chun-cheng Lin (林俊呈), Chih-cheng Lin (林至誠)

National Taiwan Normal University

hugo5689@gmail.com
cclin@ntnu.edu.tw

The purpose of this study is to investigate EFL college students' perceptions of a university self-access center (SAC) in northern Taiwan. In the 2006-2007 academic year, a total of 1,104 students were enrolled in Freshman English and encouraged to use the SAC after class. While each of their visits was documented, the students, at the end of each semester, were required to fill out a survey questionnaire of 23 items, which mainly probed into their perceptions of training their language skills, improving their performances in class, and monitoring their own learning. More than a third of the enrolled (397, 35.96%) finished both questionnaires; and, 45 students, who visited the SAC more than 20 times in either semester, were further recruited for a semi-structured interview. The analysis of the two survey questionnaires shows that the students felt satisfied with the SAC facilities and considered the SAC a convivial environment for English learning. In particular, they were grateful to the assistant's timely help and were impressed with the technology-enhanced learning programs. Not only did they appreciate speech recognition in some pronunciation programs, they also enjoyed the freedom of viewing movies in the relaxing atmosphere. Correspondingly the analysis of the interviews reconfirms that their learning in the SAC helped them achieve their language learning objectives, among which listening skills are reported to be greatly improved. Also, they applied certain compensatory strategies to solve difficulties in their learning. To sum up, the EFL college students believed that the design of the self-learning program and the support provided in the SAC facilitated their self-directed learning as well as achievements.

INTRODUCTION

In the spring of 2001, a Self-Access Classroom (SAC) was established in National Taiwan Normal University, in which students, mostly freshmen, have access to various language learning software programs. *Freshman English* instructors often required students to visit the SAC as a complement to limited instructional hours. Before students entered the college, their English classes used to be teacher-directed and exam-oriented (Hsu & Xu, 2007), but this passive learning style is incongruent with current notions of "learning how to learn," learner-centeredness, or autonomous learning. In view of this, universities across the island have set up self-access systems to help foster learner autonomy. In Chao's (2005) survey of the autonomous and self-regulated learning activities among 107 colleges in Taiwan, these activities can be classified into five categories, i.e. learning resources, self-evaluation, one-on-one conferencing, exposure to English, and learning community. A further focus interview with 119 students from 7 universities showed that network-learning, English corner, and self-instruction community are the most popular among students.

Nowadays, CALL self-access classrooms which feature interaction, resources and individualization have proliferated and been believed to help foster learner autonomy. In these self-access classrooms, learners are expected to take much responsibility for their own learning, and learn to reflect their own learning (Gardner & Miller, 1999). Nonetheless,

several researchers have cautioned that while learners are engaged in self-access mode of learning, they may not be able to direct their own learning or learn actively (e.g. Benson, 2001). Moreover, while the necessity of measuring learners' gains through self-access learning is well recognized (Gardner & Miller, 1999; Morrison, 2005), Gardner (2002) identified four aspects that may induce problems for evaluating the effectiveness of a SAC, including (1) a heterogeneous learner population with regard to learning objectives, learning backgrounds and styles, (2) the uniqueness of each SAC, owing to their different objectives, (3) difficulty of data collection, mainly due to individualized learning patterns and the concern of the evaluator to avoid intruding into the learning process, and lastly (4) the difficulty of data analysis, as it is difficult to attribute learning specifically to what learners do in a self-access language learning classroom (pp.62-63). In spite of these problems, the recent decade has witnessed some studies on evaluating the effectiveness of SAC in improving linguistic proficiency and help learners take control of their learning (e.g. Cheng, 2006; McMurry, 2005; Ning, 2008; Reinders, 2000; Wang, 2005; Zou, 2006) . The current study will continue with this line of research and explore the perceptions of EFL college students toward learning in a CALL SAC. More specifically, frequent visitors' perceptions toward learning in a SAC are explored. Two research questions are addressed:

- (1) Do EFL college students perceive learning in a CALL SAC beneficial to their English learning?
- (2) What motivates EFL college students to come to the CALL SAC?

LITERATURE REVIEW

Definition and elements of self-access language learning

Sheerin (1991) defined self-access as “a way of describing materials that are designed and organized in such a way that students can select and work on tasks on their own” (p. 143). To Gardner & Miller (1999), self-access language learning is regarded as an integration of different elements, including resources, teachers and learners, management, system, individualization, needs/wants analysis, learner reflection, counseling, learner training, staff training, assessment, evaluation, and materials development. The introduction of self-access language learning is concomitant with changing roles of the teacher and the learner (pp. 9-10). In this mode of learning, learners are planners and motivators of their own learning, have to take an increasing amount of responsibility for their own learning, reflect on their own learning, and may have to redefine their learning goals to meet their learning needs. On the other hand, teachers, while still counselors, evaluators, materials developers, and assessors, need to give over some of their control and redefine their roles (Sturtridge, 1997).

In Figure 2, Jones (1998, p. 379) conceptualized learner autonomy as a continuum, with the classwork and homework requiring the least amount of learner independence. In this spectrum, three labels, teacher-directed, self-access, and self-directed, were placed. In Jones' conceptualization, self-access involves different kinds of autonomous learning, from *full autonomy*, *teach-yourself autonomy* to *homework*.

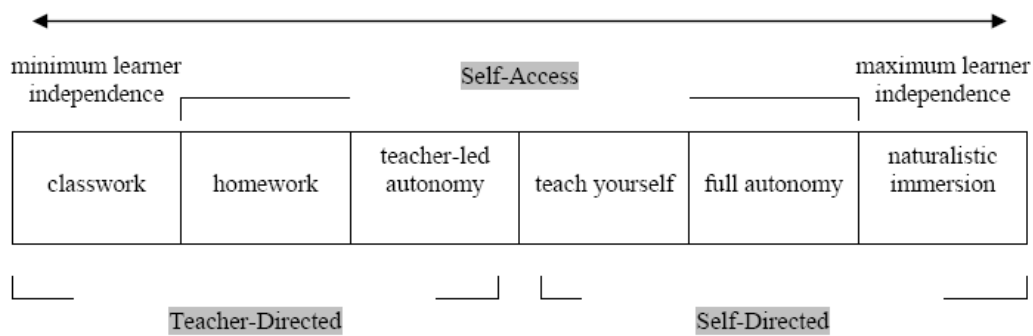


Figure 2 Jones' diagram on learner autonomy

Self-access language learning classrooms

The first self-access language learning center at CRAPEL (Centre de Recherche et d'Applications Pédagogiques En Langues) (Riley & Zoppies, 1985) was established based on the assumption that the access to a rich collection of audio-visual learning materials may provide learners opportunities to experiment with autonomous learning. Since this inception, self-access language learning classrooms have proliferated around the world, not only in the western educational settings but in the eastern context, such as the Independent Learning Centre of Chinese University of Hong Kong. In Taiwan, several self-access language centers have also been established, including the Language Teaching and Research in National Chiao Tung University (2002), and the Language Center in Soochow University (2002).

The setup of SAC is often based upon the assumption that the technology-based and resourced-based environment is conducive to students' language learning. In such an environment, learners have access to various sources of audio-visual input. Ideally, these inputs can be optimal if they are interesting or relevant to the learner, and comprehensible (Kenning & Kenning, 1990). Other than exposure to large quantities of input, learners can have opportunities to use the target language with the current automatic speech recognition (ASR) technology. While not engaged in interaction with an authentic audience, learners can have opportunities to have a conversation, produce language, and receive corrective feedback from the software. In addition, one feature that characterized the SAC is individualization, i.e., they can work competently at their own speed (Walker, 2005). In spite of the advantages of exposure, output, feedback and individualization, possible disadvantages or problems have also been brought up (Galavis, 1998; Walker, 2005), including the lack of real communication and the need to motivate students to come regularly. Hence, the effectiveness of self-access language learning classrooms, as suggested by Reinders and Cotterall (2001), can be improved in four areas: (1) self-access center administrators should explore learner beliefs; (2) students need an effective initial orientation to this learning environment; (3) administrators should provide on-going support to students (4) more links between the self-access center and the language classroom should be formed. So far, several studies have investigated different aspects of self-access language learning classrooms in counseling (Voller, Martyn & Pickard, 1999), learning goals (Carson & Allen, 2006; Ridley, 2000), and maximizing learners' access to learning resources (McMurry, 2005).

So far some studies have been conducted to explore language learners' perceptions of learning in a self-access language learning classroom and how these learners improve linguistically and may become autonomous (e.g. Klassen, Detarmani, Lui, Patri & Wu, 1998; Zou, 2006). Reinders (2000) investigated learners' perspectives on learner autonomy as well as their self-access language learning in an English proficiency program. Data collection was chiefly achieved through questionnaires and interviews with 16 students. Self-access learning

was found to be positively evaluated by the students, as a means both to develop language skills and to foster independent learning skills. A correlation between frequency and positive perception was observed. That is, students who use the resources more are more positive about its effects on their development in English. Lastly, it is reported that one crucial factor to the success of self-directed learning lies in the resources that students can have.

Aside from the studies abroad, three studies have been conducted to investigate the effectiveness of local SAC in colleges, and results generally showed that college students can benefit from learning in the SAC. Cheng (2003), mainly as a program evaluation, explored learners' beliefs and attitudes toward self-access, and their viewpoints on the self-access program and multimedia learning software in the Multimedia English Learning Center (MELC) in National Kaohsiung First University of Science and Technology. Findings suggest that most students reported to have improved their listening and speaking after their training in the center, but less than 30% of the students indicate that they have developed the independent learning ability or formed the regular learning habits after training. In another study by Wang (2005), 913 college freshmen were recruited and how cultural competence was developed in a CALL SAC was investigated. Results show that among the four language skills, students' listening ability was the most improved (38.69%), followed by writing (35.97%) and oral skills (11.44%). Via film watching, learners can appreciate different cultures and cultivate their critical thinking ability. One of the most recent studies is Ning (2008), who evaluated the self-access learning via the investigation of learners' perceptions of their learning in the SAC in National Taiwan Normal University. In that study, 1,613 students from *Freshman English* course were included, and 29 of them were further recruited for semi-structured interviews. It is found that students benefited more in vocabulary, listening, speaking, and grammar. Also, frequent visitors were found to metacognitively monitor their progress along the way, and could motivate themselves.

To sum up, the CALL SAC has been established in the educational contexts on the premises that learners can benefit from self-access mode of learning and make good use of the learning resources in the SAC. So far, several studies have confirmed these premises and implied that the more frequently learners visit the SAC, the more positive evaluations they hold toward it. It is hoped that via the current study, apart from exploring if SAC may facilitate language learning, their motivation for coming to the SAC can be further explicated.

METHODS

Participants

In the 2006-2007 academic year, a total of 1,104 college students in northern Taiwan were recruited as participants for the current study. Before the new semester began, these freshmen took a placement test composed of listening and reading comprehension questions. Based on the results, students were assigned into five levels: High, High-Intermediate, Low-Intermediate, Low and Basic. A two-hour *Freshman English* instruction was given per week over a year, and students were encouraged to visit the self-access English classroom on a regular basis. Also, one learning passport was given to each student, to record the date and learning software used during each of their visit to the center. Moreover, only those who stayed in the center for more than 40 minutes during each visit can be ratified.

Among these students, interviewees were chosen mainly based upon their visit frequencies. In the fall semester 20 learners and 25 in the spring one were chosen for a semi-structured interview. While two interviewees in the fall semester visited 8 times, the rest 18 came to the SAC from 12 to 47 times. As for interviewees in the spring semester, 23 came to the SAC from 10 to 57 times. Nonetheless, two interviewees did not come as frequently: one came to the SAC for five times, and the other eight times. Their inclusion for interview was mainly because they did not come to the SAC as frequently as they did in the previous

semester, and an in-depth interview may shed light on their attitudinal change. Table 1 presents the number of interviewees in each language proficiency level.

Table 1. *Number of interviewees in each proficiency level*

Proficiency	High	High-Intermediate	Low	Low-Intermediate	Basic	Total
Fall	4	6	5	4	1	20
Spring	7	13	5	0	0	25

Weekly assigned lessons in the SAC

The self-access classroom in National Taiwan Normal University is characterized by various multimedia learning materials and may facilitate learners' listening, speaking, reading, vocabulary and grammar. Students can come to the SAC, freely choose any material that they would like to use and learn at their own pace. In case students may not know the features of specific programs or do not know where to start, they can follow a weekly schedule, in which a software was recommended to students. In this way, those students who follow the weekly assigned lessons can get to know different multimedia programs, including *Tell Me More*, *Traci Talk*, *Tense Buster*, *Connected Speech*, *Studio Classroom* and so on.

Instruments

In the current study, two research instruments are utilized. The first one is a survey questionnaire of 23 items (See the survey questionnaire in Appendix). The first 19 items, based on a 4-point Likert scale (4 = Strongly agree, 1 = Strongly disagree), probe into learners' perceptions of the effectiveness of the SAC in training their language skills, of the help provided by the assistant, and of the physical environment of the SAC. In the rest of the items, learners are asked to select their most favorite and the least favorite software, areas of the SAC that can still be improved, and advantages of learning English in the SAC.

To substantiate quantitative data and to "[provide] details on the process that leads to a certain kind of situation" (Marquart, Li, & Zercher, 1997, p.10), the second instrument is a semi-structured interview at the end of each semester. Fourteen questions were asked of each interviewee. These questions include: (a) the *Freshman English* instructor's regulation of visiting the SAC, (b) self-set objectives, (c) the achievement of their objectives, (d) any tests that help evaluate the achievement of their objectives, (e) their learning in the SAC, (f) software selection, (g) opinions about certain software, (h) the greatest gain of visiting the SAC, (i) their reasons of keeping visiting the SAC, (j) suggestions, (k) other extracurricular English learning activity, (l) the most obvious advantages and disadvantages of the SAC, (m) the assigned lessons, and lastly (n) the change of the SAC within one semester. These interviews were conducted in Chinese, and all recordings were transcribed and translated into English.

Data analysis

Results from the survey questionnaire were computed and tabulated with the application of SPSS 13.0 (Statistical Package for the Social Science). In terms of two end-of-semester questionnaires, the ratings to the first 19 questions were calculated, followed by a paired-sample t-test analysis, via which ratings of those 397 visitors who filled out questionnaires in both semesters were compared. Via this, visitors' attitudinal change toward learning in the self-access language classroom over time was examined. Also, responses to the last four questions probing into learners' preferences toward specific learning software, the reasons that they came to the classroom, and areas that this classroom may improve, were computed and analyzed. Furthermore, the recorded interviews were coded and further

analyzed.

RESULTS

The purpose of this study is to investigate how EFL college students perceived their learning in the self-access classroom. Major findings from the survey results as well as the semi-structured interviews are presented.

Survey results

Table 2 presents the results of first 19 items of end-of-semester questionnaires in two semesters. The students' average ratings ranged from 2.97 to 3.45 in the first semester, and from 3.07 to 3.48 in the second one. In general, students in the second semester held more positive attitudes toward their learning in the center than those in the first semester, with higher average ratings in all 19 questions.

Table 2. *Results of the end-of-semester questionnaire in two semesters*

<i>N</i> =	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>
1104	(Fall)		(Spring)	
Q1	3.34	0.57	3.39	0.56
Q2	3.35	0.56	3.38	0.55
Q3	3.29	0.60	3.40	0.55
Q4	3.35	0.57	3.38	0.56
Q5	3.19	0.63	3.23	0.58
Q6	3.24	0.57	3.32	0.55
Q7	3.03	0.68	3.14	0.63
Q8	3.10	0.61	3.18	0.57
Q9	3.22	0.56	3.27	0.53
Q10	3.24	0.62	3.30	0.56
Q11	3.20	0.63	3.28	0.56
Q12	3.17	0.62	3.20	0.60
Q13	2.97	0.69	3.07	0.67
Q14	3.45	0.57	3.47	0.55
Q15	3.42	0.58	3.47	0.54
Q16	3.37	0.56	3.40	0.55
Q17	3.18	0.67	3.40	0.55
Q18	3.12	0.73	3.43	0.57
Q19	3.33	0.55	3.48	0.53
ALL	3.24		3.33	

The 19 items in the survey can be categorized into four main dimensions: (1) software (Q1-Q3) and physical environment (Q17-18) of SAC, (2) language training (Q4-Q7), (3) the assistant's help (Q14-Q16), and (4) the weekly assigned lessons (Q8-Q13). In terms of the software and physical environment of SAC, it is revealed that the language learning software is positively valued both students in two semesters. Students felt satisfied with the language learning software ($M = 3.34, 3.39$). Also, they positively evaluated the content ($M = 3.35, 3.38$) and technology ($M = 3.29, 3.40$) of these software. Nonetheless, while students in the second semester highly appreciated the environments ($M = 3.40$) and computer equipment ($M = 3.43$) in the SAC, learners in the first one did not respond so positively, with an average rating of 3.18 toward the environment and 3.12 toward the computer equipment. As for language learning, students in both semesters considered the SAC more beneficial to their listening ($M = 3.35, 3.38$) and reading ($M = 3.24, 3.32$) than speaking ($M = 3.19, 3.23$) or writing ($M = 3.03, 3.14$). Next, concerning the management and service of the assistant, learners were

generally satisfied, considering the assistant competent ($M = 3.45, 3.47$) and enthusiastic to help ($M = 3.42, 3.47$). Lastly, regarding the weekly assigned lessons, students' responses were moderately positive. They agreed that the weekly assigned lessons can boost their learning interest ($M = 3.24, 3.30$), complement inadequacy of *Freshman English* course ($M = 3.20, 3.28$), and increase certain aspects of their language proficiency ($M = 3.22, 3.27$). Nonetheless, learners did not particularly agree that these assigned lessons correspond with their language proficiency ($M = 3.10, 3.18$), nor do they highly agree with the statement that they used the weekly assigned lessons ($M = 2.97, 3.07$). Apart from these four dimensions, a general appraisal of the SAC was elicited from Q19, and the results indicated that students enjoyed learning in the SAC ($M = 3.33, 3.48$).

Among the 1,104 students, 397 (35.96%) completed survey questionnaires in two semesters. A paired-sample t-test was conducted on two questionnaires, in an attempt to investigate any attitudinal change. Table 3 presents the mean, standard deviation, t-value, and significance level of each statement.

Table 3. A paired-sample t-test analysis of attitudinal change

<i>N</i> =	Mean	SD	Mean	SD	<i>t</i>	<i>P</i>
397	(Fall)		(Spring)			(2-tailed)
Q1	3.37	0.53	3.38	0.56	-.244	.807
Q2	3.37	0.53	3.37	0.53	.173	.863
Q3	3.30	0.58	3.37	0.56	-2.026	.043*
Q4	3.36	0.55	3.34	0.57	.532	.595
Q5	3.20	0.60	3.22	0.60	-0.644	.520
Q6	3.23	0.56	3.29	0.55	-1.814	.070
Q7	3.00	0.67	3.13	0.62	-3.102	.002**
Q8	3.10	0.60	3.18	0.59	-2.243	.025*
Q9	3.23	0.53	3.27	0.55	-0.983	.326
Q10	3.25	0.60	3.29	0.57	-1.293	.197
Q11	3.20	0.61	3.28	0.55	-2.255	.025*
Q12	3.18	0.59	3.22	0.60	-1.104	.270
Q13	2.98	0.67	3.10	0.64	-2.699	.007**
Q14	3.52	0.52	3.47	0.56	1.581	.115
Q15	3.50	0.53	3.45	0.55	1.524	.128
Q16	3.41	0.54	3.42	0.57	-0.164	.870
Q17	3.16	0.69	3.40	0.56	-6.572	.000**
Q18	3.11	0.72	3.42	0.60	-7.408	.000**
Q19	3.36	0.52	3.47	0.54	-3.820	.000**

Note: * $p < 0.05$; ** $p < 0.01$

Among the 19 items, seven items reach statistical significance, including Q3, Q7, Q11, Q13, Q17, Q18 and Q19. Therefore, compared with the fall semester, students in the spring semester considered the software to be more advanced in technology, and that the weekly assigned lessons can more complement their *Freshman English* course. Moreover, more students used the weekly assigned lessons; the physical environments (3.16 vs. 3.40) and computer equipment (3.11 vs. 3.42) were more positively evaluated. Likewise, learners who had filled out survey questionnaires in both semesters enjoyed learning in the SAC more in the spring semester than they did in the previous one.

In addition to the 19 items that probed into learners' perceptions of the software and physical environments, the assistant's management, benefits to their language learning, and weekly assigned lessons, students were also asked about advantages of this CALL classroom,

and they can choose more than one. The results are shown in Table 4.

Table 4. *Students' perceptions towards advantages of the SAC*

Advantages	I can choose learning materials.	I can set my own learning schedule.	I can concentrate on my learning.	I can learn with my friends.
Fall semester Number of Entry	479 (35.48%)	381 (28.22%)	238 (17.63%)	252 (18.67%)
Spring semester Number of Entry	768 (36.26%)	620 (29.27%)	394 (18.60%)	336 (15.86%)

As shown in Table 4, students in two semesters think the greatest advantages of the SAC are that they can choose their own learning materials and can set their own learning schedules. These two advantages account for approximately two-thirds of total responses. Also, students' perceptions of the areas that the SAC can improve are explored, and the results are presented in Table 5.

Table 5. *Areas that the SAC can improve*

	Computer equipment	Listening	Speaking	Reading	Writing	Movies	Testing
Fall	568 (22.98%)	227 (9.18%)	295 (11.93%)	193 (7.81%)	378 (15.29%)	629 (25.44%)	175 (7.08%)
Spring	154 (11.79%)	130 (9.95%)	213 (16.30%)	97 (7.43%)	246 (18.84%)	344 (26.34%)	122 (9.34%)

Similar trends can be observed as students in both semesters think that the movies, the computer equipment and writing software are the three areas that the SAC can improve. Despite this, it can be noticed that computer equipment had fewer entries (154, 11.79%) in the spring semester, a finding that corroborates the results from the survey questionnaire. From the responses to the item 18 in the questionnaire, it was revealed that students in the latter semester held more positive attitudes toward the computer equipment in the SAC.

Apart from being asked to evaluate their degree of agreement with each statement, learners are also encouraged to write down comments regarding the SAC. These written comments were also collected and analyzed. It was found that these written comments corroborate the findings derived from the survey items. In total, 11 comments were on language learning, for instance, "*Friends* are good! Watching it can train my listening and reading," "Pu 101 is beneficial to my English learning," and "I improved my speaking via *My English Tutor* [MyET]." Also, 5 comments concerned how students learned content knowledge, e.g. "I learned a lot of knowledge from IMAX videos." Still other 11 comments focused on the environment and resources of the SAC, e.g., "There are many resources to choose from," and "Such an environment can boost our interest in English." Other than these positive comments, suggestions were also made. These suggestions centered upon the hardware and videos, e.g. "I want videos without subtitles," "I suggest that difficult vocabulary or phrases in the video can be explained," "I want more movies," and "It is not comfortable to wear headphones."

Interview results

This section presents the principal themes and generalizations drawn from the interviews with learners. Analysis of the qualitative data could further explicate learners' perceptions of learning in the SAC. It was found that many *Freshman English* instructors set regulations of minimum number of visits per semester. While some interviewees reported this factor to be one of the reasons that kept them visiting the SAC, approximately two-thirds of the interviewees (28 out of 45) indicated that they came to SAC owing to the resources, software, and the environment of the SAC. As the following interviewees commented,

“I think this Self-Access Classroom is a place where I can concentrate on my learning. On the other hand, there are so many kinds of software that I can use, and this is convenient for my learning.” (Chen, W. –C.)

“I want to improve my English proficiency. I also think this is a good environment, and that there are many nice movies.” (Ting, J. –C.)

As for the greatest gain of visiting the SAC, many interviewees (24 out of 45) reported improvement in their listening; other linguistic areas of improvement include reading, speaking, and vocabulary. Interviewees also commented that from American sitcoms, they learned authentic daily expressions and gained more understanding about the United States:

“I can learn the daily expressions from the *Friends*, but the speed is rather fast. So, I didn't understand each word.” (Wang, Z. –S.)

“Some parts of *Friends* are involved with culture, and afterwards, I came to understand its funny aspects.” (Kuan, P. –W.)

As their listening comprehension failed, some interviewees mentioned certain compensatory listening strategies, e.g. drawing inferences from nonverbal clues and guessing:

“When I first watched *Friends*, I had no idea of what they were talking about. Later I came to understand the content via the characters' gestures, movements, and my current English ability.” (Lin, Y. –C.)

“Over one year of use in this center, I would try to guess the meaning when I encounter the vocabulary or sentences that I don't know in watching movies. And I think my listening has improved.” (Lu, Z. –Y.)

Aside from listening, interviewees pointed out that one automatic speech recognition software, *My English Tutor* (MyET), greatly helped them detect mistakes in their pronunciation and improve their speaking:

“My favorite one is MyET. It has four rating charts and would point out my mistakes one by one. The users can repeatedly listen to their voices. Because I am weak in my speaking and listening, I would imitate its intonation patterns.” (Chen, W. –Y.)

Also, when asked about any difference of the SAC in these two semesters, 21 of the 25 interviewees noticed the renovation of the software and computer equipment, and commented that regular maintenance would greatly encourage them to come to the SAC.

DISCUSSIONS

This study investigated how EFL college students perceived their learning in one CALL self-access classroom. Different sources of data, including two end-of-semester surveys, one open-ended question in which learners could write their opinions regarding the SAC, and interviews, were collected. The following part will synthesize different data and interpret the major findings in response to the research questions. Regarding the first research question of if EFL college students perceive learning in a CALL SAC to be beneficial to their English learning, it was revealed that learners reported greater gains in their listening and reading than in speaking or writing. This finding may have much to do with the nature of the activities that the learners have been engaged with. From the learning records, it was found that many students watched videos, movies, or sitcoms in the SAC, and during the process, learners could be exposed to large quantities of authentic input, and thereby proceed to their improve their listening proficiency (Plass & Jones, 2005). Moreover, interviewees reported that encountering comprehension failure, they would utilize certain compensatory strategies, for instance, guessing or drawing inference from contextual clues. This greater improvement in listening is in line with findings from Cheng (2006), Chia (2004), and Wang (2005).

As for learners' evaluations of the weekly assigned lessons, they generally agreed that these lessons can boost their learning interest, complement the inadequacy of *Freshman English* course, and increase their language proficiency. Nonetheless, learners may find the weekly assigned materials not compatible with their linguistic proficiency, nor did all the learners utilize the weekly assigned lessons. This concern of difficulty level of materials was brought up in Beatty (2003), and Cheng (2006) attempted to incorporate five hundred graded books in the self-access program, and more than half of the participants (54.1%) positively evaluated these graded materials. Therefore, more guidance of the weekly assigned lessons and facilitation may be warranted to assist learners in their self-access learning.

In response to the second research question of the reasons that keep learners coming to the SAC, it was revealed that aside from the instructor's regulation, most interviewees came to the SAC because of the availability of various learning resources as well as the cozy physical environment. As shown in the survey results, the freedom that students can enjoy to choose their own learning materials and to set their own learning schedules was reported to be two great advantages of this CALL self-access classroom. Likewise, more than two-thirds of the interviewees reported that the cozy environment and abundant resources are strong impetuses. A further investigation into students' perceptions of the language learning software shows that its content and technological aspects were positively evaluated by learners in two semesters. Compared with this positive evaluation of software, the physical environments and computer equipment were not so well appreciated by the learners in the fall semester, but a great attitudinal change was observed in the following semester. As shown in the paired-sample t-test and the interviews, the renovation of hardware and movies was much appreciated by the learners. Moreover, learners in this study positively rated the assistant's management and service. These findings illuminate the importance of maintenance and management within the SAC (Gardner & Miller, 1999; Reinders & Cotterall, 2001).

CONCLUSION

This study probed into the perceptions of EFL college students' toward learning in the SAC. The findings reveal that learners find exposure to audio and videos beneficial to their listening and reading. Gains in their speaking, vocabulary, and cultural understanding were also reported. Moreover, other than the instructor's regulation, the freedom of choosing materials and the control of their own learning schedule in a cozy atmosphere and resourceful environment are the chief reasons that motivate learners to keep coming to the classroom. The fact that the computer equipment was not favorably evaluated in the fall semester

highlights the importance of maintenance and management within the SAC, as the quality of the hardware may greatly influence learners' willingness to come to the classroom.

Even though this study triangulated data from surveys and interviews, some limitations are still present. The first one is inherent in the self-report nature of the instruments. The inclusion of objective measures may help researchers explore the effect of learning in a CALL SAC on learners' language gains. Another limitation is the disproportion of interviewee samples, particularly a lack of interviewees from the low- and basic-proficiency learners in the second semester. Future research may aim to include interviewees from different proficiency levels, in an attempt to explore if learners of different language proficiency may differently evaluate their learning in the SAC.

REFERENCES

- Beatty, K. (2003). *Teaching and researching computer-assisted language learning*. Pearson: England.
- Carson, L., & Allen, R. (2006). A self access centre in action: Facilitating student achievement of specific language goals. *ESP Malaysia*, 12, 91-100.
- Chao, C. C. (2005). 尋找英語學習的新樂園：英語自學措施分類與分析 [Encouraging self-regulated EFL learning in Taiwanese: higher education: Activity types and analysis] *English Teaching & Learning*, 29(4), 1-20.
- Cheng, W. -w. (2006). *Learner perspectives of self-access to the multimedia English learning center at NKFUST*. Unpublished master's thesis, National Kaohsiung First University of Science and Technology, Kaohsiung, Taiwan.
- Galavis, B. (1998). Computers and the EFL class: Their advantages and a possible outcome. *The Autonomous Learner*, 36(4), 27.
- Gardner, D., & Miller, L. (1999). *Establishing self-access: From theory to practice*. Cambridge: Cambridge University Press.
- Hsu, W. -C., & Xu, X. (2007). *Learner autonomy: Does it work in the exam-oriented Taiwanese secondary context?* In Department of English, National Chengchi University (Eds.), *Proceedings of the 24th International Conference on English Teaching and Learning in the Republic of China* (pp. 1-11). Taipei: Taiwan ELT.
- Jones, F. R. (1998). Self-instruction and success: A learner-profile study. *Applied Linguistics*, 19, 378-406.
- Kenning, M. -M., & Kenning, M. J. (1990). *Computers and Language Learning: Current theory and practice*. New York: Ellis Horwood.
- Klassen, J., Detarmani, C., Lui, E., Patri, M., & Wu, J. (1998). Does self-access language learning at the tertiary level really work? *Asian Journal of English Language Teaching*, 8, 55-80.
- Marquart, J. M., Li, S., & Zercher, C. (1997). *Making the whole more than the sum of the parts: Challenges in a mixed method study of inclusion*. Paper presented at the Annual Conference of the American Educational Research Association, Chicago, Illinois. (ERIC Document Reproduction Service No. 408749).
- McMurry, B. L. (2005). *Self-access centers: Maximizing learners' access to center resources*. Unpublished master's thesis, Brigham Young University.
- Miller, L., & Rogerson-Revell, P. (1993). Self-access systems. *ELT Journal*, 47(3), 228-233.
- Morrison (2005). Evaluating learning gain in a self-access language learning centre. *Language Teaching Research*, 9(3), 267-293.
- Ning, S. -j. (2008). *EFL college students' perceptions of self-access language learning at a CALL center*. Unpublished master's thesis, National Taiwan Normal University, Taipei, Taiwan.
- Plass, J. L., & Jones, L. C. (2005). Second language acquisition with multimedia. In R.

- Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 476-488). New York: Cambridge University Press.
- Reinders, H. (2000). *Do it yourself? A learners' perspective on learner autonomy and self-access language learning in an English proficiency program*. Unpublished master's thesis, University of Groningen, the Netherlands.
- Reinders, H., & Cotterall, S. (2001). Fortress or bridge? Learners' perceptions and practice in self-access language learning. *Tesolanz*, 8, 23-38.
- Ridley, J. (2000). Towards autonomy in university classrooms: The role of learners' goals. In D. Little, L. Dam, & J. Timmer (Eds.), *Why Focus on Learning Rather than Teaching* (pp.3-15). Dublin: CLCS, Trinity College Dublin. ‘
- Riley, P., & Zoppis, C. (1985). The sound and video library. In P. Riely (Ed.), *Discourse and Learning* (pp.286-298). London: Longman.
- Sheerin, S. (1991). State of the art: Self-access. *Language Teaching*, 24(3), 143-157.
- Sturtridge, G. (1997). Teaching and learning in self-access centers: Changing roles? In Benson, P., & Voller, P. (Eds.), *Autonomy and Independence in Language Learning* (pp.66-78). London: New York.
- Voller, P., Martyn, E., & Pickard, V. (1999). One-to-one counseling for autonomous learning in a self-access center: Final report on an action learning project. In S. Cotterall & D. Crabbe (Eds.), *Learner autonomy in language learning: Defining the field and effecting change* (pp.111-126). New York: Peter Lang.
- Walker, R. (2005). Self-access learning, Module 1.4, *Information and Communications Technology for Language Teachers (ICT4LT)*. Slough, Thames Valley University. Retrieved April 2, 2009 from <http://www.ict4lt.org/en/>
- Wang, L. (2005). *EFL college students' perceptions of culture learning at a CALL center*. Unpublished master's thesis, National Taiwan Normal University, Taipei, Taiwan.
- Zou, X. (2006). Learners' autonomy in self-access centers – An exploration of the SAC in Chongqing University. *Sino-US English Teaching*, 3(9), 1-7.

APPENDIX

End-of-semester Survey Questionnaire

	Strongly Agree	Agree	Disagree	Strongly Disagree
1. I feel satisfied with the language learning software in the SAC.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. In terms of content, I think the SAC provides practical language learning software.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. In terms of technology, I think the SAC provides advanced language learning software.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. In terms of listening training , I think the SAC provides sufficient language learning software.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. In terms of speaking training , I think the SAC provides sufficient language learning software.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. In terms of reading training , I think the SAC provides sufficient language learning software.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. In terms of writing training , I think the SAC provides sufficient language learning software.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. I think that the weekly assigned lessons fit in with my proficiency.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. I think that the weekly assigned lessons increase some aspects of my English ability.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. I think that the weekly assigned lessons can boost my interest in English.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. I think that the weekly assigned lessons can complement the inadequacy of <i>Freshman English</i> .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. Even if I have finished this week's assigned lessons, I would still visit the SAC and use other software.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. I used the weekly assigned lessons in the SAC.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14. I think the assistant is competent for this job.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. I think the assistant is enthusiastic to serve for me, for instance, in software operation.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16. I think the assistant tries his best to solve my questions.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17. I think the physical environments of the SAC are suitable for language learning.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18. I think the computer equipment (including headphone, and microphones) is suitable for the software operation.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19. Overall, I enjoy learning English in the SAC.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20. I think learning English in the SAC has the following advantages:				
<input type="checkbox"/> I can choose learning materials.	<input type="checkbox"/> I can set my own learning schedules. <input type="checkbox"/> I can learn with my friends.			
<input type="checkbox"/> I can concentrate on my learning.				
<input type="checkbox"/> Others				
21. In what areas do you think the SAC can improve?				

- | | | |
|---|---|--|
| <input type="checkbox"/> Computer equipment | <input type="checkbox"/> Software (Listening) | <input type="checkbox"/> Software (Speaking) |
| <input type="checkbox"/> Software (Reading) | <input type="checkbox"/> Software (Writing) | <input type="checkbox"/> Software (Movies) |
| <input type="checkbox"/> Software (Testing) | <input type="checkbox"/> Others | |
-

22. The software that I like the most:

- Reasons :
- | | |
|---|---|
| <input type="checkbox"/> Learning content | <input type="checkbox"/> Language ability |
| <input type="checkbox"/> Interface design | <input type="checkbox"/> Software operation |
| <input type="checkbox"/> Others | |
-

23. The software that I dislike most:

- Reasons:
- | | |
|---|---|
| <input type="checkbox"/> Learning content | <input type="checkbox"/> Computer equipment |
| <input type="checkbox"/> Interface design | <input type="checkbox"/> Software operation |
| <input type="checkbox"/> Others: | |
-

Listening, Speaking and Reading for Writing: Holistic Pedagogies for an Advanced Writing Course

Grace H C Lin (林慧菁), Max M H Ho (何明軒), Paul S C Chien (簡士捷)

National Sun Yat-sen University

lingrace@faculty.nsysu.edu.tw
kilakilamax@yahoo.com.tw
paulchien@faculty.nsysu.edu.tw

This study provides a series of pedagogies conducted in an advanced writing course of senior English majors at a national university in southern Taiwan. The holistic approaches were incorporated into teaching and learning in class. Through a semester, various pedagogies based on the theory of holistic teaching and learning mentioned by Blanton (1992) and Sauve (2002) were designed for 16 participants. Teaching English writing was conducted through guiding discussion activities, instructing reading and writing strategies, and demonstrating model articles on the Internet and assigning learners one minute presentation with writing topics, playing games, and composing essays at home.

Based on the theory proposed by Celce-Muricia (1991), a linear acquisition sequence of skills of listening, speaking, reading and writing has led curriculum designers and language educators to introduce these four types of language abilities and to assist students in integrating these four modes of the target language. Indeed, L2 writing competence must not be isolated from the other three kinds of language competence. This study argues that a writing course should be a more wide-range training for the purpose of integrating four types of language proficiency in English into the final goal of English learning, practically using it as a tool of interaction in our daily life.

INTRODUCTION

During the last two decades, a holistic approach has gained prominence in studies of second language acquisition (SLA). One central importance is the holistic approach in which the four skills of a language (i.e. listening, speaking, reading and writing) were regarded as a whole, which is mutually exclusive with an atomistic approach to language. An atomistic approach attempts to analyze language into components, like grammar, vocabulary, pronunciation, or purposeful models which can be defined as independent content of different curricula. In contrast, the holistic pedagogies emphasize the significance of ability integrating components of various types of English proficiencies. Holistic pedagogies focus on making proper correlations among all proficiencies of English and truly apply English as tool of communication through speaking and writing.

Although a holistic approach might have different foci from an atomistic approach, a growing number of language educators and scholars (e.g., Blanton, 1992; Brown, 2001; Sweedler & Carol, 1993) attribute greater importance to a holistic approach. The definition of holistic pedagogy can be defined in either a simple way or a rather complex one. Brown (2001) defines a holistic approach as the one which aids students in integrating multiple dimensions of language skills, contending that the integration of the four skills appears to be the only possible approach in the theoretical framework of communication and interaction. As it can be seen, the definition of holistic teaching traditionally tends to emphasize the

integrated type of teaching and learning with four skills, including listening, speaking, reading and writing. However, from a more contemporary aspect, the whole person's intellectual, physical, social, emotional, spiritual, and environmental factors might be much more significant than the intellectual way of knowing that the teacher offers the learners (Schiller, 2007). In Schiller's (2007) terms, holistic teaching and learning should take more psychological components into consideration, such as Krashen's (1982, 2003) concern of emotional block (termed affective filter).

A holistic approach might take different forms inasmuch as different scholars with different worldviews might not interpret the same thing consistently. Despite the variances of definitions, a rich variety of merits in a holistic approach have been recognized. As described by Blanton (1992), a whole-language approach integrates language and content, thus helping students make best of their time by focusing on one type of language and related language skills. In the same vein, Brown (2001) suggests that use of one type of language skill reinforces another, such as writing and reading as well as speaking and listening, introducing various types of instruction which can foster integrated skills, like content-based instruction, theme-based instruction, experiential learning, the episode hypothesis, and task-based teaching.

Recently, there has been growing concentration in application of the holistic approach to adult education in second language (L2) writing. Egyptian scholars Mohamed and Ahmed (2006) note that "Although all of these skills are important, interrelated and can not be separated, yet developing the writing skill with all its difficulty has been neglected for a long time due to the adoption of traditional approaches of language teaching that mainly focus on speaking and listening." (p. 3) Due to the lack of the holistic type of teaching, students can not develop their English verbal and writing proficiency in a balanced mode, which four types of competence should be not too diverse. In order to resolve the orientation that students are better in speaking than writing, Mohamed and Ahmed (2006) recommend that the holistic form of teaching as well as the whole language approaches should be applied to tutor writing. In this study that they conducted, a significant enhancement in writing was achieved through training with the updated holistic and whole language type of teaching and learning. Not only were students' improvements in writing proficiency accomplished, the teacher and students' relationship was found to be much better through the non-threatening learning milieu that holistic and whole language type of teaching had created.

In China, Qiang and Wolff's (2008) study, a holistic English curriculum is created and applied. The findings show that the holistic approach appears to provide students more comprehensible input through observation, listening, reading, writing, debate, conversation, and Internet research, thereby allowing learners to have more opportunities to employ and practice different types of language skills at the same time. What is more, this approach seems not merely able to foster self-confidence and intrinsic motivation but also to build up autonomous learners and creative thinkers.

This study conducted at a national university in Kaohsiung, was designed through applying various factors and skills in order to actually show positive evidences of the updated holistic pedagogies. Indeed, in the second millennium, people on the earth are trying to be assimilated and cooperate with each other; we need to integrate our different types of English proficiencies to be a more real ability in communication, through being trained by a holistic pedagogy. Only through the holistic mode, our learners might have an opportunity to integrate and assess their listening, speaking, reading and writing abilities. More specifically, this study implies that writing competence should be integrated with speaking, listening, and reading, along with logical brain storming, and makes learners enthusiastically engaged in creating and sharing ideas for various kinds of learning purposes. Based on the holistic approach, "...students become intellectually/ cognitively engaged in language and content."

(Blanton, 1992, p. 291) Then, students might become more communicative and internationalized under the current trend of globalization.

RESEARCH QUESTIONS

The primary research question that had guided this study was “How effective did learners evaluate the holistic approach?” Furthermore, “How much the advanced-level learners can improve after the treatments of holistic type of pedagogies?” In other words, this study aimed to discover to what level the learners of Taiwan would like to accept the updated type of teaching. On the other hand, we would like to know if the holistic teaching would contribute to develop the advanced-level learners’ space in improving their composing ability.

ASSUMPTIONS

1. The assumption of this study is that the students would support the holistic types of pedagogies applied in the writing course.
2. Students might not cooperate with the updated pedagogies and answer the survey questions because they were not willing to interact and make any presentation in English, which were not traditional types of examinations.
3. The treatments of holistic pedagogies might make student improve a lot in their writing proficiency.
4. The treatments of holistic pedagogies might not have a noticeable improvement in their writing proficiency.
5. The scores before and after complete treatments of holistic pedagogies might not be significantly different.
6. The scores before and after all-encompassing treatments of holistic pedagogies might be significantly different.

METHODOLOGY

Celce-Murica (1991) has mentioned in the field of ESL teaching and learning, a linear acquisition sequence of four skills 1st: listening, 2nd: speaking, 3rd: reading and 4th: writing has guided most pedagogists instruct these four skills. Writing indeed is a complex task for the teachers to teach and for learners to develop. In order to strengthen second language learners’ writing competence, the order of four skills for integrating propensity of English have to be followed in order to well organize the effective pedagogies in the writing class. A writing itinerary should not only involve the teachers’ professional instruction of writing, but also it should include training for listening, speaking, and reading proficiencies as well. Through reviewing these four skills of English, the last skill as well as the most complicated skill, writing could be constructed. Writing expertise can be accomplished in a more effectual and well-organized system if a lecturer can follow Celce-Murica’s (1991) sequence of linear acquisition. For this reason, this study attempted to pursue the emphasized progression and designed numerous activities.

Research Setting

This study was conducted at National Sun Yat-sen University, where 16 English majors in the Department of Foreign Languages and Literature took their advanced writing course for eighteen weeks. It is a required program for graduation; the students have to pass writing one and writing two in order to take this writing level three, which is the most advanced writing class in the requirements for graduation. In this class, there are three male undergraduates and thirteen female ones. One of them was an apprentice who re-took this class since she failed the same course in the previous year. Two of the English majors were overseas Chinese from Malaysia and the Union of Myanmar, and the others were local

Taiwanese students, whose first languages were Taiwanese, Mandarin and Hakka respectively.

Besides, there were four French students from exchanging school, who from time to time came to the class and scrutinize the trainer's oration. These students were exchange students from Paris, Dijon and Niece, involving three males and one female. This group of foreign students had visited the class in its 1st and 2nd sessions. Although they had participated a few activities conducted during that period of time, their performances and test grades were not evaluated and included in this research project. The reason is that the researchers were not able to collect wide-ranging data for comparing and contrasting the effectiveness.

Listening articles on the Internet

The holistic approach applied in the first session was to help students be able to listen and read at the same time, through reading the articles on the Internet and the teachers' instructions and recitations in class. Through reading composing strategies posted on web sites at Harvard, Texas A&M, and University of Southern California, the trainer vocally instructed a series of writing strategies. At the same time, the students were able to eavesdrop and understand writings on the screen provided in class and unreservedly interact with the writing trainer in order to gain a deeper understanding about the knowledge of writing. Sometimes the trainer assigned a subsection of instructions of writing strategies for a student to read and the other students to take note.

With the apparatus equipped inside the university's classroom, the electricalized type of teaching contributed to the smoothness and efficacy of teaching and learning in the writing course. In this session, knowledge of writing about transition, summary, counterargument...etc, posted in a series of free e-learning web-pages were examined by the course members and interpreted by the trainer. In fact, during this period of training, three of class members came to the trainer during the class break and revealed that they considered the resources of Harvard College to be very intellectual, motivation-stirring, and obliging. Actually the resource of the top school had been applied by the researchers due to the concern of establishing the learners' self-assurance.

Speaking with topics of TOEFL

In the second session of writing course, the instructor applied a list of writing topics in Test of English as a Foreign Language (TOEFL) in order to carry out an activity of brain storming. Course members were divided into pairs of eight and requested to contribute their thoughts and proposals for TOEFL writing topics. Students were given guidelines, implications and modification by the instructors when they met difficulties in critical thinking, grammatical problems of structures or pronunciations...etc. After fifteen minutes of sharing ideas, students were requested to engage in a balloon game for providing ideas. The trainer started the game as the first speaker and demonstrated the short presentation for a minute. After that, the balloon was thrown to the second (next) member of the class, who was proposed to give another one-minute presentation based on the contents of brainstorming. This activity lasted for around thirty minutes, which almost all of the class members had to have a word and speak out.

Reading the trainer's articles

In the third session of this writing course, the students were distributed five articles written by the trainer, including thesis based on topics of pragmatics, teaching English as foreign language, cultural study, and linguistics. They were requested to comprehend the published articles in different levels and to give comments as if they were the editors of journal institutions. All course members were requested to do proofreading and bestow commentaries for the published papers through writing reflections on a page. In this session, trained learners had applied their reading and writing skills. According to the class members'

comments and feedback returned to the trainer, the researchers determined that almost all the learners had made sense the meanings contained in the trainer's publication. Although most of the reading assignments were done individually and independently away from the classroom, the researchers of this study were able to check their understanding of the distributed readings since their commentaries were precise and correct.

Writing essays through reading trainer's work and peer scaffolding

The final session of this holistic curriculum was a program, which suggested that students should accumulate sufficient competences in reading, listening and speaking for writing. Therefore, in the last few classes, students were requested to perform their holistic ability in essay writing based on TOEFL topics and based on proofreading the trainer's published papers. In order to stimulate the course members' intrinsic motivation for achieving superiority and perfectness in their integrated abilities of English, the instructor made students freely provide their compositions without limiting their point in time and quantities of submitting works. Hopefully, without coercing them to offer their homework unswervingly to their lecturer, but their peer first, the advanced-level learners would be able to activate their brain more relaxingly without anxiety in a habitual fashion.

It was anticipated that alternative type of writing pedagogy, peer scaffolding suggested by Lee (2008) and Sotillo (2005), would carry out the learners' preeminent presentations in writing articles. A great deal of progress might be achieved due to their relaxing emotion from knowing their articles. This could be upgraded by their peers before submitting it to the lecturer. Moreover, it was also expected the course members could make senses for the trainer's intricate papers, creating comments for proving abilities of reading in order to establish their confidence of writing, before writing for their peers' proofreading. In fact, taking students' psychological factors of learning into account, Krashen has repeated that learners' confidence should be established. In order to attain this point and reduce the learner's learning apprehension, Schiller (2006) has also contended that the professors should make students revise the articles of their classmates before the instructor corrects them.

Data Collection

In this study, qualitative data was collected. At the end of the students' learning, the instructor distributed a piece of survey sheet based on the holistic types of pedagogies that had been conducted in the class, and asked students to fill out their reflections through numbers. The research questions of this study were primarily about students' reflections of the holistic type of teaching for the writing class and the pedagogies applied in the holistic approaches. In order to diminish potential negative influence on students' learning, the researcher found only nine volunteers instead of a whole class were willing to fill out the survey in five minutes during the class break.

Secondly, students' performances before and after the entire treatments of holistic pedagogies were quantitatively measured. Their basic writing proficiencies were tested through two assignments of five-paragraph techniques of writing. With subject matters from the TOEFL writing topics, their presentations in their midterm and final examination (before and after complete) holistic training were evaluated. This study was anticipated to see the perceptible differences before and after complete training through holistic type of pedagogies. That is to say, the researcher of this study expected to see students make progress in their writing ability because of being trained in an alternative pedagogy.

Instruments for Data Collection

Firstly, a consent page with survey questions was allocated to every member in the advanced writing course. In order to make the class members honestly and freely answer the questions associated with the effectiveness and efficiencies of holistic pedagogies, the researchers of this study did not compel the learners to collaborate and sign the agreement; therefore, only nine of all class members returned their survey sheets to the researchers. For

the purpose of not influencing learners' effectiveness of learning due to wasting time filling out forms, the survey questions only included seven items, which were related to barely primary methodologies applied in their advanced writing course.

In addition, the second instrument is the TOEFL rating criteria designed by ETS. There are points from 6 to 0 for describing seven levels of the writers' achievements. For the highest level 6, the trainer and the rater of this study applied a range of twenty points to more precisely measure the learners' performances. Since the learners in this class were all highest level as well as the advanced-level English majors, who had taken 12 credits of English writing courses of intermediate and basic levels, their performances can not be in the lower two levels of 0 and 1 designed by ETS. Hence, these two levels were considered and designed to be the same levels of twenty points, although they should not be applied in this group of writing learners. Based on the ETS criteria, the students' written texts were calculated by the trainer as well as the rater, and the grades were given to the students for self-inspection.

LIMITATIONS

1. The limitation of this study is that the some students might not support the holistic types of approach due to their intentions for traditional examination preparation. They did not fill out surveys cooperatively. Since seven of class members did not return the survey sheets, the conclusion of holistic pedagogies might be a bit biased and too optimistic.

2. The limitation of this study is that some students might not answer the question honestly because they worried that their scores graded by the researcher (as well as their teacher) would be lower if they did not support the applied holistic approach. In other words, volunteers might always answer in a positive way, because they might want to please the researchers as well as the trainer in order to gain a higher final score.

RESULTS AND DISCUSSIONS

The results of this study based on qualitative and quantitative investigation both showed that the holistic types of writing pedagogies are important and necessary. Positive attitude toward the applied pedagogies through listening, speaking, reading, and actual writing was found and declaimed by volunteers. At the same time, learners' performances between incomplete holistic teaching and complete holistic teaching were evaluated, which demonstrated the complete one contributing to the learners' improvement in a significant way. The following two paragraphs will more clearly interpret the collected data.

Based on Qualitative Data

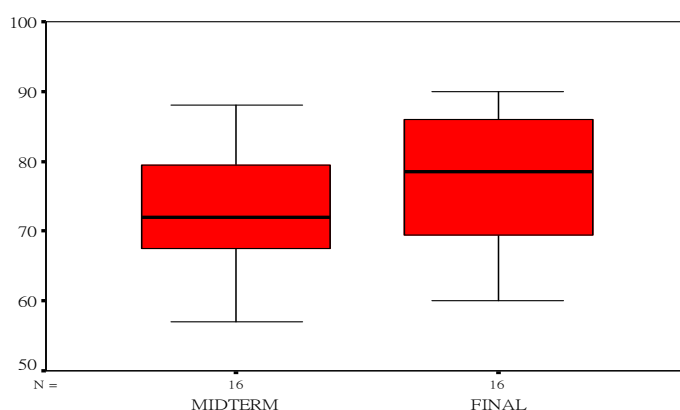
After analyzing the qualitative data collected from nine volunteers, the result of this study revealed that most students welcome the trainer's holistic approaches applied in class. Seven students expressed that through reading, listening, and speaking, they could make more progress in writing. Five students (> 50%) believed especially for advanced writing, reading, listening, and speaking should be contained in the writing course. Eight out of nine students believed that four skills should interactively influence each other. They were of the same opinion that the holistic approach should be applied. About the activities conducted in the classroom, seven students (> 50%) students expressed that they could absorb the teachers' recitations for the writing strategies on Harvard's website. Five students (> 50%) felt benefited from teacher and classmates' lectures in class. Six students agreed that one minute presentation based on TOEFL writing topics was useful for making progress in their writing.

Based on Quantitative Data

Moreover, based on the performances before and after complete holistic types of pedagogies, the class members' scores in writing essay based on TOEFL topics revealed the effectiveness of holistic pedagogies in this advanced writing course. In the middle of the

semester, students had experienced learning writing through listening and speaking. Their average grade of writing evaluated through criteria (see Appendix B.) designed by ETS was seventy-three. A significant difference in the second measurement after experiencing listening, speaking, peer reading, and reading teachers' works, was with an average score of 77.

Based on the graphs analyzed by Statistics Package for Social Science (SPSS), it is uncomplicated to discover, the holistic pedagogies did assist students make a certain level of progresses in their writing proficiency. The statistical data in the T-test revealed that there was a significant difference before and after complete holistic training, since the p-value was close to zero. In other words, dissimilarity up to four points between two means in midterm and final examination reveals that the holistic pedagogies through listening, reading, peer correction, and writing are apparently effective and efficient.



Graph 1. Box plots displaying significant differences between two sessions of complete and incomplete holistic pedagogies

One-Sample Test

	Test Value = 0					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
MIDTERM	32.647	15	.000	73.00	68.23	77.77
FINAL	33.794	15	.000	77.00	72.14	81.86

Graph 2. Statistical numerical data revealing the significant grade difference between two sessions of incomplete and complete holistic pedagogies

CONCLUSION

From the results of the survey questions, this study has revealed that students tended to support the holistic approaches applied in the writing course. There is no activity conducted in the writing course and mentioned in the surveys which were rejected by the students. All of them were supported by more than fifty percents of the volunteers. Hence, this study argues that holistic approach should be applied in teaching writing itinerary, and also it can be accomplished through Celce-Muricia's (1991) linear acquisition sequence.

Furthermore, from the measurements of before and after complete holistic types of training; class members' performances were notably different. This is the best illustration that holistic types of pedagogies for writing need to be included and even promoted. As English teachers, we usually are worried about how much improvement students would make after

taking our class. With the holistic type of the curriculum design, we do not need to worry to any further extent about students might squander time due to taking our class. In other words, if we can teach through four types of English skills and make our learners relaxingly immerse in the English environment, we do not have to be afraid that we will dissipate our learners' precious time taking out class. The reason is that their improvement might be predictable after holistic types of training.

Unflinchingly, autonomous learning might be stimulated through teacher's concern for the learners' psychological status. A holistic approach based on Schiller's theory, some games or music has to be applied in order to stir up the learners' learning motivation. Moreover, based on Blanton's (1992) theory, integrating language and content is a more meaningful way of teaching and learning. The four skills of language should be developed and established in a balanced manner. Students should not be taught in traditional ways, i.e., to teach four skills separately without integrating them into meaningful English interaction. In sum, this study emphasizes that listening, speaking, and reading should be also taught in a holistic manner, in a writing course in order to integrate students' language expertise into their writing aptitude.

REFERENCES

- Blanton, L. (1992). A holistic approach to college ESL: Integrating language and content. *ELT Journal*, 46(3), 285-293.
- Brown, H. (2001). *Teaching by principles: An interactive approach to language pedagogy*. New York: Longman.
- Celce-Murcia, M. (1991). *Teaching English as a second or foreign language*. Boston, MA: Heinle & Heinle.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.
- Krashen, S. (2003). *Explorations in language acquisition and use*. Portsmouth: Heinemann.
- Lee, L. (2008). Focus-on-form through collaborative scaffolding in expert-to-novice online interaction. *Language Learning & Technology*, 12(3), 53-72.
- Mohamed, A., & Ahmed, A. (2006). *The effect of using the whole language approach on developing some composition writing skills in English for experimental secondary students in Egypt*. Helwan University, Cairo, Egypt.
- Qiang, N., & Wolff, M. (2008). China EFL: What is holistic English Retrieved February, 9,, 2009, from <http://www.usingenglish.com/esl-in-china/what-is-he.pdf>
- Sauve, V. (2002). Storyweavers: Holistic education for ESL/EFL learners. *TESL Canada Journal*, 20(1), 89-112.
- Schiller, S. (2007). *Creating the Joyful Writer: Introducing the Holistic Approach in the Classroom*. Lanham, MD: Rowman & Littlefield Education.
- Sotillo, S. (2005). Corrective feedback via instant messenger learning activities in NS-NNS and NNS-NNS Dyads. *CALICO Journal*, 22(3), 467-496.
- Sweedler, B., & Carol, O. (1993). The effects of ESL errors on holistic scores assigned by English composition faculty. *College ESL*, 3(1), 53-69.
- Writing Resource from Harvard College, Retrieved on November, 1, 2008 from <http://www.fas.harvard.edu/~wricntr/resources.html>
- Writing Resource from University of Southern California, Retrieved on November, 9, 2008 from <http://college.usc.edu/writingcenter/>
- Writing Resource from Texas A& M University, College Station, Retrieved on November, 11, 2008 from <http://writingcenter.tamu.edu/>

Appendix A. **Consent form and Survey Sheet**

Consent form for Data Collection

My name is _____. I am a volunteer who have willing to provide answers toward my teacher Grace Hui Chin Lin's class. The data that I offered will not be displayed with my name in order to protect my privacy. The survey with data will be located in Grace Lin's office for five years.

Survey for Holistic Pedagogies applied in Writing III

- A. Agree
- B. Disagree

1. I have trained through speaking in one minute prompt lecture with topics of TOEFL. I think this is helpful in improving my composition ability.
2. I have read many articles including my teacher's writings, writing strategies through Websites of Harvard University and Texas A&M University. I think this is helpful to my progress of writing thesis.
3. I have listened to my teacher's and classmates' lecture focused on one subject. I think this is helpful to my ability of expressing and writing.
4. I think writing should be a combined proficiency of all four English skills, so I enjoy the holistic pedagogies that my teacher had applied in my advanced writing class.
5. I feel through "reading", "listening", "speaking," I can make more progress in my writing.
6. Because of the four skills should interactively influence each other, I agree to be taught through the way my lecturer is conducting her class.
7. I feel reading listening and speaking should be contained in the writing class especially for the highest level of writing.

Appendix B. **Criteria of Writing Assessments for Midterm and Final examination**

TOEFL Writing Section Scoring Criteria

There are scores, relating to the Writing portion of the TOEFL Ñomputer based and Paper based TOEFL test, provided below.

Your essay will receive a score of 6 (100) in case if:

- It addresses the writing task in an effective way.
- The essay can be considered as well organized and well developed.
- There are clear and appropriate details, which provide strong support to the thesis statement, and / or which illustrate the main ideas, used in the essay.
- The essay displays consistent facility in the application of language.
- There is a syntactic variety and appropriate word choice, demonstrated in the written TOEFL essay.

Your essay will receive a score of 5(80) in case if:

- Some parts of the task are addressed in more effective way, in comparison with the others.
- The essay can generally be considered well organized and developed.
- There are details, which provide support to the thesis statement, and / or which illustrate the main ideas, used in the essay.
- Your TOEFL essay displays facility in the application of the language.
- There is at least some level of syntactic variety and range of vocabulary, demonstrated in the essay.

Your essay will receive a score of 4 (60)in case if:

- Parts of the task do not correspond to the topic of the essay, although in general it addresses its topic in quite an adequate way.

- The essay can be considered to be organized and developed adequately.
- There are details, which provide support to the thesis statement, and / or which illustrate the main idea, used in the essay.
- There is an adequate, although somewhat inconsistent facility with syntax, as well as there can be contained some errors, which can obscure the meaning occasionally, in your TOEFL essay.

Your essay will receive a score of 3 (40) in case if:

- The essay is considered to be organized and developed in an inadequate way.
- It is either inappropriate or there are no sufficient details, which can be used to support or illustrate generalizations, in the TOEFL essay.
- The choice of words or of word forms is noticeably inappropriate.
- There are multiple errors in the structures and / or usage of sentences.

Your essay will receive a score of 2 (20) in case if:

- There is a serious disorganization and / or underdevelopment noticed in the TOEFL essay.
- There are little or no details, which provide support to either thesis statement or which illustrate the main idea of the essay.
- The errors, done in sentence structure or usage are either serious or frequent, or both.
- There are serious problems with the ability to get focused on the main idea of the essay.

Your essay will receive a score of 1 (0) in case if:

- The essay is considered to be incoherent and underdeveloped.
- There are severe and persistent writing grammatical, syntactical and other types of errors, contained in the essay.

Your essay will receive a score of 0 (0) in case if:

- There is no response, contained in the TOEFL essay.
- The essay does not copy the topic, or if it copies it merely.
- The TOEFL essay is considered to be off its topic.
- It is written in some foreign language or consists only of keystroke characters.

字典音標教國小學童英語發音：以協同行動研究為例

Ho-hui Lin(林和蕙), Samuel H. Wang (王旭)

Fortune Institute of Technology
Yuan Ze University

francesintaiwan@hotmail.com
onghiok@saturn.yzu.edu.tw

自教育部於民國 90 年正式宣佈國小五、六年級開始全面實施英語教學後，雖提昇了小學生的英文能力，但也產生部分學生用注音符號標示英文發音的嚴重問題。目前台灣只有少數的研究著重在小學的發音教學成效上，因字典音標有讓學習者容易辨識，不易造成混淆，和協助記憶拼音等好處，所以，用字典音標作為增進小學生發音之相關研究也就格外重要。

本研究為量化與質化混合型，透過實驗教學與行動研究，來探討字典音標做為小學生英文的啟蒙音標的可行性。研究發現部分除了比較字典音標、KK 音標與自然發音等三種英語發音之實驗教學成效外，也將探討任教師執行協同性行動研究時，所發現之教學問題及可行的解決方案。另外，在會中也將分享部分老師之行動研究報告，藉此呈現老師發現問題及嘗試解決問題過程。

背景

教育部從民國 87 年開始籌備，計劃讓學生正式學英文的時間，能夠提早從國小開始，以取代傳統上從國中一年級才開始學英文的作法 (Butler, 2004)。經過多年的規劃，教育部於民國 90 年正式宣佈全面自五、六年級開始實施英語教學。實施至今，小學生的英語能力，整體上是有了適度的提昇，然而也相對的產生了小學生以注音符號標示英文發音的問題。教育部對於這些問題，至今尚無理想的解決方案。

目的

本研究之目的，在於透過行動研究與兩年的實驗教學，來比較字典音標、KK 音標與自然發音等三種英語發音教學成效之異同，進而探討字典音標做為小學生開口說英文的啟蒙音標的可行性。另一方面，也經由教師集體協同的行動研究來探討將字典音標運用於教學的問題及可行的解決方案。本計劃結果將提供國內國小英語教師及教育行政決策單位參考。為達到此目的，本研究問題可分為可分為學生學習成就及教師協同性行動研究兩方面。經由比較字典音標、KK 音標與自然發音組學生學習成就來分析哪一種發音方法較易幫助小學生學習英語，同時也藉由分析各組授課教師所執行之行動研究，來看出其所遭遇之問題和可能解決之道。

學生學習成就方面：

1. 在音素切割方面，學生的學習成就是字典音標、KK 音標還是自然發音法較佳？
2. 在字詞辨認方面，學生的學習成就是字典音標、KK 音標還是自然發音法較佳？
3. 在字母拼寫方面，學生的學習成就是字典音標、KK 音標還是自然發音法較佳？

佳？

4. 在一般字認讀方面，學生的學習成就是字典音標、KK 音標還是自然發音法較佳？
5. 在音標認讀方面，學生的學習成就是字典音標、KK 音標還是自然發音法較佳？

教師協同性行動研究方面：

1. 字典音標、KK 音標及自然發音法各組教師在實驗教學中時，發現到哪些問題？
2. 當老師發現這些問題時 他們是如何嘗試去解決的？

文獻探討

在國小實施英語教學之初，國內許多學者指出，國小英語教學主要方向是在培養學習英文的興趣以及簡易的聽說能力（王啟琳，民 90；施玉惠、周中天、陳淑嬌、朱美惠，民 87；謝良足，民 89），之後教育部也在〈國中小英語教學手冊〉及〈中小學九年一貫課程綱要〉等文件，明定國小英文課程目標是以聽說教學為主，讀寫為輔，以培養學生基本的溝通能力（教育部，民 92，民 95）。

為配合以上的英語教學政策，許多學者、老師們紛紛建議國小的發音教學以自然發音法（字母拼讀法）為主要啟蒙學生開口說英文的方式。在以自然發音法為主的發音教學，學生是不用學習或背誦一套發音符號，而是藉著經由英文字母與字音的關係（王旭，personal communication, July 31, 2007；王啟琳，民 92；張玉玲，2004；謝欽舜，1994；謝良足，民 89），讓學生記憶背誦字母與單音之間的對應規則，以讀取新字（如 c, c, c, [k], [k], [k], a, a, a, [æ], [æ], [æ], t, t, t, [t], [t], [t] 用來學習 cat 這個字）。許多學者都期望藉著這樣的學習方式，讓初學英文的小朋友，在熟悉英文字母之後，可以在較輕鬆的環境之下，引發他們的學習興趣。

這樣的方式在初學時，是可達到口說英文某種流利的程度，但是隨著單字的變長及逐漸複雜，在學習單字與發音的應對規則時，便產生了許多嚴重的問題。如民國 92 年立法委員孫大千先生就曾在立法院質詢時提出，不少小學生有用中文勺勺口標示音英文的問題（王正寧，民 92；鄭任汶，民 92；韓青秀，民 92；中國時報，民 92），他指出「五六年級小學生沒有輔助音標可以輔助發音，只好採取自力救濟方式以注音符號注音」（中國時報，民 92），他並且指出「國小學生以注音符號記英文發音，唸出來怪腔怪調，幾乎無人能解」（張武昌，民 92）。

針對這個小學生使用中文注音符號來標示英文發音的問題，早在正式實施全面國小英語教學之前，就被許多學者發現了徵候。許多學者老師們陸續建議，學生似乎仍可使用 K K (Kenyon & Knott) 音標，作為輔助發音的工具，於是許多高年級學生（五、六年級）學習 K K 音標，以作為輔助自然發音法（字母拼讀法）的發音的研究，就陸續的展開了。

在教學教法方面，學者各提出不同的建議。有的學者建議，小學生可在學習英文字母兩年後，開始學習 K K 音標（韓青秀，民 92；謝良足，民 89），以輔助教學，有的則是建議在音標教學時，將子、母音搭配教授，比子、母音分開教授，可獲得較好的學習效果（葉麗雪，民 94，頁 70）。

在實驗教學的研究方面，對於學習成果，學者則有不同的發現。謝良足自民國 89 年開始，以五、六年級學童兩年的實驗教學研究方式，探討比較全英語、中英文混合、K K 音標、自然發音、自然發音加 K K 音標，以及 K K 音標加自然發音等六組教學之成效（謝良足，民 89，民 90），結果發現，其中 K K 音標加上自然發音的教學組，期末學

習評量較優。但是，最近則有葉麗雪於民國 94 年探討 K K 音標輔助英語教學，在經過對六年級學生實驗教學，以字母拼讀法、字母拼讀法輔以 K K 音標(子母音分別介紹)及字母拼讀法輔以 K K 音標(子母音搭配介紹)發現自然發音組和自然發音+K K 組等兩組學生之間的英語學習成就，並無顯著差異(葉麗雪，民 94)。

陸續也有學者從語音學及學習者的角度表示對 K K 音標的質疑。因為 K K 音標本身是一種音標系統，與傳統的英語拼音方式分離。另言之，K K 音標本身即是另一套符號，對學習者容易造成記憶上的負擔(王旭，民 77，頁 61；林淑玲，民 92；張武昌，民 92)，且對學習者造成相同符號、不同念法的混淆與苦惱(王旭，民 77，頁 63；張武昌，民 92；趙碧芬，民 95)，因而大大降低了語言學習的效率。

也有許多實際從事兒童美語教學的老師們，在網路上或是書刊中表達對這個問題的關心。敦煌書局親子學習報導網站作者趙碧芬就用淺顯的例子來敘述以 K K 音標來因應小學生注音符號的結果。她說：「以小學 K K 音標來因應小學生注音符號的問題，則好比是牙痛找來郎中燻耳朵抓蟲一般」(趙碧芬，民 95)，她指出，這是因為 K K 音標與字母十分相似，但是相同的母音符號在 K K 音標的念法卻是完全不相同，例如字母 e 在英文裡發音為「一」，而在 K K 音標中，卻發音為「ㄟ」(趙碧芬，民 95)，如此的情形，非常容易產生 K K 音標與英文發音的混淆。就是在以上所提及的負擔與混淆的問題下，許多小學生在面對字母與 K K 音標混淆以及需要記憶背誦一套新的發音符號的困擾下，便選擇放棄了學習 K K 音標，轉而用自己熟悉的中文ㄅㄆㄇ來拼英文音。

除了嘗試應用 K K 音標外，王啟琳(民 90)則建議使用改良式自然發音法(whole-to-parts)，以解決因自然發音法中字母與發音對應繁瑣，而導致小學生學習興趣低落的問題。但是，因為教師需要較多的訓練，備課時間耗時，且因小學授課時間不足等問題，教師對改良式拼讀法的推廣，有所保留。

面對 K K 音標以及改良式自然發音法，並不是完全適合國小學生學習英文發音的輔助工具之情況下，王旭所倡導的字典音標，則具有易學易記，且有讓學習者容易辨識讀音而不易造成混淆的優點(Wang, 1991)，似乎可以用來解決小學生以注音符號注英文發音的問題。

因此，本研究者認為，可採用一般英文字典中常見的字典音標(Dictionary phonetic symbols)，來協助小學生有效率地學習英文發音，以下進一步介紹字典音標，以及敘述其足以解決問題的原因。

什麼是字典音標？

所謂音標乃是根據語音所擬的一套發音符號(劉璧君，民 77)。而字典音標即是在英語系國家學生所使用的字典上所看到，讓讀者知道如何發音的一套音標符號。一般而言，分為英式字典音標與美式字典音標兩大類。英式字典音標大多數的是以國際音標，IPA (International Phonetic Alphabet)為主。而國際音標則是以本質上一符一音為原則，並以歐洲語言習慣來創立記憶音的符號。不同於的是，美式字典音標大部分來自於韋氏新世界字典(Webster's New World)。不像國際音標用另一組新符號，此一字典的母音是用英文字母代表母音(Wang, 1991；傅一勤，1995)。

字典音標建構原則，基本上是來自 26 個英文字母。因為英文的發音是多於 26 個字母，大約在 40 到 45 之間，視其英式或美式的發音不同而定(Šuštaršič, 2005)。因此每一種類的字典音標，視其發音的需求，由字典的編輯者，以一符一音(one symbol signifies one sound)，以及直接標記(phonetics without respelling)的原則，創造出所需的音標符號(Wang, 1991)。主要的目的是讓學習者對於一個音只需要學一個符號的音標，且儘量讓音標發音與字母的發音相像，避免造成混淆。一般而言，美式的字典音標傾向於讓讀者能在發音時，儘量就能識別出正確的拼字(orthography)。美式

的字典音標編輯者，除了使用基本的 26 個字母，作為基礎的發音符號外，通常會使用一些二合字母 (digraphs) 和區別發音符號 (diacritics) 來補足其發音上的需求，也就是說，學生只需要根據英文的 26 個字母的發音，再加上學習一些必需性的附加符號 (二合字母及區別發音符號)，即可用來記住英文的發音。美式英文字典中台灣學生所熟悉的韋氏字典 (Webster Dictionary) 或是美國傳統字典 (American Heritage Dictionary) 即是按照這種方式編輯而成。

為什麼選擇字典音標？

與 KK 音標比較，由於字典音標能較直接地表示語音與文字的關係，有不易造成混淆及協助記憶拼字的好處，因此是較合適做為教學的音標 (王旭，民 77; Wang, 1991)。KK 音標是沿襲國際音標而定，是以歐洲語言習慣來創立記憶音的符號。而字典音標的音標符號之創造原則是則是基於英文拼字的習慣，也就是發音時，以某一語音的英文字母或再加上一些簡易的附加符號，例如：see，注為 /sē/ (王旭，民 77，頁 65)，所以，學習者在學習字典音標時，由於音標與文字符號類似，容易見音標發音。與 KK 音標的一種符號，兩種念法的學習方式比較 (例如字母 e，寫作音標符號是 [e]，之後念作 ㄝ)，是相對的減輕了學習者的許多負擔。在小學生的邏輯能力不及國中學生的情況下，若需同時應付 KK 音標中兩種陌生的語言，實在太難了 (趙碧芬，民 95)。又因字典音標有音標符號與文字符號類似，學會之後會讓學生獲得會發音即會拼字的好處 (Wang, 1991, p. 20)。例如 “late” 一字注為 /lāt/，所需要注意的是加上不代表語音的 “e”，故極易記憶，這樣的優點對於需要大量記憶力來輔助其學習的台灣學子實有其必要性 (王旭，民 77，頁 67)。

在考慮字典音標的教學內容時，王旭於 95 年 7 月 24 日回覆本研究者的電子郵件中，認為美國傳統字典的音標具有明確的附加符號來代表長、短母音，可提醒學生這些是音標符號而非字母的特性，且因為經由這些附加符號，可讓學生更清楚的知道同一個音標，可以有長短音的區分。這種特性，對學生學習英文發音助益甚大，因此本研究字典音標的教學內容，以美國傳統字典的字典音標為主。

基於字典音標有容易辨識，不會造成混淆，且會協助拼字的好處，因此本研究者建議，可以認真考慮運用字典音標作為國小英文之教學音標，以輔助小學生學習英文發音及拼字。國小英語教學，正式實施已有將近五年的時間了，小學生使用注音符號來注記英文的問題，也日趨普遍。在面對學生發音錯誤的問題，李櫻 (民 89) 認為一旦在學習初期階段養成了錯誤的發音習慣，在國中階段再來補救，「也很難改正幼時已習得的模式」 (頁 9)。因此，以字典音標作為小學生啟蒙發音的教學音標的建議，似乎可成為解決此一難題的方案。

研究方法

本研究是以質化研究為主，量化研究為輔的行動研究實驗教學，藉著老師的教學時所執行之行動研究與學生的學習成就，以瞭解字典音標作為小學生開口說英文的啟蒙音標的可行性。本研究量化研究部份，透過兩年之實驗教學的前後成果差異，來瞭解字典音標、K. K. 音標與自然發音等三種英語發教學法的教學成效，進而探討字典音標做為小學生啟蒙音標的可行性。本研究在第一年中，分兩階段，分別於上、下學期實施。質化研究部份，採用行動研究 (action research)，其行式選擇以協同行動研究 (collaborative action research) 為主，主要是藉著群體的力量達到有效解決問題的目的 (Burns, 1994, 1999)。

參與實驗教學的六名授課老師，於 96 年 8 月 15 日完成招募，是由某技術學院四年級英文系，在校表現較為優秀，且未來有志於從事兒童美語教學工作之同學中挑選出

來的。參與實驗的對象學校為高雄市中正國小五年十三班、五年音樂班和高雄縣溪寮國小五年級忠班、孝班等共 4 班。

量化實驗教學分上下學期二階段，從 96 年 10 月 5 日到 97 年 5 月 23 日止，96 學年度上下學期各約 3 個月。學生學習前、中、後，各接受一次測驗。英文能力測驗內容參考許雪芳（民 92）在「培養兒童的文字認讀與拼寫能力：三種發音教學的成效研究」中，針對五年級學生的英語發音教學，所研製的英文能力前測改編而成。各次測驗，均有五項試題組，前測與中測分別為「音素切割」、「字詞辨認」、「字母拼寫」、「一般字認讀」與「偽字認讀」。在後測時，前四項維持不變，第五項試題組則略有修正，四個自然發音組繼續測試「偽字認讀」，而二個字典音標組則改成測試字典音標的「音標認讀」，二個 KK 音標組則改成測試 KK 音標的「音標認讀」。

而行動研究實施的步驟，則在 96 學年度上下學期各實施三循環（個月）。任教老師在每一個學期執行三個循環的行動研究，每一個循環為期一個月。行動研究實施時間同實驗教學時間也是 96 年 10 月 5 日到 97 年 5 月 23 日止。實驗教學上課時間後，接著是行動研究討論會。研究步驟採用 van Lier（1994）的週期行動研究的模式（Cycle of action research）。van Lier 改良 Kemmis 和 McTaggart（1982）的模式，用計畫、行動、觀察、反思等四項為一個週期的行動研究步驟，有系統地蒐集資料並分析在投入行動後的結果，以達到回答研究問題之目的。本研究之任教老師執行行動研究的步驟，根據 van Lier 的模式，採取四個步驟一個循環的方式進行，每週依序執行一個步驟後，將心得帶到討論會來討論。第一週主題為「計畫」（plan），第二週為「行動」（act），第三週為「觀察」（observe），而第四週為「反思」（reflect），反思其預定目標是否達成，並設立下一行動研究的目標，以此類推。討論主題是以為四週一循環為主。

結果討論

實驗教學探討

本研究前測於學生參加實驗教學開始前一週實施，中、後測則是在各階段教學後舉行。為確保比較字典音標、KK 音標與自然發音等三種英文發音學生學習成就資料的一致性，本研究所有資料分析，以在中測與後測等二次測驗均參加的 88 位同學成績資料，做為比較分析的基礎。

全部資料經由 t-檢定的顯著度測試後，八組學生在中測與後測「音素切割」、「字詞辨認」、「字母拼寫」、「一般字認讀」與「音標認讀」等五個試題組平均答對率差異情形進步或退步的顯著程度並不一致，分述如下：

首先為音素切割試題組的成績比較情形。音素切割試題組的設計，是要了解學生分辨一個英文字音素個數之能力。該項題組在第二階段開始與結束時的中測與後測裡，均各有 10 題題目，每題各有一個 3~8 個音素英文字，其中 1 個音節的字有 6 題，2 個音節的字有 2 題，3 個音節的字有 2 題。每題 1 分，共 10 分。表 1 彙整二校四班八組共 88 位學生，在中測與後測「音素切割」試題組平均答對率的比較情形。

表 3 第二階段分組教學前後各組同學在「音素切割」試題組平均答對率比較表

試題組別		中測百分比	後測百分比	平均數差異	t	顯著性(雙尾)	
中正國小	音樂班	字典音標組 14 人	57.9	49.3	-8.6	-2.05	0.061
		自然發音組 13 人(字典對照)	51.5	30.8	-20.8	-3.44	0.005
	13 班	KK 音標組 16 人	33.1	32.5	-0.6	-0.12	0.906
		自然發音組 15 人(KK 對照)	37.3	30	-7.3	-1.28	0.221
溪寮國小	五忠	字典音標組 8 人	38.8	38.8	0.0	0.00	1.000
		自然發音組 8 人(字典對照)	33.8	40	6.3	1.00	0.351
	五孝	KK 音標組 7 人	33.3	33.3	0.0	0.00	1.000
		自然發音組 7 人(KK 對照)	28.6	40	11.4	1.26	0.256

* $p < 0.05$

以 t 檢定，檢視八個教學組中測與後測平均答對率的差異值，發現除了有一組自然發音組(中正音樂班)同學在「音素切割」試題組的平均答對率，有顯著的退步情形外，其餘各組平均答對率，雖有進步與退步的差異現象，但統計上都不具顯著意義。自然發音組(中正音樂班)退步 20.8%，其 p 值為 0.005，小於 0.05，具有顯著意義。而其餘七組的平均答對率則都沒有顯著差異：字典音標組(中正音樂班) p 值為 0.061、KK 音標組(中正 13 班)為 0.906、自然發音組(中正 13 班)為 0.221、字典音標組(溪寮五忠)為 1.000、自然發音組(溪寮五忠)為 0.351、KK 音標組(溪寮五孝)為 1.000，而自然發音組(溪寮五孝)為 0.256，由於七組的 t 檢定 p 值均大於 0.05，顯示此七組學生在中測與後測「音素切割」試題組平均答對率的差異，在統計上全部不具顯著意義。

其次為字詞辨認試題組的成績比較情形。字詞辨認試題組的設計，是要了解學生分辨兩個相似字的能力。該項題組在第二階段開始與結束時的中測與後測裡，均各有 10 題，每題 1 分，共 10 分。表 2 彙整八組 88 位學生，在中測與後測「字詞辨認」試題組平均答對率的比較情形。

表 4 第二階段分組教學前後各組同學在「字詞辨認」試題組平均答對率比較表

試題組別		中測百分比	後測百分比	平均數差異	t	顯著性(雙尾)	
中正國小	音樂班	字典音標組 14 人	94.3	80	-14.3	-5.70	0.000
		自然發音組 13 人(字典對照)	96.9	89.2	-7.7	-3.83	0.002
	13 班	KK 音標組 16 人	93.1	83.8	-9.4	-2.80	0.014
		自然發音組 15 人(KK 對照)	94	83.3	-10.7	-2.62	0.020
溪寮國小	五忠	字典音標組 8 人	81.3	80	-1.3	-0.11	0.915
		自然發音組 8 人(字典對照)	92.5	82.5	-10.0	-5.29	0.001
	五孝	KK 音標組 7 人	96.7	91.7	-5.0	-1.17	0.296
		自然發音組 7 人(KK 對照)	95.7	87.1	-8.6	-2.12	0.078

* $p < 0.05$

以 t 檢定，檢視八個教學組中測與後測平均答對率的差異值，發現在「字詞辨認」試題組的平均答對率呈現顯著退步情形有五組，分別為一組字典音標組(中正音樂班)、一組 KK 音標組(中正 13 班)與三組自然發音組(中正音樂班、中正 13 班與溪寮五忠)。而其餘字典音標組(溪寮五忠)、KK 音標組(溪寮五孝)與自然發音組(溪寮五孝)等三組的平均答對率，雖有退步的差異現象，但在統計上都不具顯著意義。字典音標組

(中正音樂班)退步 14.3%，其 p 值為 0.000，自然發音組(中正 13 班)退步 10.7%、p 值為 0.020，自然發音組(溪寮五忠)退步 10.0%、p 值為 0.020，KK 音標組(中正 13 班)退步 9.4%、p 值為 0.014，而自然發音組(中正音樂班)退步 7.7%、p 值為 0.002，由於五組平均答對率差異的 t 檢定 p 值均小於 0.05，所以在統計上具有顯著意義。剩下的三組平均答對率則都沒有顯著差異，字典音標組(溪寮五忠) p 值為 0.915、KK 音標組(溪寮五孝)為 0.296，而自然發音組(溪寮五孝)為 0.078，由於這三組的 t 檢定 p 值均大於 0.05，渠等在中測與後測「字詞辯認」試題組平均答對率的差異，在統計上全部不具顯著意義。

第三組的比較為字母拼寫試題組的成績。字母拼寫試題組的設計，在於瞭解當在每一個英文字中，減少一個字母，然後要求受試者在聽到老師讀音之後，能夠填入遺漏字母，使之成為完整單字的能力。該項題組在第二階段開始與結束時的中測與後測裡，均各出 10 題，每題 1 分，共 10 分。表 3 彙整八組 88 位學生，在中測與後測「字母拼寫」試題組平均答對率的比較情形。

表 5 第二階段分組教學前後各組同學在「字母拼寫」試題組平均答對率比較表

試題組別		中測百分比	後測百分比	平均數差異	t	顯著性(雙尾)	
中正國小	音樂班	字典音標組 14 人	77.1	84.3	7.2	1.28	0.224
		自然發音組 13 人(字典對照)	67.7	73.8	6.2	0.76	0.459
	13 班	KK 音標組 16 人	50	43.1	-6.9	-1.79	0.094
		自然發音組 15 人(KK 對照)	49.3	41.3	-8.0	-2.26	0.041
溪寮國小	五忠	字典音標組 8 人	56.3	50	-6.3	-1.26	0.250
		自然發音組 8 人(字典對照)	58.8	47.5	-11.3	-0.89	0.402
	五孝	KK 音標組 7 人	33.3	55	21.7	2.89	0.034
		自然發音組 7 人(KK 對照)	41.4	51.4	10.0	1.53	0.177

* $p < 0.05$

以 t 檢定，檢視八個教學組中測與後測平均答對率的差異值，發現在「字母拼寫」試題組有顯著進步的，只有一組 KK 音標組(溪寮五孝)，而呈現出顯著退步的，則也有一組自然發音組(中正 13 班)，其餘各組平均答對率，雖各有進步與退步的差異現象，但在統計上都不具顯著意義。差異具顯著意義的 KK 音標組(溪寮五孝)進步 21.7%、p 值為 0.034，而另一組具顯著意義的自然發音組(中正 13 班)則退步 8.0%、p 值為 0.041，由於兩者 p 值均小於 0.05，因此各自進步或退步的差異，在統計上具有顯著差異。其餘六組的平均答對率則都沒有顯著差異，各組的 t 檢定 p 值，分別為字典音標組(中正音樂班)的 0.224、自然發音組(中正音樂班)的 0.459、KK 音標組(中正 13 班)為 0.094、字典音標組(溪寮五忠)的 0.250、自然發音組(溪寮五忠)的 0.402 與自然發音組(溪寮五孝)的 0.177，由於這六組的 t 檢定 p 值均大於 0.05，此六組學生在中測與後測「字母拼寫」試題組平均答對率的差異，在統計上全部不具顯著意義。

接續則比較一般字認讀試題組的成績。一般字認讀試題組的設計，是用來測量學生認讀一般字的能力。每次測驗都有 10 個字，以音素來計分，前測共有 43 個音素，配分 43 分，中測及後測皆為 35 音素，各配分 35 分。表 4 彙整八組 88 位學生，在中測與後測「一般字認讀」試題組平均答對率的比較情形。

表 6 第二階段分組教學前後各組同學在「一般字認讀」試題組平均答對率比較表

試題組別			中測 百分比	後測 百分比	平均數 差異	t	顯著性 (雙尾)
中正國 小	音樂 班	字典音標組 14 人	81.2	74.5	-6.7	-0.992	0.339
		自然發音組 13 人 (字典對照)	79.8	80.4	0.6	0.26	0.799
	13 班	KK 音標組 16 人	53.0	65.7	12.7	3.087	0.008
		自然發音組 15 人 (KK 對照)	53.0	54.1	1.1	0.23	0.822
溪 寮 國 小	五 忠	字典音標組 8 人	53.6	45.4	-8.2	-2.37	0.050
		自然發音組 8 人 (字典對照)	46.1	59.6	13.5	1.59	0.161
	五 孝	KK 音標組 7 人	52.9	76.7	23.8	2.56	0.051
		自然發音組 7 人 (KK 對照)	47.8	58	10.2	2.15	0.075

* $p < 0.05$

以 t 檢定，檢視八個教學組中測與後測平均答對率的差異值，發現在「一般字認讀」試題組的平均答對率上，呈現顯著進步情形的，有一組 KK 音標組（中正 13 班）。其餘七組平均答對率，雖有進步或退步的差異現象，但在統計上都不具顯著意義。進步幅度具顯著意義的 KK 音標組（中正 13 班）進步 12.7% 的、p 值為 0.008，由於 p 值小於 0.05，所以該組在中測與後測平均答對率的進步差異，確實具有顯著的意義。其餘七組的平均答對率則都沒有顯著差異，各組的 t 檢定 p 值，分別是字典音標組（中正音樂班）的 0.339、自然發音組（中正音樂班）的 0.799、自然發音組（中正 13 班）的 0.822、字典音標組（溪寮五忠）的 0.050、自然發音組（溪寮五忠）的 0.161、KK 音標組（溪寮五孝）的 0.051，與自然發音組（溪寮五孝）的 0.075，由於此七組的 t 檢定 p 值皆不小於 0.05，這七組在中測與後測「一般字認讀」試題組平均答對率的差異，在統計上全部不具顯著意義。

最後，則是比較音標認讀試題組的成績。此部分是直接比較這四組學生在中測「音素切割」、「字詞辨認」、「字母拼寫」、「一般字認讀」與「偽字認讀」等五項音標試題組平均答對率，與後測「音標認讀」平均答對率的差異情形，進而從兩項平均答對率的差異顯著程度，來推論學生學習字典音標與 KK 音標的成效好壞。表 5 為字典音標組與 KK 音標組兩校四組學生分組教學後的學習成效比較表。

表 7. 字典音標組與 KK 音標組「音標認讀」學習成效表

	中測平均 (五項音標試題組)	後測平均 (「音標認讀」)	進步或退步 百分比	t	p 值
中正音樂字典音標組	77.2	74.3	-2.9	-0.79	0.443
溪寮五忠字典音標組	56.5	48.7	-7.7	-1.02	0.341
中正 13 班 KK 音標組	58.4	64.6	6.2	2.04	0.060
溪寮五孝 KK 音標組	57.2	87.2	30.0	8.65	0.000

* $p < 0.05$

以 t 檢定，檢視此四個教學組中測與後測平均答對率的差異值，發現只有溪寮 KK 音標組有顯著的成長差異外，其餘三組都沒有顯著差異。溪寮五孝 KK 音標組的 t 檢定 p 值為 0.000，小於 0.05，顯示該組學生學習 KK 音標的進步成效，在統計上有顯著的意義。其餘三組的差異則不明顯，中正音樂字典音標組的 p 值為 0.443，溪寮五忠字典音標組的 p 值為 0.341，而中正 13 班 KK 音標組的 p 值為 0.060，三組的 p 值均小於 0.05，不具備統計上的顯著差異 ($p = 0.341 > 0.05$)。

三種發音系統學習成效探討

從上述分析的結果顯示，八組學生在中測與後測五項試題組平均答對率的差異情形雖然有各自不同的進步或退步情形，但是在只有呈現出統計上的顯著意義時，才能被推論具有顯著不同的進步或退步狀況。表 6 彙整各組在中測與後測「音素切割」、「字詞辯認」、「字母拼寫」、「一般字認讀」與「音標認讀」等五個試題組平均答對率的顯著差異情形。

表 6 字典音標、KK 音標與自然發音的學習成效比較表

音標組別	所屬班級	「音素切割」	「字詞辯認」	「字母拼寫」	「一般字認讀」	「音標認讀」
字典音標組	(中正音樂班)	△	X	△	△	△
	(溪寮五忠)	△	△	△	△	△
KK 音標組	(中正 13 班)	△	X	△	○	△
	(溪寮五孝)	△	△	○	△	○
自然發音組	(中正音樂班)	X	X	△	△	
	(中正 13 班)	△	X	X	△	
	(溪寮五忠)	△	X	△	△	
	(溪寮五孝)	△	△	△	△	

註：○顯著進步；△進步或退步差異不具顯著意義；X顯著退步

二個 KK 音標組在五個試題組所呈現的學習成效，顯著進步的項目有三項，而顯著退步的有一項。呈現顯著進步的三項，分別是「字母拼寫」(溪寮五孝 KK 音標組)，「一般字認讀」(中正 13 班 KK 音標組)，以及「音標認讀」(溪寮五孝 KK 音標組)。呈現顯著退步的則有「字詞辯認」一項(中正 13 班 KK 音標組)。剩下的「音素切割」之學習成效，則不具顯著意義。

二個字典音標組在五個試題組所呈現的學習成效，顯著退步者有一項，而其餘四項的學習成效則不具顯著意義。學習成效呈現顯著退步的為「字詞辯認」(中正音樂班字典音標組)，而其餘「音素切割」、「字母拼寫」、「一般字認讀」與「音標認讀」等四項之平均答對率的進步或退步差異，都不具顯著的意義。

最後，四個自然發音組在五個試題組所呈現的學習成效，則是顯著退步有三項，而無顯著意義的則有一項。呈現顯著退步的三項，分別為「音素切割」(中正音樂班自然發音組)、「字詞辯認」(中正音樂班自然發音組、溪寮五忠自然發音組、溪寮五孝自然發音組)，以及「字母拼寫」(中正 13 班自然發音組)。而不具顯著意義的一項則為「一般字認讀」。

綜合前五項英文能力試題組的比較分析結果，顯示各個 KK 音標組的學習成效似乎比自然發音組與字典音標組的成效要好，特別在「一般字認讀」、「字母拼寫」，以及「音標認讀」等三項。其中中正 13 班 KK 音標組在「一般字認讀」，進步 12.7%，從中測的 53.0% 成為後測的 65.7%。相同地，溪寮五孝 KK 音標組在「字母拼寫」，進步 21.7%，從中測的 33.3% 成為在後測的 55%。在「音標認讀」，更是大幅地進步 30.0%，從中測的 57.2，成為後測的 87.2%。造成這種情形的可能原因有許多，如該組的任教老師教得特別好，或是該組的學生特別積極，或是 KK 音標較容易被學習的，或是大部分的學生之前有學過 KK 音標。

從授課老師執行的行動研究中得知，KK 音標的學習成效比較好的原因，似乎是因為學生有學過 KK 音標的經驗所導致。在第二學期第一循環的行動研究時，教中正音樂

字典音標組的 Amy 老師就發現到，她班上的 14 學生，大約有 57% 已經學過 KK 音標（8 人學過，完全沒學過者 6 人），而在學習字典音標符號時，這些學生常表示會與 KK 音標混淆的情形。

另外，學生在中測時填答基本資料有關是否學過 KK 音標時，也呈現過半學生先行學過現象。在「一般字認讀」有呈現進步的中正 13 班 KK 音標組 16 位同學中，學過 KK 音標的比例也有 44%（7 人學過，完全沒學過者 9 人）。同時，在「字母拼寫」，以及「音標認讀」有大幅進步的溪寮五孝 KK 音標組 7 人中，則也有約 57% 以上的同學表示曾學過 KK 音標（4 人學過，完全沒學過者 3 人），在兩組都有約 50% 學生具學習 KK 音標經驗的前提下，接受第二階段的 KK 教學，對他們來說，實有複習與加強的效果。

因此，由以上的兩項推論看來，KK 音標組的學習成效較佳，有可能是因為大部分學生之前的 KK 音標學習經驗所導致。繼續學 KK 音標的學生獲得了加強的機會，而接受字典教學的學生，則受到兩種音標系統的混淆。

而在字典音標學生學生面臨與 KK 音標的混淆下，在謹經過約三個月學習後，與長時間接觸自然發音發音的學生相較，其學習成就仍較自然發音組為佳。研究發現，二個字典音標組在「音素切割」、「字詞辨認」、「字母拼寫」、「一般字認讀」與「偽字/音標認讀」五個試題所呈現的學習成效，顯著退步的只有「字詞辨認」一項，而四組自然發音組在五個試題組所呈現的學習成效，則是顯著退步的除「字詞辨認」外，尚有「音素切割」、以及「字母拼寫」二項。特別是在「字詞辨認」此項中字典音標組中只有中正音樂班一組退步，而自然發音組則有中正音樂班、中正 13 班與溪寮五忠等三組是顯著退步的。此外，參加實驗的中正及溪寮國小學生中，分別是從小學三年級及一年級即開始學英文，使用的都是以自然發音為主的國內佳音及康軒版的英文教材。自然發音組學生雖有長時間學習自然發音經驗，但在測驗成績上卻未能表現出其成效。而字典音標的學生只經過短時間的學習，其學習成就卻較自然發音組佳，由此，更可證明字典音標是幫助學童較有效率學習英文發音與拼字工具之一。

行動研究探討

根據 96 上學期二位自然發音老師 (Amy 及 Nancy)，及 96 下學期四位老師 (教字典音標的 Amy、教 KK 音標的 Alice 及教自然發音的 Anna 和 Nancy) 所發現的 36 行動研究報告分析結果，課堂問題主要可分為 1) 音標與音標混淆，2) 發音規則困難，3) 一符多音，4) 相似音與 5) 其他等五大類。這些問題都是任教老師在行動研究討論會中提出的，在集合其他老師或助教的建議後，老師們都有嘗試實行不同的方案，以期能解決問題。經其在課堂中多次的實施解決方案的結果，大部分的問題都有得到改進。

第一類，則是音標與音標混淆。所謂音標乃是根據語音所擬的一套發音符號 (劉璧君，民 77)。本研究所指的音標乃為 KK 音標與字典音標。音標與音標混淆是指老師在教授 KK 音標或字典音標時，學生對其發音符號所代表的音/符號，與英文字母、其他音標系統的發音符號或附加符號 (二合字母及區別發音符號) 產生混淆的情形。屬於此類的行動研究共計 12 個。

第二類，則是發音規則困難。所謂發音規則是，當學生看到字母後，根據某些既定的發音規則來唸出對應的發音，進而結合音素唸出整個字的讀音。當老師在教完字母對應的發音之後，該如何教同學唸出整個字的讀音，會常常根據老師的方法不同而產生不同的結果。在 36 個行動研究中，有 9 個屬此類。其中 5 個有關發音規則困難的行動研究，是由教授自然發音的老師所提出的，其中 96 學年度上學期有 1 個，下學期有 4 個。

第三類是一符多音。所謂一符多音就是在看到字母，唸出對應的音會有不同。如字母 e 在 sell 和 seal 中的發音是不同的，在學習時，學生面臨一符多音的情況，難免

發生混淆。在 36 個行動研究中，有 8 個是屬於此類。

第四類，是屬於相似音的問題。所謂相似音，是指兩個單字的讀音非常類似，在學習過程中，常造成學生誤以為是同音同字。在 36 個行動研究中，有 4 個是屬於此類。

第五類，是「其他」部分。在此是指蒐集到的行動研究，目前暫時無法適當分類，而稱為其他。在 36 個行動研究中，有 3 個是屬於此類。

以下為任教字典音標的 Amy 老師，在中正國小 96 學年度下學期第一循環上課時，所執行過的行動研究報告分享。

學生初次認識字典音標符號，容易和 KK 音標符號搞混。Amy 老師在進行課本 (P. 4) 的聽音辨音習題時，發現有些學生把音標的寫法與 KK 音標混淆，例如：lamp(lämp) 卻寫成 lamp[læmp]。她便詢問學生是否有學過 KK 音標，有不少學生表示有，也表明容易與字典音標混淆。對於學生所產生的問題，討論會中有 3 位老師提供建議，分別是 1) 可以在字典音標書前放濃縮對照表符號，讓學生去對照，2) 作習題的時候，讓學生去翻課本先前所教過的地方，讓學生有音標參考，3) 可以參考字典音標課本 P. 71 的附件：認識字典音標，讓學生從單字反推該符號的發音。

針對這些建議，她覺得利用課本來做對照是方便又有效的，因為不管是做習題也好，複習也好甚至是玩遊戲，學生都可以利用課本來做對照。更進一步地，為了加深學生的印，她還做了一個簡介字典音標，內容如下的小講義：

什麼是字典音標？
字典音標基本上是由三部分組成，分別是：

1. 基本英文字母 (26 個)；例：a、b、c、d、e……z
2. 二合字母；例：(hw)、(kw)、(ng)…….
3. 區別發音符號；例：˘，-，^，· (這些符號都只標在母音上面，如：(ä)、(ë)…….)

第二週時，她把課本收回且讓助教在教室後訂正回家作業時，開始利用黑板舉例說明講義上的內容。她首先說明字典音標主要是由基本英文字母 (26 個)、二合字母和區別發音符號等三部分所組成，然後針對每一部份都舉例說明。為了讓學生更注意老師所出的題目，及深刻記住今天所教的，她抽問了學生並對答錯的部份，立即糾正。以下為其將講解過程中與學生互動錄音片段：

老師：這是什麼？對對對對對！！就是你…！（老師手指黑板上的“^”符號）

學生：這是短音

老師：很好！這是什麼？

學生：這個是那個什麼…區別發音符號 (diacritics)

老師：很好！好～我在來一個～這是什麼？不要看這個

學生：也是基本字母

老師：很好！再加上這個符號～是什麼？

學生：區別… (不確定的口氣)

老師：區別？區別？你看～講義上有嘛！

學生：區別發音符號

老師：對嘛！區別發音符號，不要這麼緊張嘛！再來一個～

老師：來～請問這是什麼？(同學幫答聲響起) 你看你同學對你多好～

學生：(全班偷偷笑～)……

課程進行完後，於是她對全班實施測驗。目的是想了解學生們是否已經真正有初步的概念，也知道區別發音符號只能出現在母音上面，題目類型為圈選正確的字典音標，例如 (ë) 與 (ε)。測驗結果是 14 位學生，每人 5 題，計有 70 題，答對 67 題，答對率 95%。如此的測驗結果，答對率雖然看似很高，她並不滿意，因為還是有三位學生錯在和 KK 音標混淆的地方。而且，在行動研究討論會時，助教提到，考試時不時有學生東

張西望，因此她決定利用黑板更清楚呈現字典音標的規則。

在第三週時，她為了要更清楚地說明什麼是字典音標規則，她在黑板上呈現了視覺的區分。不同於第二週的是，她在黑板上第一排寫正確的音標符號，第二排寫錯誤的，第三排式混合著錯和對的。她帶著學生唸正確的字典音標後，也特別強調了區別符號只能加在母音的符號上。在進行到第三排時，學生們也發現了錯誤的音標符號，因為子音上面是不該加上區別發音符號的。當時她與學生也一一地更正。在複習後，也實施了與上週類似的圈選測驗。這次的測驗的目的是想了解學生是否有深入的認知字典音標，且不會與KK音標搞混。測驗結果是14位學生，每人5題，計有40題，答對66題，答對率94%，成績結果比上次的測驗進步4%。如此的成績表現讓她認為是清楚地重復規則是有效的，同時也推論學生似乎是不再混淆了。所以，她決定下週會再次採用此方法，而且在未來的課程裡，也會多次複習。

第四週時，在進入複習後，在她確定幾乎全班已經會發音和知道區別發音符號後，接著抽報雙號數學生進行測驗，共抽點了七人，一人五題，共35題，答對率94%。對於與上週相同的答對率，她也並不是非常的滿意。因為助教在討論會時說：「改了很多次考卷，但是學生們每次都有錯，而且每次都是同樣的學生錯同一個地方」。她也覺的有一些困惑，為甚麼多次的復習結果，仍然無法超越那幾位學生的原有KK音標的印象。

雖然在這一次行動研究中，她覺得對自己的教學成效有些挫折，但是畢竟大部份的學生都會分辨了。面對學生字典音標符號和KK音標符號混淆的問題，她認為多多複習，加深學生的印象可能是較有效的方法，但是對於那些“改不過來”的學生，她也學到了，即使是一點點的混淆，也可能是學習上的致命傷，這種衝突(interference)也是需要長久的學習才能改過來的。

結論與建議

本研究的目的是在於透過行動研究與兩年的實驗教學，來比較字典音標、KK音標與自然發音等三種英語發音教學成效之異同，進而探討字典音標做為小學生開口說英文的啟蒙音標的可行性。在實驗教學的研究結果顯示，發現三種音標系統的教學成效，由高而低，依序為KK音標教學、字典音標然後是自然發音。總體上來說，KK音標組在「字母拼寫」、「一般字認讀」與「音標認讀」的學習成果較佳。自然發音組在「音素切割」、「字詞辨認」、「字母拼寫」等三項的學習成效顯著退步。字典音標組與自然發音在五項英文能力的表現居中，沒有任何一項呈現顯著進步。KK音標組表現較佳的情形，究其原因，可能是由於部分學生學過KK音標的經驗影響。字典音標在經過短時間學習，在三種發音系統中成效居中，因此建議相關教育決策單位可推廣此一音標系統，讓學童學習音語發音時，有多一種輔助工具選擇。

本研究結果發現，在選擇參與實驗教學學生時，學生的特色，例如是否有學過KK音標的經驗，是較難掌控的變數。目前在小學階段的學生，由於父母望子成龍的心切，在小學階段父母即將送去補習班學習KK音標的比率是相當高的，因此在本研究者實驗教學分組時，是很難刻意排除有KK音標學習經驗的學生，為增進實驗教學結果之信度與效度，本研究建議未來的研究可以在低年級實施。

參考文獻

- 王正寧(民92, 2月17日)。另類“自然發音法”小學生自行注音加強記憶，影響將來發音與聽力，教育部考慮採行音標。聯合晚報，四版。
- 王旭(民77)。從音韻學及語言學習的觀點看英語教學的音標。中華民國第五屆英語文教學研討會論文集，61-70。台北：文鶴。

- 王啟琳 (民90)。國小英語教學試用改良式自然發音法的探討。(行政院國家科學委員會專題研究計畫成果報告, (NSC-89-2411-H-029-021))。台中市：東海大學。
- 王啟琳 (民92)。從英語字音字形對應分析，論字母拼讀教學。東海大學文學院學報。44 (2), 280-308。
- 李櫻 (民89)。談兒童英語教學中的發音問題。英語教學, 25 (2), 4-15。
- 林淑玲 (民92, 2月24日)。小三新學三語文，教長怕出狀況。中國時報。上網日期：95年，7月4日。網址：<http://eng.tceb.edu.tw/>
- 施玉惠、周中天、陳淑嬌、朱美惠 (民87)。國小英語教學實施現況與未來規劃方向。第七屆中華民國英語文教學國際研討會論文集, 759-778。台北：文鶴。
- 許雪芳 (民92)。培養兒童的文字認讀與拼寫能力：三種發音教學的成效研究。國立台北師範學院兒童英語教育研究所碩士論文。
- 張玉玲 (民93)。全語言兒童英語教學。台北：敦煌書局。
- 張武昌 (民92, 2月21日)。都是「自然發音法」惹的禍？中國時報，論壇。
- 傅一勤 (民84)。也談國中英語的音標教學。國立編譯館通訊, 8 (1), (46-48)。
- 教育部 (民92)。國民中小學英語教學手冊。台北：教育部。
- 教育部 (民95)。國民中小學九年一貫課程綱要—語文學習領域 (英語)。台北：教育部。
- 教育部：自然發音法符合國小英語自然學習順序 (民92年, 2月19日)。中國時報。上網日期：民95年7月4日。網址：<http://web.zhes.tc.edu.tw/vicky/9203/news92030107.htm>
- 葉麗雪 (民94)。KK音標輔助英語教學之研究—以六年級學童為例。國立中山大學教育研究所碩士論文。
- 趙碧芬 (民95)。何時學習KK音標。上網日期：95年8月10日。網址：http://bn8720937.myweb.hinet.net/kid&montessori/kid/kid_12.htm
- 劉璧君 (民77)。音標與發音教學。英語教學, 12 (3), 3-7。
- 鄭任汶 (民92, 2月17日)。小學生用勺勺唸ABC。中時晚報，焦點新聞。
- 謝良足 (民89)。國小英語發音教學研究 (I)。(行政院國家科學委員會專題研究計畫成果報告, NSC-89-2411-H-251-001)。屏東市：屏東商業技術學院
- 謝良足 (民90)。國小英語發音教學研究 (II)。(行政院國家科學委員會專題研究計畫成果報告, NSC-89-2411-H-251-002))。屏東市：屏東商業技術學院。
- 謝欽舜 (民83)。發音教學：Phonics 教學的基本內容與方法。CET (2), 35-41。
- 韓青秀 (民92, 2月24日)。ABC竅門：先說2年英語再學音標。聯合晚報，四版。
- Burns, A. (1994). *Introduction to research methods*. Melbourne: Longman.
- Burns, A. (1999). *Collaborative action research for English language teachers*. Cambridge: Cambridge University Press.
- Butler, Y. G. (2004). What level of English proficiency do elementary school teachers need to attain to teach EFL? Case studies from Korea, Taiwan and Japan. *TESOL Quarterly*, 38(2), 245-278.
- Kemmis, S., & McTaggart, R. (1982). *The action research planner* (2nd ed.). Burwoed, Victoria: Deakin University Press.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis* (2nd ed.). Thousand Oaks, CA: SAGE
- Šuštaršič, R. (2005). Dictionary Transcriptions Representing Standard British and American Pronunciations and Their Application in Teaching English Phonetics, Phonetics Teaching and Learning Conference, July 30, University College, London, <http://>

- [//www.phon.ucl.ac.uk/home/johnm/ptlc2005/pdf/ptlcp41.pdf](http://www.phon.ucl.ac.uk/home/johnm/ptlc2005/pdf/ptlcp41.pdf).
van Lier, L.(1994). Action research. *Sintagma (Spain)*.
Wang, H. S.(1991). An Orthography-Based Symbolic system for TESL. *English Teaching Forum*, 24, 3, 16-21.

Picture-Book-Making in EFL Writing Classes: Third-Graders' and Teacher's Perception

Ya-ling Lin (林雅玲), Fang-chi Chang (張芳琪)

National Chiayi University

s0961131@mail.ncyu.edu.tw

engivy@mail.ncyu.edu.tw

In southern Taiwan, an ongoing writing project has been conducted in a primary school since two years ago. The project started with providing students with opportunities to draw and/or write freely based on what they had learned in English classes in Semester 1. In the following year, structured writing activities were planned and implemented. Incorporated into the fifth semester of the ongoing project was the picture-book-making project which started in September 2008 and of which the plan was to have students create two picture books in groups. This paper reports what the 122 third-grade students' and two teacher instructors' perception are about the picture-book-making project. The data collected revealed that the majority of the students held a positive attitude towards this project; they especially liked to see their classmates' works presented in class. More than eighty percent of the interviewed students of different language proficiency levels stated that this activity made them like English learning more. The teachers named several contributions of the picture-book-making group work, such as motivating learning and involving, facilitating learning of English, learning about peer cooperation and interaction, and so on.

INTRODUCTION

According to the Guidelines of Language Arts (English) in Grade 1-9 Curriculum from the Ministry of Education (MOE), the learners' development of listening and speaking skills is the major focus, whereas writing is not considered much. Some researchers (e.g. Davies & Pearse, 2000; Paul, 2003), however, assert that what primary school students' really need is an integrated-instruction. Davies and Pearse deem that four language skills have the same degree of importance and the most appropriate language learning program for second or foreign language learners is the one which integrates four language skills.

Paul (2003), likewise, claims that most Asia governments pay more attention to listening and speaking, which may result in a "serious misperception of the needs of these children and what they are capable of" (p. 83). Writing, according to some researchers (e.g., Chiu & Chang, 2006; Davison & Dowson, 2003; Paul, 2003; Sejnost & Thiese, 2007), has numerous advantages such as facilitating language learning and enhancing one's critical thinking ability. Vygotsky (1978) also postulates that in children's whole learning process, writing is a crucial element and it also acts as a tool for children to enhance their cognition. Therefore, it seems that an English curriculum which incorporates writing as one of the essentials is more beneficial for language learners.

In southern Taiwan, the importance of writing has attracted a primary school's attention. This school has been cooperating with two academics to implement an English writing project, which started with the students enrolling in the first grade in September 2006. In Semester 1, the students drew and/or wrote freely about what they had learned in English. In Semester 2, introduced was a brainstorming technique, which was employed to generate

ideas for writing. In Semester 3, lexical parts of speech were instructed in order to help students combine words appropriately to make sentences. In Semester 4, the practice of self- and peer-proofreading was implemented for the students to correct their and others' writing. These were the activities conducted in the two years, four semesters, previous to the current study. For the fall semester of 2008, the fifth semester of the ongoing project, picture-book-making was the major task. This report presents the picture-book-making project and the participants' perception of this picture-book-creation activity.

LITERATURE REVIEW

Reviewed in this section are children's writing development, picture books and language learning, the cooperative-learning principles, Vygotsky's (1978) social interaction and zone of proximal development.

Development of Children's Writing

Drawing and writing, in fact, go hand in hand with children at the early age and they are both regarded as graphic symbol systems for children (Jalongo, 2007). Children, at first, may not distinguish the differences between drawing and writing, but before long, they will weave drawing and writing together (Galda, Cullinan, & Strickland, 1997). Drawing first helps children make use of the communicative strategies they are comfortable with. Later, as soon as children become confident in drawing, they are more and more comfortable to use drawing to help their writing (Lamme, Fu, Johnson, & Savage, 2002).

Children will gradually generate more and more complex scribbles. Their scribbles may become the collections of lines and loops which resemble letters. Or children may imitate the words they read from the signs. Besides, they may ignore any punctuation or just mix all capital letters they know in their productions (Cunningham, Moore, Cunningham, & Moore, 2004). After a period of time, children start using invented spellings which may sound correct but violate the spelling rules. Temple, Nathan, Temple and Burriss (1993) propose that the more opportunities for children to experience their invented spelling, the more possibility that they will become better writers and readers in the future. In conclusion, children's writing development may start with drawing or scribbling, but with adequate opportunity to experiment written language, children will become mature writers. These development characteristics are documented on young EFL learners and ESL learners as well (Galda et al., 1997).

Picture Book and Language Learning

"Why not engage students with picture books and improve their writing skills simultaneously," asked Paquette (2007, p. 156). Since writing starts with drawing and since young children are fascinated by picture books, combining picture-book reading and picture-book making seems to be a good way for language learning. Researchers (Culham, 2004; Jalongo, 2007; Morrow, 1996) do regard using picture books, which contain closely related illustrations and texts, to help students become good writers as an effective strategy (Culham, 2004; Jalongo, 2007; Morrow, 1996). A plenty of activities can be employed by using picture books with students of all ages or various proficiency levels. For example, students may orally discuss about the texts and illustrations, interpret the illustrations with various angles, label pictures with different stories and even use the picture books as the inspiration of creating their own picture books (Peregoy & Boyle, 2005).

Morrow (1996) notes that creating picture books allow students to use their own voices, existent English knowledge, and imagination to practice what they have known about writing and to draw what they want their readers to appreciate. Besides, picture-book-making serves as the connecting bridge for listening, speaking and reading activities. For example, before the writing activity, the teacher and students may discuss the topic and the main ideas. After the picture book is completed, the product can serve as the reading materials for the

next lessons. Therefore, creating picture books not only helps learners learn the written language, but also enhances the other three skills.

The above notions suggest a revised version of Paquette's (2007) question, that is, "Why not engage students with picture-book reading and writing in group settings to foster language learning?" In the following, cooperative learning, and Vygotsky's (1978) social interaction and zone of proximal development are reviewed to support this advocacy.

Cooperative Learning, Social Interaction and Zone of Proximal Development

In the light of Vygotsky's (1978) learning theory, learning itself, is a social activity. Only when the learning condition is under the assistance of social interactions, will learning become effective for the learners (Ghosn, 1996). In terms of Vygotsky's (1978) theory of zone of proximal development, there is some distance between children's developmental and proximal levels. Through social interaction with peers, learners will advance to the next stage. Paul (2003) believes that "group activities encourage cooperative learning, and lead to social interaction to a much greater extent than pair activities" (p. 42). Johnson, Johnson, & Holubec (1994) also propose that when we put students in the cooperative learning groups, they can gain greater comprehension of the content and skills they are studying. In cooperative learning groups, peer interaction can provide sufficient input and many opportunities for output in the warm supportive learning context.

Language acquisition thus occurs when learners experience a complex interaction of a number of input, output and context, and cooperative learning has a positive impact to it (Kagan, 1995). Temple et al. (1993), for example, note that through observing and interacting with more skilled others, children learn about writing more efficiently. In the present study, the writing activities involved several group activities which encourage individuals to seek outcomes that are not only beneficial to themselves but also to all their group members. Hence, the concept of cooperative learning and social interaction is employed in the current research. It is hoped in such cooperative learning groups, learning result will be the most beneficial.

To sum up, researchers have claimed that writing is beneficial to the development of other language skills and learning can best occur in the cooperative learning groups. Moreover, picture-book activities may serve as good ways to facilitate language learning. Inspired by the advocacy and being eager to make students' learning the most profitable, the researchers conducted picture-book-making activities in four EFL third-grade classrooms to investigate some issues and this paper reports the findings on two of the issues:

RQ 1: What is the students' perception of the picture-book-making activities?

RQ 2: What is the instructors' perception of the picture-book-making activities?

METHOD

Setting

The School, which is located in southern Taiwan, was established in 2003, and defined as a bilingual school by the government. That is to say, the School has more freedom and resources in designing their language curriculum than other primary schools. Students of the School, therefore, have three periods of English lessons per week instead of two periods that most primary schools do and furthermore, the School lets their first graders start to learn English.

In the summer of 2006, the School began to cooperate with two academics to design an English curriculum for the new coming first graders. Perceiving the value and the benefits of starting writing early, the two academics incorporated writing into the early EFL curriculum. Hence, some writing activities like free drawing and/or writing, brainstorming, and proofreading were conducted sequentially on the EFL young learners in the past two years.

Participants

The student participants were 122 third graders, aged eight or nine, who belonged to four different classes taught by two Chinese English instructors respectively. These student participants have been engaged in the writing activities introduced above since they were in the first grade. Before entering the elementary school, most of them took about one year or two years of some English instruction in kindergarten. The other participants were two teacher participants, Ms. Hsu and Ms. Chang, two academics, and a graduate student, the first author of the report.

Weekly Activities

A complete class period each week was scheduled for this picture-book-making task. The semester schedule allowed each group to create two picture books. The goal of the English curriculum on the writing portion was thus to have students work in heterogeneous groups to produce two picture books.

The participants of each class worked in six heterogeneous groups to make their groups' picture books. Based on the notions of three English writing approaches, *Language Experience Approach*, *Interactive Writing Approach* and *Process Writing Approach*, the researchers designed some picture-book-making activities. It took seven weeks as a cycle to complete one picture book with the following activities sequenced: (1) introducing the picture book, (2) brainstorming and drafting, (3) self- and peer-proofreading, (4) creating front and back covers, (5) shared-reading and conferencing, (6) making the picture book and practicing oral presentation, and (7) presenting publicly. The details of classroom activities for the first seven weeks on the first picture book were briefly explained as follows.

(1) Week one: introducing the picture book

Zoo-Looking (Fox, 1996) was introduced and served as the theme for the student participants' first picture books.

(2) Week two: brainstorming and self-drafting

The student participants of each class were assigned into six heterogeneous groups. The instructors first led the whole class to generate ideas by clusters on the blackboard and then each group generated their ideas on the A4-sized paper. Afterwards, based on the clusters each student drafted sentences alone. The student participants were asked to write at least five sentences.

(3) Week three: self- and peer-proofreading

In week 2, the student participants wrote at least five sentences individually and this week, the student participants of the same group exchanged drafts for self- and peer-proofreading. This activity was to help writers themselves and their peers to find possible mistakes on the writings.

(4) Week four: creating front and back covers

The instructors led the whole class to generate ideas for some possible beginning and ending sentences first. Second, group members discussed and wrote their book titles, beginning and ending sentences for their group books.

(5) Week five: shared-reading and conferencing

The teacher presented all students' written products (the drafts after self- and peer-proofreading) on a big screen for the whole-class shared-reading. She then guided students to discuss about the good and weak parts of the products in the hope of helping students discover some drawbacks or merits from their own or their peers' writing. Subsequently, the students' written products were returned to them for a conference on choosing the sentences they wanted to be published. The chosen sentences were proofread and corrected by the researcher-observer to ensure the accuracy of the publication.

(6) Week six: making the picture book and practicing oral presentation

After receiving the written products proofread by the researcher observer, the student participants copied their sentences onto the pages, which were bounded to make a book. Afterwards, the instructor demonstrated how to do the oral presentation and let students practice their own presentation in groups.

(7) Week seven: presenting publicly

The groups took turns to present their group's production. The student participants read aloud the pages they were responsible for. In addition, the writers might ask the listeners some questions on the book.

For the following weeks, the aforementioned writing activities were repeated until the end of the semester, with *Market Day* (Ehlert, 2000) as the second picture book to be presented as the stimulus for the production of a second picture book.

Data Collection and Analysis

Data were collected during a period of about four months from September 2008 to January 2009 through two methods: (a) questionnaires in Mandarin (Appendix A and C) read aloud to the student participants for them to fill out on a separate sheet for responses and (b) interviews (Appendix B and D) with some focal student participants and two teacher participants (Appendix E). The interviews with the focal student participants were conducted for further information after the student participants filled out Book One's and Book Two's questionnaires respectively for further information. The principle of selecting student interviewees was first, choosing students from the same group and second, choosing some special students who circled 'like' or 'dislike' for all items. Furthermore, the second interview included the Book 1 interviewees because the researchers wanted to investigate whether the Book 1 interviewees changed their conception of picture-book-making activities or not. Hence, 29 student participants were interviewed for Book 1 and 48 interviewees for Book Two. Finally, interviews with two teacher participants were conducted respectively as well in the end of the semester. All the interviews were conducted in Chinese and they were recorded for later analysis.

FINDINGS AND DISCUSSION

The first research question is on the students' perception of the writing activities. It is to be answered with the questionnaire responses given by the students after they completed Book 1 and Book 2 respectively. Meanwhile, the interview data from some focal students serves as the supplementary findings for researcher question 1. As for research question 2, the data was retrieved from the interviews of two teacher participants. Generally speaking, whether student or teacher participants, they all had a very positive response to the picture-book-making activities. The details of the findings and discussion are elaborated in the following.

Student Participants' Perceptions of Picture-Book-Making Activities

To give answers to research question 1, the student participants responded to two questionnaires (for Book 1 and Book 2 respectively) and overall, most student participants held positive attitudes towards the writing activities. Table 1 represents some general findings of the two questionnaires. Generally speaking, these students liked these picture-book-making activities; 86 percent of them liked the activities on Book 1 and 89 percent, Book 2. Some further findings related to their perception are explained in the following

Table 1 *Student Participants' Perception of Picture-Book-Making Activities*

	Book 1	Book 2
1. I like these picture-book-making activities.	86%	89%
2. I like to use clusters to generate ideas with my group members.	81%	66%
3. I like to write sentences based on the clusters I made.	66%	66%
4. I like to proofread my sentences.	89%	89%
5. I like to proofread my classmate's sentences.	84%	75%
6. I like to discuss with my group members about our book title.	89%	79%
7. I like to discuss with my group members about our book's beginning and ending sentences.	80%	71%
8. I like to see my works presented by my teacher.	80%	63%
9. I like to see my classmates' works presented by my teacher.	90%	93%
10. I like to cooperate with my group members to choose the sentences we want to be published.	84%	61%
11. I like to draw pictures for the pages I am responsible for.	85%	78%
12. I like to read out the sentences I wrote in front of the class.	78%	68%

Task Familiarity Reducing the Resistance

For all the writing activities, students reported with preference of more than 60%. They particularly liked to see their classmates' works presented by the teacher (90% on Book 1 and 93% on Book 2) and proofread their own sentences (89% on Book 1 and 89% on Book 2). One of the reasons for their enjoying these activities may be the task familiarity. For the past two years, the student participants had experienced free writing and/or drawing, brainstorming and peer- and self-proofreading activities with the same instructors and accordingly, they had been pretty familiar with the most major writing activities of the picture-book-making task. Such familiarity may eliminate some certain degree of students' fear of trying a new task. Meanwhile, according to the interview data, 18 out of 29 interviewees stated that this task was easy on Book 1, whereas 37 out of 49 on Book 2. This phenomenon indicated that more and more students felt such task was not difficult after completing two books. Hence, being familiar with the task may help students reduce some uncertainty of trying a new task.

Moreover, the questionnaire data showed that the top one reason for students to like this task is *it is interesting*. This data revealed that students regarded the picture-book-making task as a joyful activity and only in this kind of learning environment, learning can obtain its greatest benefit (Li, 2004; Paul, 2003). The second reason is *I like drawing*. In students' products, they had to decorate their own pages in which they wrote sentences. Quite a few interviewees also indicated that they enjoyed drawing the most. Drawing might have added some more joyful elements for the picture-book-making task.

Group Work Increasing the Possibility

Not only did students report that the picture-book-making task was interesting and they liked drawing, but also stated that they enjoyed working with their classmates. The original design of the picture-book-making task was to let the student participants create their own books individually. However, the teacher participants worried that this might be a heavy burden for the student participants. The design, therefore, was shifted from individual work to group work. According to the interview data, all interviewees pointed out that while they worked with their group members, everyone was willing to help. Whenever they encountered any problems such as spelling, their group members would teach or help them to spell out the words.

They taught each other how to write sentences or vocabulary, or they provided various ideas for group discussion. And most importantly, they felt happy to do so. Furthermore, one

of the expectations of the picture-book-making task was to let students understand the importance of cooperative learning. It seems that the students had obtained good impression of cooperative learning. The two teacher participants, according to their observation, also noticed that students were more and more willing to assist each other, especially on Book 2. Researchers have claimed that learning is more effective when it is under the assistance of social interaction (Ghosn, 1996; Vygotsky, 1978) and in the present study, all participants, including students and teachers, found the delight of working in groups.

Group Rearrangement Increasing Disagreement

The teacher participants asserted that students should not only learn how to cooperate with others, but also learn from various students. Consequently, groups were rearranged on Book 2. This also means that students were asked to adjust to new parties as well as practice on a new topic which differed from that of Book 1. On Book 2, 33 students “did not like to generate ideas in groups.” The top three reasons given are: a) Disagreement (8 students), b) It is boring (7 students), c) It is difficult (5 students). Group rearrangements may be one of the reasons that cause disagreement among students since students have to adapt themselves with the new group members and it may be unavoidable to have some argument. In addition, when students made Book 2, they followed the same steps as they created Book 1, so some students might feel it is boring to repeat the same procedures. Perhaps that is why seven students mentioned that this activity was boring on Book 2, but only two felt bored about this activity on Book 1. Furthermore, five students felt generating ideas with groups was difficult on Book 2, whereas only two students had the feeling on Book 1. The teacher participants encouraged the students to write longer sentences on Book 2 and perhaps this requirement had added some more difficulty on this activity.

Drafting Sentences Alone was the Hardest Activity

Question 3 about whether students like to write sentences by themselves or not received 66% support on Book 1 and Book 2 as well (Table 1). Though this is not the lowest percentage, the researchers would like to discuss this phenomenon. In fact, according to the interview data, drafting sentences alone was the most difficult activity. On Book 1, 12 out of 29 (41.4%) interviewees had this opinion and on Book 2, 20 out of 48 (41.7%) had the same statement. Most interviewees did not provide concrete answers about why they felt difficult; a few students stated that: “It is difficult to write sentences and I have to spend a long time to think about it.”

Writers have to concentrate in order to come up with letters and words and then they try to combine the words into sentences. Since children’s attention span is short (Paul, 2003), sitting quietly to write proper sentences becomes a challenge to them. Certainly, we cannot deny that students of the present study have very limited English knowledge. Students may have some ideas of writing, but this limitation may impede them while thinking of correct sentences.

Comparing Book 1 and Book 2, the researcher observer found that the students made great progress on Book 2. The researchers randomly selected six students’ (three boys and three girls) drafts from each class and counted for the average words and sentences. The data revealed that on Book 1 the average is 39 words and 9.13 sentences, and on Book 2, 53 words and 8.58 sentences. Though the average sentence on Book 1 and that on Book 2 were similar, the difference between the two numbers of average words seemed to imply that the students made longer sentences on Book 2 than on Book 1. The teacher participants also agreed that students’ writing improved a lot on Book 2. Though students may be not aware of their own progress in English learning and some of them even did not like to write sentences, their progress was observed and witnessed by the teacher participants and researchers as well.

Shy Character Resisting Self-Presentation

In terms of self-presentation, whether in speaking and writing, it is less favored by

students. Some students did not want their writing to be presented in front of the whole class or they did not like to read out in public. This may be related to Taiwanese students' shy character. Four out of 24 students reported that they felt shy when their written products were presented by their teacher and that was why they "dislike the sharing task." Eighteen out of 45 students had the same feeling on Book 2. Besides, on Book 2, 14 students pointed out that they did not like to see their writing presented in public because people might laugh at their works. Shyness also explained their feeling nervous or unconfident to read in front of all their classmates.

Children enjoy others' compliment. Receiving more positive feedback and praise, they may gradually deem themselves as writers more confidently (Dickson, 2004). One of Cambourne's (1995) suggestion of helping children learn naturally is to let children writers receive informative and non-threatening feedbacks. If students are confident about their products, they may not feel shy to present them. Perhaps, these students in this present study need more encouragement.

Longer Sentences Increasing Difficulty

One of the questions asked about whether students liked to cooperate with their group members to choose sentences and the data revealed that 19 students circled "dislike" on Book 1 whereas 47 on Book 2. On Book 1, the three main reasons for holding the negative attitude are: a) It is troublesome (5 students), b) It is tiring (3 students), and c) It is boring (3 students). However, on Book 2, the top three reasons are: a) It is too difficult to choose (10 students), b) It is troublesome (7 students), and c) It is boring (5 students). Based on the aforementioned data, we found that students had more difficulty on selecting sentences on Book 2 than on Book 1.

As mentioned, the students wrote longer sentences on Book 2 than on Book 1. Their being able to produce longer sentences might imply that they had gained more confidence of themselves as writers when they wrote Book 2. They might be thus proud of and cherished all the sentences produced and that might make the sentence-choosing task difficult.

Picture-Book-Making, Language Learning and Group Work

Did these students perceive this picture-book-making task as being beneficial to their English learning? Did they prefer working alone or in groups? Three questions were added in questionnaire 2. The questions and results are as shown in Table 2.

Table 2 *Picture-Book-Making, Language Learning and Group work*

Question	Response
1. Does the peer-proofreading activity make your writing better?	Positive: 70%
2. Do the picture-book-making activities make your English better?	Positive: 60%
3. Do you prefer to make a book alone or in groups?	Group: 79%

Question 1 received 70% positive support, which means that 85 out of 122 students considered peer-proofreading activity being helpful. Moreover, 41 students gave reasons for this response and the top two reasons are: a) My sentences are better (29 students), and b) My vocabulary is more correct (6 students). This finding indicates that more than half of the students admitted that peer feedback is helpful to them. Since most students have perceived the benefit of this peer-proofreading activity, maybe this peer-proofreading practice should be continued. Students who can criticize others' writing have more chances to become critical readers and then revise their own writing more productively (Rollinson, 2005). Practicing more proofreading skills may help students become better writers and readers as well. Sixty percent of them felt that the writing activities helped them learn English better. Seventy-nine percent of them preferred to work in groups.

Teacher Participants' Perceptions of Picture-Book-Making Activities

Research question 2 discusses what the instructors' perception of the picture-book-making activities is. The data was collected by interviewing, in Mandarin, two teacher participants with planned interview questions. Some of their responses to the interview questions are presented below.

Reflecting Teaching

The two teacher participants both claimed that examining students' writing products helped them reflect on their own teaching. For example, responding to the interview question *What have you learned from this picture-book-making project*, Ms. Hsu stated,

"I thought students had understood the words, but when they wrote, some of them either could not know how to write them or make a lot of mistakes. By examining their writing, I realized on which part students might need more help."

Paul (2003) asserts that another benefit of writing is its characteristic-visible. He states that by examining the learners' written products helps the instructors understand what the children have comprehended and what they haven't and more importantly, their products help both the instructors and learners themselves clearly understand their learning progress. Ms. Hsu's reflection exactly corresponded to Paul's (2003) assertion.

Applying What Have Learned

Both teacher participants were asked what they expected their students to learn from this picture-book-making task. Ms. Hsu expressed,

"I hope students will be able to apply what we have taught on their writing and it is obvious to see their application on Book 2...."

Ms. Chang stated,

"I hope their sentences are full of abundance, and not merely short and simple sentences. On Book 2, their sentences became longer and they used many adjectives to describe their sentences...."

The themes of the picture-book-making activities were "zoo" and "market." They were chosen because the student participants had learned some related vocabulary and sentence patterns on them. Since the topics were associated with the students' learned knowledge, students could make the best of their acquired knowledge.

Hard to Keep Classroom Order

One interview question dealt with which part of the picture-book-making task was the most difficult and easiest for teachers to handle. Ms. Hsu thought the most difficult part of her teaching process was keeping the classroom order because the student participants often lost attention while discussing with peers. For Ms. Chang, time limitation was a problem. As she stated as follows:

"... We have time pressure and that is why we cannot provide enough time for students to write...."

Hard to Write, Easy to Brainstorm

Responding to the question on which part was the most difficult and easiest for their students, Ms. Hsu explained that the student participants had limited English knowledge about vocabulary and sentence patterns and this resulted in the difficulty of students to write beginning and ending sentences. When the students wrote beginning and ending sentences, they had to think about how to connect their beginning and ending pages with their responsible pages as a whole book, and thus this certainly increased the difficulty of their writing. As for Ms. Chang, she stated that most student participants encountered problems when they drafted their sentences individually. She stated,

"When students wrote three layers of clusters, they basically just used those three words to write sentences. They repeatedly use the same words."

Students were taught to use clusters to generate ideas, but they may be too persistent on the

words they had written and did not think of another words or sentence patterns. However, both teacher participants expressed that the student participants could brainstorm to complete clusters effortlessly since they had experienced such activity for almost two years.

Language Skills Enforced and Cooperation Learned

On which part did the students get the most and the least benefits? Ms. Hsu discovered that not only the student participants' writing ability was reinforced, but also their reading ability. As she claimed,

“Students are more acute of writing, and they are more familiar with the basic grammar concept of English. They had improved their sentence writing, vocabulary knowledge and reading. Specifically, when they presented their work, their faces were full of satisfaction.”

Writing skill is enhanced by practicing writing and furthermore, the other language skills are reinforced as well (Davison & Dowson, 2003). Ms. Hsu's observation seemed to prove Davison and Dowson's assertion. If the writing project can go through six years, the two teachers both hope their students would be able to write a short paragraph which contains complete content. Besides, Ms. Chang believed that the best benefit at this stage was that the student participants understood the concept of cooperation which was also her expectation of the writing activities.

Just Try it, Writing

Finally, both teacher participants pointed out that writing instruction was not difficult. As long as teachers give students the opportunity of doing so, students should be able to do it though it was in fact a tiring work. As Ms. Hsu claimed that:

“I went to two schools to share my experiences of teaching writing. The listeners all felt it was too tiring, but actually, when I saw students' learning achievement, I do not feel tired at all. You just put writing into your schedule and then you will realize it is possible to do it because I really have seen the benefits of writing on students.”

CONCLUSION AND IMPLICATION

Overall, the participants in the current research reported to have a positive attitude towards the picture-book-making activities. The student participants enjoyed in those activities while the teacher participants also discovered the benefits of incorporating writing into the curriculum. About 60% of the student participants felt such activities were helpful to their language learning and furthermore, they wanted to do it again in the following semester. The teacher participants claimed writing instruction resulted in the win-win situation. Not only did it facilitate students' learning, but it also helped teachers to monitor themselves to see whether students needed some more assistance.

The student interviewees' data revealed that the student participants considered these picture-book-making activities as being beneficial to their English learning and this was confirmed by their instructors. The teacher participants, according to the interview data, expressed that they had witnessed the students' huge progress in English learning, such as the length and diversity of sentences and the use of various vocabulary words, especially in Book 2. Though writing has stereotypes that are to be too difficult to teach, the teacher participants both claimed that it was not so unreachable. Perceiving the dramatic progress on their students even persuaded them of doing the right things for students.

Researchers had proposed that writing is conducive to learners (Chiu & Chang, 2006; Sejnost & Thiese, 2007), students need a more balanced curriculum design (Davies & Pearse, 2000; Paul, 2003), and cooperative learning can make learning the most beneficial (Johnson, R. T. Johnson, & Holubec, 1994; Kagan, 1995). In the current research, the findings confirmed the aforementioned scholars' view. EFL instructors may question the appropriateness of employing writing on young beginning learners; however, the findings of

the current research indicated that young beginning learners were able to do it. Also, for either the instructors or learners, writing in fact is not an impossible task. It is achievable as long as the instructors want to conduct it. Children may not decide for themselves what to learn, but adults should not deprive children of learning opportunities (Scott & Ytreberg, 1990). Since writing has numerous advantages, why not giving children the opportunity of experiencing it?

Certainly, we cannot deny that such curriculum design needs a lot of people's effort. In the present study, the two teacher participants are lucky to have three helpers, two academics and a graduate student, but perhaps other English teachers are not as lucky as they. Nevertheless, curriculum design is flexible. The instructors can adjust the curriculum to make it more suitable for their own learners depending on the school's needs. One thing we can keep in mind is that writing is not unreachable. If we can, only when we can provide students with the opportunity of learning writing, students then will gain the benefits of learning writing. Even though with just a 5-minute activity of free writing, students can still experience how writing elicits their current language knowledge. Students in EFL context do not have so many opportunities to speak out or practice what they have learned, but writing on the other hand, may provide them with a resource of performing their language proficiency. Leki (2005), in particular has indicated that writing "may well reflect the learners' best performance" (p. 88). And this reflection also helps teachers to monitor, to detect, to improve their way of teaching. When they understand in which part their students may need some help, their teaching will be the most profitable. It is hoped the findings of the current research can shed some light on the understanding of how to incorporate writing into EFL settings. More related research is needed in order to examine whether writing should be integrated into EFL's primary school curriculum design or not.

REFERENCES

- Cambourne, B. (1995). Toward an educationally relevant theory of literacy learning: twenty years of inquiry. *The Reading Teacher*, 49(3), 182-190.
- Chiu, L. C., & Chang, M. Y. (2006). *Using predictable books to teach writing in an EFL setting*. Paper presented at the Fifteenth International Symposium and Book Fair on English Teaching, Taipei.
- Culham, R. (2004). *Using Picture Books to Teach Writing with the Traits*. NY: Scholastic Inc.
- Cunningham, P. M., Moore, S. A., Cunningham, J. W., & Moore, D. W. (2004). *Reading and writing in elementary classrooms: Research based K-4 instruction* (5th ed.). Boston, MA: Allyn & Bacon.
- Davies, P., & Pearse, E. (2000). *Success in English teaching*. HK: Oxford University Press.
- Davison, J., & Dowson, J. (2003). *Learning to teach English in the secondary school* (2nd ed.). London and New York: RoutledgeFalmer.
- Ehlert, L. (2000). *Market Day*. San Diego: Harcourt.
- Fox, M. (1996). *Zoo-Looking*. NY: Scholastic.
- Galda, L., Cullinan, B. E., & Strickland, D. S. (1997). *Language, literacy and the child* (2nd ed.). Orlando, FL: Harcourt Brace College Publishers.
- Ghosn, I. K. (1996). *Whole language EFL with children's literature: The way it worked in one kindergarten class*. Paper presented at the 28th Annual Meeting of the Teachers of English to Speakers of Other Languages, Baltimore, MD.
- Jalongo, M. R. (2007). *Early childhood language arts*. Boston: Pearson/Ally and Bacon.
- Johnson, D. W., R. T. Johnson, & Holubec, E. J. (1994). *The new circles of learning: Cooperation in the classroom*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Kagan, S. (1995). We can talk: Cooperative learning in the elementary ESL classroom. *ERIC*

Digest, ED382035.

- Lamme, L. L., Fu, D., Johnson, J., & Savage, D. (2002). Helping kindergarten writers move toward independence. *Early Childhood Education Journal*, 30(2), 73-79.
- Leki, I. (2005). *The challenges of teaching EFL writing*. Paper presented at the 14th International Symposium on English Teaching and Book Exhibition., ETA
- Morrow, L. M. (1996). *Literacy Development in the Early Years* (3rd ed.). Boston: Allyn and Bacon.
- Paquette, K. R. (2007). Encouraging Primary Students' Writing through Children's Literature. *Early Childhood Education Journal*, 35(2), 155-165.
- Paul, D. (2003). *Teaching English to children in Asia*. Hong Kong: Longman Asia ELT.
- Peregoy, S. F., & Boyle, O. F. (2005). *Reading, Writing, and Learning in ESL: A Resource Book for Teaching K-12 English Learners*. MA: Allyn and Bacon.
- Rollinson, P. (2005). Using peer feedback in the ESL writing class. *ELT Journal*, 59(1), 23-30.
- Scott, W. A., & Ytreberg, L. H. (1990). *Teaching English to children*. Essex, England: Longman.
- Sejnost, R. L., & Thiese, S. (2007). *Reading and writing across content areas*. Thousand Oaks, Calif.: Corwin Press.
- Temple, C., Nathan, R., Temple, F., & Burris, N. A. (1993). *The beginnings of writing* (3rd ed.). Boston, MA: Simon & Schuster, Inc.
- Vygotsky, L. (1978). *Mind in society: Development of higher psychological process*. Cambridge: Harvard University Press.
- Minister of Education (MOE) 教育部 (2005)。國民中小學九年一貫課程暫行綱要，修訂版。
- Dickson, D. (2004). 紐西蘭小一學生的寫作歷程：寫作世界真奇妙 (謝幸儒, Trans.). In 蘇妙慧 (Ed.), *灌溉熱情讓英語學習得種子萌芽 All about passion*. Taipei: 東西文化.
- Li, S. H. (2004). 李信宏 健康面對英檢-語法和語用平衡的教學思維. In 蘇妙慧 (Ed.), *語言功夫 Language empowerment*. Taipei: 東西出版社 East & West Book Co., Ltd.

APPENDIX A Questionnaire for student participants on Book 1

(presented only part of the questionnaire)

Part I : Question sheet for the teacher participant to read

1. 你喜歡這次的圖畫書創作分組活動嗎?
2. 你喜歡英文課時和你那一組同學一起畫泡泡圖的活動嗎?
3. 和同學一起完成泡泡圖後，有一個功課是要根據你自己的泡泡圖，自己寫 5 個句子，你喜歡這個寫句子的活動嗎?

APPENDIX B Interview questions for student participants on Book 1

(presented only part of the questionnaire)

1. 你喜歡這次的分組活動嗎?為什麼?
2. 你喜歡英文課時和你那一組同學一起畫泡泡圖的活動嗎?為什麼?
3. 和你那一組同學一起畫泡泡圖後，有一個功課是要根據你自己的泡泡圖，回家寫 5 個句子，你喜歡這個活動嗎?為什麼?

APPENDIX C Questionnaire for student participants on Book 2

(presented only part of the questionnaire)

1. 有一個活動是同組成員一起挑選小書的句子，你喜歡這個活動嗎?為什麼?
2. 有一個活動是根據自己挑選的句子，回家畫插圖，你喜歡這個活動嗎?為什麼?
3. 你喜歡看到自己的作品被老師展示給所有同學看嗎?為什麼?

APPENDIX D Interview questions for student participants on Book 2

(presented only part of the questionnaire)

1. 你覺得整體而言，這個整組一起塗塗寫寫的活動會很難或很簡單嗎?為什麼?
2. 哪個部分比較難或比較簡單?為什麼?
3. 你比較喜歡自己塗塗寫寫還是和同學一起創作?為甚麼?

APPENDIX E Interview questions for teacher participants

(presented only part of the questionnaire)

1. According to your English teaching experience, what are the problems you have seen for primary school students in terms of English learning?
2. Before you cooperated with the two professors on this writing practice two years ago, did you ever try with any English writing activities in your class before? Why or why not?
3. Before you cooperated with the two professors on this writing practice two years ago, what was your perception of writing in EFL curriculum?

Scoring Collaborative Oral Interaction in EFL Classrooms:

What is Fair?

Angie H.C. Liu (劉蕙君)

Chung Yuan Christian University

angiehcliu@cycu.edu.tw

Despite of reported benefits of collaborative language learning activities, the assessment of group-constructed oral interaction is often avoided in EFL classrooms due to challenges related to the scoring of group performance and concerns about test fairness. This study investigated the issue of fairness in scoring collaborative oral interaction based on the perception of major stakeholders of a test – test takers. A total of 114 EFL college students, in a group of three, participated in two group-based oral discussion assessment tasks. Immediately after the test, they were asked to rate the performance of their group as well as their own. They were also asked to reflect on the fairness and validity of two scoring approaches (i.e., group-based versus individual-based scoring). Overall, the agreement between ratings of group and individual performance was found to reach at a moderate level. Analysis of the participants' retrospective written response further showed that the support for the two scoring approaches was equally strong. The main argument for the individual-based scoring grounded in its perceived fairness – that is, group-based assessment needed to account for differential contributions from individual test takers. On the other hand, the primary argument for the group-based scoring lied in its perceived validity – that is, individual test performance is built upon collaborative efforts and hence, could not and should not be evaluated independent of group performance. The pedagogical implications of the study findings for ELL teachers were also discussed.

INTRODUCTION

As more language educators, policy makers, and learners of English language (ELLs) recognize the critical role of language use and the beneficial impact of collaborative learning on the development of second language proficiency, there has been increasing popularity of task-based, interactive instructional activities in second language classrooms (Lantolf, 2000; Naughton, 2006; Nunn, 2000; Swain, 2000). It is contended that oral interaction in second language classrooms provides the platform for learners to engage in mutual scaffolding through the interactive process of meaning negotiation and metacognitive reflection on linguistic forms (Naughton, 2006; Swain & Lapkin, 1998). Aligned with the movement of communicative language learning, performance-based testing via contextualized tasks has also received great attention in the field of language assessment in recent decades (Hudson, 2005). It is maintained that group-based tasks provide learning opportunities for ELLs to practice important interaction skills such as distributing and competing for opportunities to speak, holding the floor, adjusting to the contribution of other speakers and negotiating mutual understanding when engaging in opinions sharing (Nunn, 2000). Additionally, it is argued the format and design of group-based oral assessment places ELL test takers in equal power position in relation to each other and thus, brings out the kind of interaction that is not possible to be elicited between an expert-assessor and a test taker (Luoma, 2004).

Considering the beneficial backwash of performance-based group oral tests on second language learning, there are compelling reasons to align instructional tasks with assessment practice in EFL classrooms. Nonetheless, despite of reported benefits of group-based language learning activities, the assessment of group-constructed oral interaction in EFL classrooms is often avoided due to problems and concerns related to the scoring of such a complex and dynamic construct. When the test discourse is jointly created by multiple test

takers, the accuracy, consistency and fairness of scoring the collaborative performance poses real challenges to raters considering the various parameters involved, such as language proficiency level and communication style of the individual test group member, and familiarity among group members (Brown, 2003; McNamara, 1997). Can the oral interaction co-constructed by multiple ELL test takers be assessed in a fair and valid manner? This study aims at researching the issue of test fairness and validity in scoring collaborative oral interaction in EFL classrooms from the perspective of major stakeholders of a test – test takers.

Research Questions

Specifically, the empirical investigation of scoring collaborative oral interaction in EFL classrooms was operationalized through the study of several interconnected questions:

1. Whose performance on the collaborative oral interaction should be assessed?
2. How should the scoring of collaborative oral interaction be conducted – on a group-basis or individual-basis?
3. To what extent does the scoring approach affect the assessment outcome of collaborative oral interaction?

REVIEW of RELEVANT STUDIES

Nature and Dynamics of Collaborative Oral Interaction

A number of studies (Fulcher, 1996; Morrison & Lee, 1985; Reves, 1991; Shohamy, Reves & Bejarano, 1986) reported the use of group-based tasks to assess the oral proficiency of English language learners in various countries (i.e. Israel, Zambia and Hong Kong) and institutional settings (i.e., high schools and universities). These studies found that English language learners tended to hold positive views toward the use of group-based assessment tasks, especially if they also participated in the assessment process. It was also found that the performance condition of group oral tasks in general reduced the test taker's anxiety level to some extent so that it was easier for them to engage in the interaction with other test participants. Similarly, Fulcher (1996) reported that the group oral discussion was perceived by test takers to be a much more natural task. Nonetheless, this positive perception of group-based assessment has yet been supported by empirical evidence at the discourse level.

The acquaintance between group members of collaborative tasks and its impact on the task performance was also studied by a number of second language testing researchers (e.g., O'Sullivan 2002; Porter, 1991). Porter found no evidence in support of the hypothesis that the familiarity with one's partner in an interactive task positively affected performance. Contradictory to Porter's finding, O'Sullivan reported evidence of an 'acquaintanceship' effect such that ELL test takers achieved higher scores when collaborating with friends. The result of this study appeared to be consistent with a number of second language learning studies, which suggested that learners modified their language when interacting with speakers of various degree of familiarity (Plough & Gass, 1993; Tarone & Liu, 1995). However, analysis of the test performance indicated that there was no effect on linguistic complexity and that there was a 'sex-of-interlocutor by acquaintanceship' interaction effect on linguistic accuracy. Cao and Philip (2006) studied the relationship between interactional context of second language communication and the language learner's willingness to communicate by comparing their behavior in whole class, group and dyadic interaction and concluded that contextual factors such as group size, familiarity with interlocutor(s) and discussion topics, self-confidence and medium of communication, did exert real influence on their language output. Moreover, Weir (1990) cautioned against the use of collaborative assessment tasks as it was highly likely that test takers of higher proficiency may dominate the interaction.

Assessment Tasks Used to Elicit Collaborative Oral Interaction

The design of assessment prompts for eliciting group-based oral interaction is a complex

issue and has been the subject of numerous studies in the fields of second language acquisition, instruction and assessment (Fulcher, 1996). Quite a few studies examined various aspects of elicitation tasks of group-constructed oral interaction, such as performance conditions and contexts, types and characteristics of assessment prompts and execution variability and their impact on the oral performance elicited. Empirical findings of Fulcher's study concluded that the design of the assessment produced limited and negligible impact on the final performance scores as long as the rating scale did not refer to test method facets in its descriptors. That implies, despite of the interaction between attributes of assessment tasks and the ability underlying group-constructed oral interaction, generalization from one task to another is possible if an appropriate rating scale is developed and effectively implemented.

In terms of the task type for eliciting group-constructed interaction, Shehadeh (2001) suggested that consensus-reaching / decision-making tasks to be most effective in that this type of task structure tended to provide all participants equal opportunities to engage in the group interaction. In addition, studies investigating the impact of pre-task planning time on the production of speech samples generally reported positive effects, especially in terms of fluency and complexity; however, mixed results have been reported in terms of language accuracy (Crookes, 1989; Ellis, 2005; Ortega, 1999). For example, Yuan and Ellis (2003) compared the effects of both pre-task and on-line planning on L2 oral production and found that pre-task planning enhanced grammatical complexity while on-line planning positively influenced accuracy, grammatical complexity, fluency and lexical richness.

Scoring Group-based Performance

The scoring of group work is a complex and intricate task – on the one hand, the performance outcome is a collective effort of groups and on the other hand, it is the individual that contributes toward the ultimate group product. Some researchers (e.g., Nicolay, 2002) opposed the assignment of individual scores to group-based projects, contending that the group should be evaluated as a unit and that any attempt in isolating individual contribution negated the principal of group work. Yet, oftentimes, for stakeholders of assessment (e.g., students, teachers, parents, schools and employers), there is a pre-conceived notion of unfairness associated with differential contributions from individual test takers and a practical demand in assigning scores to individual test takers based on their contributions to the group outcome (Nordberg, 2008).

A good number of studies have researched ways that the scoring of group-based performance can be achieved in a more valid and equitable manner (e.g., Boud et al., 1999; Falchikov, 1986; Hughes & Large, 1993; Lejk & Wyvill, 2001; Johnson & Lynden, 2004; Nordberg, 2008). Some argued that the scoring of group projects should be conducted by members of the team or individual members themselves because they were in a much better position than external assessors to judge individual contributions to the group work from various aspects, such as generation of ideas, participation in group discussion, preparation of project materials and contribution to study design, analysis and reporting (e.g., Johnston & Miles, 2004; Magin, 2001). Others criticized the unreliability of peer and self judgments and contended that expert assessors were more qualified to evaluate group performance objectively (Falchikov, 1991; Krause & Popovich, 1996; Lejk & Wyvill, 2001). Most, however, agreed that both peers and expert assessors provided valid information toward the process and product of group work and hence, feedback from both groups should be incorporated into the scoring of group-based performance (Lejk & Wyvill, 2001; Nordberg, 2008; Sharp, 2006). Nonetheless, they differed in terms of how these two sources of evaluation should be combined. One line of proposals contended that all individuals should basically receive the same score (i.e., the group score), adjusting only for exceptionally substandard or outstanding individual test takers (Lejk et al., 1996). The majority of researchers; however, adopted a statistical approach in developing a weighting scheme to

combine ratings of individuals (on the basis of peer and/or self-assessment) with the group rating to arrive at individual scores (Lejk & Wyvill, 2001; Nordberg, 2008). For example, it was suggested that the peer assessment scores formed a fixed fraction of the individual students' overall score with the other fraction coming from the group performance score (e.g., Bushell, 2006). Alternatively, it was proposed that the peer assessment outcome could be used as a multiplying factor for individual scores based on the group score (e.g., Johnston & Miles, 2004; Nordberg, 2008).

METHOD

Study Design

First, each study participant was randomly assigned to a test group of three. Next, each test group was randomly assigned to one of the two discussion tasks. Each group was then allotted five-minute pre-task planning time for a four-minute group discussion. At the same time, they were informed of the scoring approach (i.e., group-based or individual-based scoring) that would be used to assess their discussion performance. A counter-balanced task administration procedure was implemented to avoid any potential interaction between assessment tasks and scoring approaches. Immediately after completing each collaborative task, participants were asked to evaluate the oral performance of their own and their test groups using the rating scale provided. Training on the use of the rating scale was provided prior to the assessment phase of the study. Finally, they were asked to respond to the open-ended reflection questionnaire, designed to elicit their perception toward the validity and fairness of group-based oral interaction in the classroom setting.

Participants

A total of 114 college freshmen from two intact classes in Chung Yuan Christian University were recruited for this study, with 59 of them from the Department of Teaching Chinese as a Second Language and 55 of them from the Department of Special Education. The average performance of the two classes on the Michigan English Proficiency Test administered at the beginning of their freshman year is approximately comparable (mean_{AC} = 56.34 and std_{AC} = 13.12; mean_{SE} = 55.87 and std_{SE} = 11.65) and their overall level of English proficiency can best be characterized as intermediate.

Instruments and Materials

Assessment Prompts. Two non-convergent discussion prompts were constructed to elicit the collaborative oral interaction among members of the test groups. Both prompts require participants to react to a crisis by discussing the actions they would take to handle it (see the Appendix). All three members of the test group received exactly the same task prompt and were specifically instructed that they did not have to reach a consensus on the solution. In order to successfully complete the task, participants needed to discuss the scenario posted by stating their views and opinions, expressing agreements and disagreements and supporting their arguments with facts and logical reasoning. Key words of the assessment prompts were supplemented with Chinese translation.

Rating Scale. A ten-point scale (i.e., 1 – very poor to 10 -- outstanding) was used to record participants' evaluation of their own and group oral performance on the collaborative tasks. Assessors were asked to take four specific criteria into consideration – fluency, content (quality and quantity of ideas presented), communication (clarity of expression), and interaction (participation in discussion) while evaluating test performance. Nonetheless, they were asked to produce a holistic assessment outcome based on their evaluation.

Data Analysis

The first and second research questions concerning the object and method of scoring collaborative oral interaction were examined through the tabulation of fixed responses

selected. That is, frequency tables were generated for each option of the scoring object and method. In addition, trends and patterns of responses on the open-ended questions were identified, categorized and summarized based on qualitative analyses of participants' retrospective written reflection. The third research question concerning the interaction between the scoring approach and assessment outcome was studied by correlation analysis and paired-samples T-test between scores assigned on the basis of individual and group performance.

RESULTS

Research Questions One and Two

The survey results of the perceived measurement object of collaborative oral interaction showed that approximately equal proportions of study participants viewed either individual performance (47.4%) or group performance (50.0%) as the legitimate target of assessment (see Table 1), with strong advocates for each. The top five reasons reported in support of the validity of individual-based and group-based scoring approaches were displayed in Tables 2 and 3, respectively. The key aspects underlying the test takers' perceptions toward group-based oral assessment include parameters relating to the concepts of test fairness, accountability, accuracy, group dynamics and group interaction.

Table 1. *Perceptions toward Scoring Target of Collaborative Oral Interaction*

Scoring Target	Count	Percent
Individual performance	54	47.4
Group performance	57	50.0
Both	1	0.9
No Response	2	1.8
Total	114	100.0

From the perspective of fairness and accuracy, advocates of individual-based scoring of group-based oral interaction maintained that individual members of the test group contributed differentially to the collaborative interaction and therefore, it was not fair to award all members with the same score. They argued that if the same score were assigned to all members of the group, it would imply that the level of participation and oral proficiency of each group member was the same, which clearly did not reconcile with their evaluation of the collaborative task process and performance. From the perspective of accountability, they claimed that individual-based scoring, in a way, forced each member of the group to actively participate in the group discussion since each of them would be held accountable for their individual performance and could not rely on other group members to cover for them. They further contended that since each member's performance was affected by those of other members, assigning scores at the individual level was the only way to minimize the performance interference among group members.

Table 2. *Top Five Reasons Supporting Individual-based Scoring of Collaborative Oral Interaction*

	Reason
1	Differential contributions from individuals toward the group performance ought to be recognized.
2	It's not fair to award the same credit to individuals when they don't perform at the same level.
3	Individual-based scoring can reflect each individual's participation more accurately.
4	Scores for individual test takers should not be affected by other members of the group.
5	Individual-based scoring encourages active participation from all members of the group because each one of the group will strive for demonstrating their best performance.

On the other hand, advocates of group-based scoring viewed collaborative oral interaction to be a joint product of all group members and thus, could not be individually dissected and scored. From the perspective of group dynamics and interaction, they claimed that despite of differential participation levels and contributions of individual members to collaborative oral interaction, everyone engaging in the group discussion worked toward the common goal of producing the best group performance. It was an interactive and dynamic process, requiring all members of the group to participate in the co-construction of the assessment discourse. Moreover, they contended that assigning scores at the group-level encouraged collaboration and motivated interaction among group members, which subsequently resulted in better group outcome. The group discussion was less likely to be dominated by higher-proficiency individuals when the group performance was the basis for the final score. When the floor for oral discussion was more evenly distributed among all members of the group, it ended up producing a more natural and 'real-life' type of oral interaction with more frequent turn-taking. Therefore, they argued that from the perspective of validity, group-based scoring was a more accurate indicator of ELL's ability of collaborative oral interaction.

Table 3. *Top Five Reasons Supporting Group-based Scoring of Collaborative Oral Interaction*

	Reason
1	Collaborative oral interaction is a joint product of all group members and thus, the performance cannot be individually dissected and assessed.
2	Group-based scoring encourages collaboration and motivates interaction among group members, which subsequently results in better group outcome.
3	The discussion floor won't be dominated by higher-proficiency individuals and thus, all members obtain more equitable shares of participation opportunity.
4	Group-based scoring provides a more accurate assessment of group collaboration, which is the key to a more natural, 'real-life' type of oral interaction.
5	Group-based scoring reduces performance anxiety.

Research Question Three

Overall, the agreement between ratings of group and individual performance on the collaborative oral discussions, as shown in Table 4, was found to be at a moderate level ($r=.509$, $df=114$, $p<.0001$). Specifically, the ratings of collaborative oral discussions (see Table 5) showed that EFL test takers' averaged assessment of group performance (mean=7.04 on a 1-10 scale) were higher than those of individual performance (mean=5.82 on a 1-10

scale). At the same time, the variance of group-based performance (standard deviation=1.381) was found to be smaller than that of individual-based performance (standard deviation=2.015). Results of the paired-samples T-test further indicated that the difference between individual- and group-based outcomes of collaborative assessment tasks was statistically significant at the .01 level (see Table 6). That means, the assessment outcome of collaborative oral interaction was statistically different depending on the scoring approach that was used to assign ratings.

Table 4. *Correlation between Assessment Outcomes of Individual and Group Scoring*

Assessment Outcome		Individual	Group
Individual	Pearson Correlation	1	.509**
	Sig. (2-tailed)		.000
	N	114	114
Group	Pearson Correlation	.509*	1
	Sig. (2-tailed)	*	
	N	.000 114	114

Note. ** indicates correlation is significant at the .01 level based on a 2-tailed test.

Table 5. *Individual-based versus Group-based Assessment Outcomes*

	Score Mean	Std. Deviation	N
Group Performance	7.04	1.381	114
Individual Performance	5.82	2.015	114

Table 6. *Comparison between Individual-based versus Group-based Scoring*

	Paired Difference			
	Mean	Std. Dev.	df	Sig. (2-tailed)
Group Score – Individual Score	1.222	1.770	113	.000

DISCUSSION

The Target and Method of Scoring Collaborative Oral Interaction

The validity of a test is built upon the strength of the evidence in support of the accuracy of a test in measuring what it intends to measure and thus, the foundation of test validation lies in defining the measurement object of a test. When test discourse is co-constructed by multiple test takers in real time, how should the construct underlying the joint performance be defined? And, what is the scoring target of collaborative oral interaction? EFL test takers' reflection in this study clearly showed that their views toward the assessment target of collaborative oral interaction were equally split between individuals and groups. Such a result appeared to echo the current debate between psychometric and dynamic approaches of language assessment. The traditional psychometric assessment derives from an individual-based model of language development. In this framework, the test taker's solo performance is assumed to account for 100 percent of the observed performance variations and represents a complete picture of the individual's underlying language competence in an

ideal world (Poehner & Lantolf, 2005). The dynamic approach of assessment, on the other hand, takes into account the mediation received from other test participants and aims at assessing the extent to which the assistance benefits the test taker's ultimate performance (Sternberg & Grigorenko, 2002). The ability to incorporate the assistance received and to transcend the newly developed ability to the assessment task is viewed as a true indication of the test taker's underlying language competence.

These two different views toward the intended measurement object of collaborative oral interaction naturally lead to different interpretations of test validity and fairness. The reasons reported by proponents of individual-based scoring clearly are in line with the fundamental assumption of psychometric testing – that is, the test taker's underlying language competence is stable in the course of assessment despite of the observed performance variations and the standardization of test administrations is a key component of test reliability and validity. Consequently, differential contributions from the individual test taker to the collaborative interaction need to be separately identified and rewarded accordingly. Any scoring procedure that fails to achieve such a goal is viewed as 'invalid' and 'unfair'. However, in the framework of dynamic assessment, the constant change and modification of mediated language performance during assessment process is viewed as an indispensable component of the procedure rather than the cause for test unreliability (Lantold & Poehner, 2004). Therefore, collaboration and assistance from members of the test group is viewed by proponents of group-based scoring to be an integral part of the test construct and not considered to be a source for test unfairness or invalidity. In other words, the concept of test fairness developed under the psychometric framework does not apply to the case of group-based assessment because the assessment prompt used to elicit the test taker's best performance is no longer static. The notion of standardization does not work either due to the dynamic nature of collaborative oral interaction. Therefore, test fairness cannot be viewed from the perspective of administration procedure, assessment context or subgroups of test population. Instead, it is the view toward the intended measurement object of collaborative oral interaction that determines the notion of test fairness. Simply stated, it is the answer to the central question – 'whose performance on the collaborative oral interaction should be assessed?' - that defines what fairness is and how it can be achieved.

Interaction between Scoring Method and Assessment Outcome

The statistical analysis of the EFL test takers' evaluation of individual and group performance revealed that the assessment outcome of collaborative oral interaction was indeed affected by the scoring method employed (i.e., individual scoring versus group scoring). Specifically, the averaged test score assigned on the basis the group performance was higher than that of the individual performance. The statistically significant score difference between the group and the individual on the same assessment task implies that many EFL test takers of group oral discussion task recognized the differential contributions and performance from individual members of the test group and adjusted their assessment of the ability underlying the collaborative oral interaction accordingly. Such an outcome suggests that despite half of the EFL participants expressed the view that the group performance should be the intended measurement object of collaborative oral interaction, the scores assigned by them revealed a different story. It seems that philosophically, many of them took the stance that the same test score should be awarded to every member of the group since collaborative oral interaction is based on joint performance and a product of interactive process in which individual test takers participate and co-construct the discussion. Nonetheless, when it was time to assign scores, their action, which perhaps derived from the notions of fairness and accountability in the traditional psychometric framework, pointed at a different direction. Another possible explanation for the discrepancy between self-reported reflections and results of test scores could be that most EFL students who are trained in the

educational system in Taiwan are accustomed to an assessment system that traditionally rewards individual excellence as a result of competition among individuals. Therefore, they are not capable of focusing only on the collaborative performance of the group and ignoring differences of individual test takers.

CONCLUSIONS AND PEDAGOGICAL IMPLICATIONS

The concept of test fairness is highly critical in an assessment system because test users often view it as the foundation for the validity and accountability of the system. This study investigated the issue of fairness in scoring collaborative oral interaction in EFL classrooms from the perspective of one major shareholder group of a test - the examinee, and the result revealed a complex and somewhat contradictory picture of group oral assessment. On the one hand, both individual and group performance received approximately equal support from the test takers to be the intended measurement object of group-based collaborative oral assessment. On the other hand, the assessment outcome of collaborative oral interaction by EFL college test takers was found to be statistically different depending on the scoring approach adopted.

Testing is a means to collect performance information on test takers so that inferences can be made about their ability in non-test situations. Different kinds of assessment tasks elicit different types of test taker response, which subsequently facilitates different types of learning behaviors. Therefore, it is essential to match the assessment tasks with the type of response that will facilitate the language learning that is intended to be promoted. If the ability to use English to carry out group discussion is indeed a learning objective valued in college EFL classrooms, English teachers need to find a way to assess the collaborative performance in a valid and fair manner. Since both individual-scoring and group-scoring received comparable amount of support from EFL test takers, perhaps a scoring scheme that combines/integrates both methods is something worth pursuing. For example, one score is assigned to the group as the base score and then any exceptional performance of individual member is rated and adjusted on top of the base score.

The current study of collaborative oral assessment in EFL classrooms is limited by the empirical evidence collected from EFL test takers only. It is recommended that future studies of group-based assessment examine and analyze actual turn-taking patterns of the co-construction process underlying collaborative interaction in the assessment setting as well as the quality of the test discourse product when different scoring methods are used to determine test outcomes.

REFERENCES

- Boud, J., Cohen, R. & Shampson, J. (1999). Peer learning and assessment. *Assessment and Evaluation in Higher Education*, 24(4), 413-26.
- Brown, A. (2003). Interviewer variation and the co-construction of speaking proficiency. *Language Testing*, 20(1), 1-25.
- Bushell, G. (2006). Moderation of peer assessment in group projects. *Assessment and Evaluation in Higher Education*, 31(1), 91-108.
- Cao, Y. and Philp, J. (2006). Interactional context and willingness to communicate: A comparison of behavior in whole class, group and dyadic interaction. *System*, 34, 480-93.
- Crookes, G. (1989). Planning and interlanguage variation. *Studies in Second Language Acquisition*, 11, 367-83.
- Ellis, R. (2005). *Planning and task performance in a second language*. John Benjamins Publishing Co.
- Falchikov, N. (1986). Product comparison and process benefits of collaborative peer group

- and self assessments. *Assessment and Evaluation in Higher Education*, 11(2), 146-65.
- Falchikov, N. (1991). Group process analysis: self and peer assessment of working together in a group. In S. Brown & P. Dove (Eds.) *Self and peer assessment*. Birmingham: SCED Publications.
- Fulcher, G. (1996). Testing tasks: issues in task design and the group oral. *Language Testing*, 13(1), 23-51.
- Hughes, I.E. and Large, B.J. (1993). Staff and peer-group assessment of oral communication skills. *Studies in Higher Education*, 18, 379-85.
- Hudson, T. (2005). Trends in assessment scales and criterion-referenced language assessment. *Annual Review of Applied Linguistics*, 25, 205-27.
- Johnston, L. & Lynden, M. (2004). Assessing contributions to group assignments. *Assessment and Evaluation in Higher Education*, 29(6), 751-68.
- Johnston, L. & Miles, L. (2004). Assessing contributions on group assignments. *Assessment and Evaluation in Higher Education*, 29(6), 751-68.
- Krause, J. & Popovich, N. (1996). A group interaction peer/self assessment process in a pharmacy practice course. *American Journal of Pharmaceutical Education*, 60, 136-45.
- Lantolf, J. P. (2000). *Sociocultural theory and second language learning*. Oxford: Oxford University Press.
- Lejk, M., Wyvill, M. & Farrow, S. (1996). A group interaction peer/self assessment process in a pharmacy practice course. *American Journal of Pharmaceutical Education*, 60, 36-45.
- Lejk, M. & Wyvill, M. (2001). The effect of the inclusion of self-assessment with peer assessment of contributions to a group project: a quantitative study of secret and agreed assessments. *Assessment and Evaluation in Higher Education*, 26, 551-61.
- Luoma, S. (2004). *Assessing speaking*. Cambridge: Cambridge University Press.
- Magin, D. J. (2001). A novel technique for comparing the reliability of multiple peer assessments with that of single teacher assessments of group process work, *Assessment and Evaluation in Higher Education*, 26(2), 139-52.
- McNamara, T. (1997). 'Interaction' in second language performance assessment: whose performance? *Applied Linguistics*, 18(4), 446-66.
- Morrison, D. and Lee, N. (1985). Simulating an academic tutorial: A test validation study. In Y. P. Lee (Ed.) *New directions in language testing* (pp. 85-92). Oxford: Pergamon Institute of English.
- Naughton, D. (2006). Cooperative strategy training and oral interaction: Enhancing small group communication in the language classroom. *The Modern Language Journal*, 90(2), 169-84.
- Nicolay, J. A. (2002). Group assessment in the on-line learning environment. *New directions for Teaching and Learning*, 91, 43-52.
- Nordberg, D. (2008). Group projects: more learning? Less fair? A conundrum in assessing postgraduate business education. *Assessment and Evaluation in Higher Education*, 33(5), 481-92.
- Nunn, R. (2000). Designing rating scales for small-group interaction. *ELT Journal*, 54(2), 169-78.
- Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, 21, 108-48.
- O'Sullivan, B. (2002). Learner acquaintanceship and OPT pair-task performance. *Language Testing*, 19, 277-95.
- Plough, I. and Gass, S.M. (1993). Interlocutor and task familiarity: effects on interactional structure. In G. Crookes & S. M. Gass (Eds.) *Tasks and language learning* (pp. 35-56). Clevedon: Multilingual Matters.
- Porter, D. (1991). Affective factors in language testing. In J. C. Alderson & B. North (Eds.)

- Language testing in the 1990s* (pp. 32-40). Modern English Publications in association with the British Council. London: Macmillan.
- Reves, T. (1991). From testing research to educational policy: a comprehensive test of oral proficiency. In J. C. Alderson & B. North (Eds.) *Language testing in the 1990s* (pp. 178-88). London: Modern English Publications and the British Council.
- Sharp, S. (2006). Deriving individual student marks from a tutor's assessment of group work. *Assessment and Evaluation in Higher Education*, 31(3), 329-43.
- Shehadeh, A. (2001). Self- and other-initiated modified output during task-based interaction. *TESOL Quarterly*, 35(3), 433-48.
- Shohamy, E., Reves, T., & Bejarano, Y. (1986). Introducing a new comprehensive test of oral proficiency. *ELT Journal*, 40(3), 212-20.
- Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.) *Sociocultural theory and second language learning* (pp. 97-114). Oxford: Oxford University Press.
- Swain, M. & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *Modern Language Journal*, 82, 320-38.
- Tarone, E. & Liu, G. Q. (1995). Situational context, variation, and second language acquisition theory. In G. Cook & B. Seidlhofer (Eds.) *Principle and practice in applied linguistics* (pp. 107-24). Oxford: Oxford University Press.
- Weir, C. (1990). *Communicative language testing*. New York: Prentice Hall.

APPENDIX

Tasks for Oral Group Discussion

Type: Non-convergent Discussion (*Opinions Exchange)

Planning Time: 5 minutes

Discussion Time: 3-4 minutes

Scoring Method: Group scoring or Individual scoring

Assessment Prompts:

Task 1: As a group, discussion what you will do in the following scenario:

“ You and a group of people are hiding from some violent criminals. A baby starts to cry. If it is not stopped soon, you will be found and killed. One way to stop the baby's crying is to smother (窒息) it with a towel. Assume it is Not your baby.”

Task 2: As a group, discuss what you will do in the following scenario:

“A fierce storm hit two passenger cruises, which started to sink. Both ships carry the same number of passengers. You are the leader of the coast guards, who came to the rescue. Unfortunately, your crew and you can only rescue one ship at a time. The second ship to be rescued would have for sure sunk to the bottom of the ocean by the time you get to them.”

Holistic Learning of Academic Writing: Impacts of Service Learning on Writing Identity and Knowledge Construction

Yi-chun Liu(劉怡君)
National Cheng Chi University
Liuyc77@nccu.edu.tw

In traditional writing classrooms, students evoke inspiration by reading writing samples, retrieving personal experiences or relying on imagination. Writing, in a sense, becomes an isolated classroom exercise or personal activity constrained within imitation of writing samples or decontextualized memories retrieval. Oftentimes, the focuses of writing are on writing conventions, forms, and mechanical problems at sentence levels.

However, from the perspective of critical academic writing, writing texts should be situated in contexts, and they are shaped by interactions among the writer, the reader and the community. In order to help students develop writing literacy from a holistic base and encourage students not only to make critical arguments about social issues but also to construct their authorial identity in their writings, service learning (S-L) method is incorporated into a writing course curriculum.

S-L, rooted in experientialism, is considered as an effective method to help students learn content related knowledge through participation and hands-on engagement. The real context experiences and service roles enable students to observe issues behind surface as insiders, hence empower them to make their writing arguments more critically with more insights. In particular, the complexity of real-world projects help students develop their rhetorical knowledge beyond linguistic skills, such as constructing their writing identity and voice, being familiar with community discourses, and recognizing writers' negotiation between texts, context and audience. This study attempts to research writers' identity and knowledge construction through service learning.

INTRODUCTION

In traditional EFL writing classrooms, students evoke inspiration by reading writing samples, retrieving personal experiences or relying on imagination. Writing, in a sense, becomes an isolated classroom exercise or personal activity constrained within imitation of writing samples or decontextualized memories retrieval. Writing teachers usually follow textbooks to teach writing skills, but oftentimes, the focuses of writing are on writing conventions, forms, and mechanical problems at sentence levels. However, from the perspective of critical academic writing, writing texts should be socially situated in contexts, and they are shaped by interactions among the writer, the reader and the community (Canagrajah, 2002). In order to help students develop writing literacy from a holistic base and encourage students not only to make critical arguments but also to construct their textual authorial identity and create professional knowledge of their own, service learning (S-L) method is incorporated into a writing course curriculum.

LITERATURE REVIEW

Rooted in experientialism, service learning (S-L) method has a long history in higher education in the U.S.A., which can be dated back to 1960s. Although S-L method has never

been the predominant pedagogy in teaching writing in the U.S.A., it has been recognized as significant in various aspects of holistic learning. Through participating in volunteering services and hands-on engagement, students are reported to be able to acquire content related knowledge (Bringle and Hatcher, 1996; Eyler and Giles, 1999), develop writers' awareness of citizenship (Eyler and Giles, 1999) as well as enhance learning motivation (Eyler and Giles, 1999; Bryant and Hunton, 2000), critical thinking and other meta-cognitive competences (Deans, 2000; Batchelder and Root, 1994). Ellen Cushman (1999) further theorizes S-L as a combination of postmodern ethnography and notions of reciprocity (p.332). Through others' perspectives, a S-L ethnographer defamiliarizes the ordinary in order to find extraordinary within, or "render the familiar strange" (Sullivan, 1996, 99), or turning the strangeness into a profession for accumulating knowledge. In particular, the real context experiences and the complexity of real world projects enable students to observe issues behind surface as insiders, and empower them to make their writing arguments more critically. Therefore, it is believed that S-L helps students develop their writing literacy beyond linguistic skills, such as constructing their writing identity, being familiar with community discourses, recognizing writers' negotiation among texts, context and audience, and create knowledge ownership through experience. However, incorporating S-L method in teaching EFL writing is rather an unconventional practice or even an agitated pedagogy in Taiwan. Under intense competition, Taiwanese students generally are test driven and purpose oriented for learning; however, the interconnectedness of academic writing and community services is not manifest to most teachers and students. This study attempts to report application of S-L method at a national university in Taiwan; moreover, it also explores construction of students' textual identity and development of writers' content knowledge ownership.

METHOD

The writing course is designed based on the notion of "Writing about the Community" suggested by Deans (2000). Student participants (N= 20) are requested to take volunteering community services which can be freely chosen according to their interest. Students can either team up with peers or work individually at least one hour per week, therefore, have total about 18-20 hours in a semester. Besides volunteering services, students are requested to use their service experience as writing sources to complete three writing tasks: narration, comparison and contrast, and argumentation. Students also need to write three reflection journals to record their observation, difficulties and solutions associating with writing or service experience. A text based interview and a survey are conducted to all the students in the end of the semester to inquire students' identity construction process and how students convert their real life experience into disciplinary knowledge. Based on the interview protocols, surveys, students' journals, and writing texts, the researcher attempts to answer the following research questions:

FINDINGS

According to the survey, 70% of the students agreed that their S-L experience influences their textual identity. And 80% of the students agree that experiences acquired from service learning benefit their academic writing. In this presentation, I will analyze one student's case in order to illustrate the two research questions respectively.

Analysis of Case Study

RQ1: Yu-Shin was a freshman student from the school of foreign language. As a typical Taiwanese student who has never gone abroad, Yu-Shin has learned English since junior high school; she went to cram schools and sometimes listened to radio English programs for extra English practices. She admitted in her journal that even though she had no interest in animals, she chose Wild Bird Society (WBS) as her service learning site simply because her first

priority to be an English tutor for dropouts had been opposed by her parents. WBS is a none-profit organization which dedicates to research and preserves wild birds and their habitats. Yu-Shin was responsible for data entry about the survival and proliferation of wild birds at first at WBS, and occasionally she got chances to take care of wounded birds found and brought by people after she had gained some experience through observation.

According to her text-based interview, Yu-Shin explained her inventing process of topic finding and idea generation. Yu-Shin said she was invited to “observe” how the experienced volunteers took care of a wounded owl and later allowed to take care of it under assistance. She said she was afraid of the wild owl at first because it is her “first time to see an owl with [her] own eyes,” but soon she learned the skills for getting alone with it. In her second journal she wrote, “When I saw the wounded little owl, I wonder what has it gone through and where are its families...” She consulted the experienced staff members of WBS about possible reasons that birds get hurt and finally decided to use first person to write the owl’s story according to the senior staff’s suggestion and her imagined storyline for her narrative essay. Yu-Shin said in her interview that it was difficult to use owl’s eyes to view the world because “it could not communicate with me in language.” She tried to imagine the flying feeling if she had wings, and she tried to picture how the world looked like if she had been at the size of a bird. She described the wounded owl from its perspectives, “It is difficult for me to balance myself with a broken left leg, and my endeavor to stand has caused my left feathers fell or broke easily because I can only use my left wing to sustain my body...” “My left claw, the weapon for hunting delicious mouse, was badly injured, but luckily I can still move my the other claw to grab the worms the human feed me...” The intentional textual identity created by Yu-Shin is an owl, but apparently her human identity still inevitably leads the role. The hybrid identity not only intertextually crosses the boundary of human and an owl, but also affectively influences Yu-Shin’s perception toward animals. In her journal, she reflected, “after days of observation of my subject owl, I gain some acquaintance with it...My writing exercise also changed my views of animals, and I started to be concerned about the issues of birds.” That is, on one hand, S-L experience inspires Yu-Shin to test a totally different textual identity; moreover, S-L experience allows her scrutinizly observing the subject, which successfully enables Yu-Shin to immerse in the created role. On the other hand, voicing owl’s experience with an owl’s identity help Yu-Shin establish affection for this little wild bird and develop sense and sensitivity of how to take care of it.

RQ2: From the perspective of ethnography, ethnographers study the “others” by stepping in fieldsites in order to investigate the target issues from the insider’s view; they “step out” from their belonging community in order to adopt the “outsider stance” to examine target issues of their own. Through observation, recording, interpreting and analyzing, ethnographers try to create, construct and deconstruct knowledge by making the unknown become familiar or making the familiar strange (Sustein and Chiseri-Strater, 2002). Service learning site provided Yu-Shin opportunities to build up knowledge. Yu-Shin in her second writing project compared and contrasted sparrows and owls. She first contrasted the physical features of sparrows and owls with details. Then, she contrasted their habitudes and reproduction. For example, “Sparrows’ eyes, located at the two sides, have much more cone cells, which are good at recognizing colors...they can turn their necks 180 degree around to expand their vision...For owls, their eyes were at the front of their heads, making them have stereoscopic vision like human. Since owls have much more rod cells, they can only identify dark and bright...owls cannot turn their eyes around since their column eyes established by rod cells, instead they can turn their head 270 degree around.” Sparrows and owls are both common birds in Taiwan, but after scrutinizing their differences, Yu-Shin originated information which re-scripted the common birds in differential manifestation. Although students in traditional writing classroom may access similar information through reading, and

Yu-Shin also searched and read related information, her content knowledge is obtained through both reading and personal observation. The decontextualized information which read from printed articles is proved by in person hands-on observation and personal discernment; hence, the information can be rooted in as acquired knowledge. Therefore, S-L Students will be empowered with ownership of the knowledge which is constructed through interactive examination of received information from books and perceived information from observation in the real world.

CONCLUSTION

S-L not only affects writers' construction of textual identity but also positively influences knowledge construction. With real world experience, S-L students may have more confidence on making authorial voice; moreover, writing becomes an knowledge construction process rather than a knowledge delivery process.

REFERENCES

- Batchelder, T. H., and Root, S. (1994). Effects of an Undergraduate Program to Integrate Academic Learning and Service: Cognitive, Prosocial cognitive, and Identity outcomes. *Journal of Adolescence*. 17, 341-355.
- Bringle, R. G. and Hatcher, J. A. (1996). "Implementing Service Learning in Higher Education." *Journal of Higher Education*. 67, 2, 221-239.
- Bryant, S. M., and Hunton, J. E. (2000). The Use of Technology in the Delivery of Instruction: Implications for Accounting Educators and Education Researchers. *Issues in Accounting Education*. 15, 1, 129-162.
- Canagrajah, A. S. (2002). *Critical Academic Writing. Michigan Series on Teaching Multilingual Writers*. The University of Michigan Press: Ann Arbor.
- Cushman, E. (1999). Opinion: The Public Intellectual, Service Learning, and Activist Research." *College English* 61(3). 328-336.
- Deans, T. (2000). *Writing Partnership: Service-Learning in Composition*. Urbana, IL: National Council of Teachers of English.
- Eyler, J. and Giles, D. E. (1999). *Where's the Learning in Service-Learning?* San Fransisco: Jossey-Bass Publisher.
- Sullivan, P. A. (1996). *Ethnography and the Problem of the "other."* In. P. Mortensen and G. E. Kirsch. Ethics and Representation in Qualitative Studies of Literacy. 97-144. NCTE: Illinois.
- Sunstein, B. S. & Chiseri-Strater, E. (2002). *FieldWorking: Reading and Writing Research*. 2nd Edition. St. Martin's.

Bridging Test Construct and Beneficial Washback Effects: Revising the GEPT High-Intermediate Reading Test

Tung-mei Ma (馬冬梅), Shu-fan Li (李淑芳)
The Language Training & Testing Center
td@lrtc.ntu.edu.tw

Based on the findings of a project which related Taiwan's General English Proficiency Test (GEPT) to the Common European Framework of Reference for Languages (CEFR), there was recognition of a need to revise the High-intermediate Reading Test of the GEPT (GEPT-HR), whose level corresponds to the CEFR B2 level. The findings indicated that longer texts better facilitate higher levels of cognitive processing. In addition to including longer texts in the GEPT-HR, increasing the variety of text types was also deemed necessary based on the results of a comparison between the GEPT-HR reading specifications and the CEFR B2 level descriptions established by the Council of Europe. An investigation of the lengths and types of texts found in the reading textbooks used in Taiwan's universities and in several major English reading tests supported the need for revision. To investigate the appropriateness of new items, two pretest forms based on the partially revised specifications were pretested. Examinees' performances and their feedback on both the pretest and the operational GEPT test were statistically analyzed. Although the examinees claimed that the long passages were more difficult, their performances did not support that claim. Their performances were satisfactory in terms of difficulty and discrimination indexes. To triangulate the data, university teachers' opinions were collected. Whether the revised test form could exert more beneficial washback effects on the teaching of English reading in Taiwan was also considered when this project was launched.

INTRODUCTION

This study aimed to enhance the construct validity of the GEPT High-Intermediate Reading Comprehension Test (GEPT-HR) through partial revision of its specifications. Another equally important aim was to achieve positive washback effects on the teaching of English reading in Taiwan. The results of a previous questionnaire survey indicated that this should be a goal (Wu, 2008:8). Fully 90% of high school and university teachers who took the survey agreed that external exams such as the GEPT influenced their teaching and assessment practices. This study first experimented with testing longer passages of approximately 350 and 400 words in length, passage lengths which would better match the can-do descriptors at the CEFR B2 level established by the Council of Europe. The testing time of 50 minutes and word-count of approximately 3,000 words remained the same as in the operational GEPT-HR. Second, to include a broader selection of text types, one argumentative text was used in the reading comprehension part. The results of the pretest are presented in this paper.

Based on Bachman's Assessment Use Argument for language design and development (2005:1-34), it was decided that partial revisions of text types and text length would be made; two test forms with parallel difficulty were then compiled based on the revised specifications and pretested so that examinees' performance results could be statistically analyzed in order to assess the feasibility of these revisions. Both test forms were pretested on 1,201 examinees randomly selected from among examinees scheduled to take an operational GEPT-HR. After taking the pretest and the operational GEPT-HR, examinees were asked to provide feedback by completing questionnaires. Examinees' opinions on the difficulty level of the two long passages of approximately 350 and 400 words in each test were analyzed. The effect of the

degree of difficulty was examined quantitatively by analyzing examinees' performances. In addition, further interviews with six university English teachers regarding the revisions to the current operational GEPT-HR were conducted.

RELATING GEPT TO CEFR

In 2005, the Ministry of Education (MOE) adopted the Common European Framework of Reference for Languages (CEFR) as a new standard for assessment in Taiwan. This framework, which "provides a common basis for the elaboration of language syllabuses, curriculum guidelines, examinations, and textbooks across Europe," describes foreign language proficiency at the A1, A2, B1, B2, C1, and C2 levels. Each level has descriptors in the form of can-do statements relating to five language skill areas: listening, reading, spoken interaction, spoken production, and writing (Council of Europe, 2001 & Little, 2005:321-336). In 2007, the LTTC undertook a calibration project to relate the GEPT to the CEFR (Wu & Wu, 2007). The results showed that the difficulty level of the GEPT High-Intermediate Level (GEPT-HI) is approximately situated at the CEFR B2 level. The general level descriptions of both the GEPT HI level and the CEFR B2 level are presented in Table 1. Table 2 shows the descriptions of overall reading comprehension of both levels.

Table 1. *The general level descriptions of the GEPT-HI level and CEFR B2 level*

<p>GEPT High-Intermediate Level</p> <p>An examinee who passes this level has a generally effective command of English; he/she is able to handle a broader range of topics, and although he/she makes mistakes, these do not significantly hinder his/her ability to communicate. His/her English ability is roughly equivalent to that of a university graduate in Taiwan whose major was not English.</p>
<p>Independent User: B2 (2001:24)</p> <p>Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation.</p> <p>Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party.</p> <p>Can produce clear, detailed text on a wide range of subjects and explain viewpoints on a topical issue, giving the advantages and disadvantages of various options.</p>

Table 2. *The descriptions of overall reading comprehension of the GEPT-HI level and CEFR B2 level*

<p>GEPT High-Intermediate: Reading Skill</p> <p>An examinee who passes this level can read written messages, instruction manuals, newspapers, and magazines. At work, he/she can read general documents, abstracts, meeting minutes, and reports.</p>
<p>Independent User: B2 Overall Reading Comprehension (2001:69)</p> <p>Can read with a large degree of independence, adapting style and speed of reading to different texts and purposes, and using appropriate reference sources selectively. Has a broad active reading vocabulary, but may experience some difficulty with low frequency idioms.</p>

According to the CEFR descriptor of reading for orientation at the B2 level, readers can "scan quickly through long and complex texts, locating relevant details" (2001:70). They are independent readers who can comprehend the main ideas of complex texts on both concrete and abstract subjects related to their field and interests (2001:24). Likewise, the

reading skill description of the GEPT High-Intermediate level indicates that successful examinees at this level achieve a satisfactory level of comprehension while reading articles of various text types and genres both in daily life and at work. The above indicates that the length and types of texts are primary considerations in the development of reading tests at high-intermediate level.

TEXT LENGTH OF READING PASSAGES

A study on the reading subtest of the new TOEFL by Cohen and Upton, which incorporated fewer but longer (600-700 vs. 300-400 words) texts than those used in previous TOEFL test designs, supported the use of longer texts for testing academic reading skills at the university level. "Longer texts better represent the academic experiences of students and they better facilitate the development of Reading to Learn purposes in the test design" (2007: 213). Their study indicated that reading to learn and inference items "call for the academic skills of identifying logical connectors and other markers of cohesion and determining how sections of passages interrelate in an effort to establish passage coherence" (2007:214). These reading skills are essential for academic readers at the university level to gain both a local and general understanding of the text passages.

According to Bachman and Palmer, test designers must consider test qualities including reliability, validity, authenticity, impact, and practicality when designing a language test, (1996:38). From the viewpoint of authenticity, the practice of using longer texts for academic purposes "reflects more closely the situation where students have to read and study long texts in college" (2000:109). Therefore, it is necessary for the GEPT-HR to include longer passages in reading comprehension tests.

REVISION OF GEPT-HR TEST

The GEPT-HR consists of three parts: Part I: Vocabulary and Structure, Part II: Cloze, and Part III: Reading Comprehension. The total number of test items is 50, including 15 items for Part I, 15 items for Part II, and 20 items for Part III, the largest portion of the test. All have equal value in scoring.

Reduction of the Number of Vocabulary and Grammatical Structure Items

In order to keep the same total word-count of approximately 3,000 words as well as the same overall degree of difficulty, and based on the consideration that the design of a reading test for this level should mainly focus on testing examinees' discourse-processing abilities, it was decided that five items in Part I would be eliminated, leaving five vocabulary items and five structure items. The total number of questions in Parts II and III remained unchanged. Thus, the total number of items for the entire test was reduced to 45. The operational test and the pretest are compared in Table 3.

Table 3. A comparison between the operational test and the pretest

	Operational Test (Current Test Format)	Pretest (Future Test Format)	Revisions
Part I: Vocabulary & Structure	15 items: 7 vocabulary+ 8 structure	10 items: 5 vocabulary+ 5 structure	1. Reducing 2 items on vocabulary 2. Reducing 3 items on grammatical structure
Part II: Cloze	2 items: 7+8 deletions	2 items: 7+8 deletions	Unchanged
Part III: Reading Comprehension	6 passages: 20 questions (The longest one: approx. 250 words)	5 passages: 20 questions	1. Using long passages of 350 words and 400 words 2. Adding one argumentative text
Total number of items	50 items	45 items	
Testing time	50 minutes	50 minutes	
Total number of words	Approx. 3,000	Approx. 3,000	

Note. The shadowed boxes indicate a reduction of the number of test items in the pretest and revisions of the test format.

Using Longer Texts and Adding the Argumentative Text Type

A comparison of the total number of words in individual passages from different reading comprehension tests including FCE, TOEFL, IELTS, etc. and from the GEPT-HR demonstrated that most of these passages contained at least 300 words, approximately 50 words more than the longest passages in the operational GEPT-HR. The results of the comparison are presented in Table 4.

Table 4. *A comparison of English proficiency tests of similar levels*

	GEPT-H R (Taiwan)	CEAE English Test (Taiwan)	JCEE English Test (Taiwan)	CET-4 (China)	FCE (UK)	TOEFL (U.S.)	IELTS (UK)	EIKEN ¹ (Japan)
Testing time	50 min.	100 min.	80 min.	55 min.	75 min.	60~100 min.	60 min.	90 min.
Tasks	Voc. & Structure, Cloze, R.C.	Voc., Cloze, R.C., Trans., Comp.	Voc., Cloze, R.C., Trans., Comp.	Skimming & Scanning, Reading in depth, Cloze	R.C.	R.C.	R.C.	Voc., Cloze, R.C., Comp.
Total no. of items	50	56	56	30	35	n/a	40	41
Total no. of words in individual passages	130~250	200~260	150~320	300~350	350~700	Approx 700	450~800	300~550
Total no. of items for each passage	3~5	3~4	3~4	5	7~8	12~14	9~16	3~4

Note: Voc.–Vocabulary / R.C.–Reading Comprehension / Trans.–Translation / Comp.–Composition

The findings supported the need to include long passages of 350 and 400 words in addition to the original passages of 250 words. The maximum number of words in the longest passage in the pretest was approximately 400 words.

To incorporate fewer but longer reading passages in the GEPT-HR pretest than those used in the current operational GEPT-HR, the number of passages in Part III was reduced from six to five. A comparison of the text types and the number of items in the reading comprehension part of the operational test and the pretest is presented in Table 5.

¹ CEAE: College Entrance Ability Exam, also known as Basic Proficiency Test or Scholastic Ability Test

JCEE: Joint College English Exam

CET-4/6: College English Test, Level 4 and Level 6 (CET is a national English as a Foreign Language in China)

FCE: Cambridge Main Suite Exam: First Certificate in English

IELTS: the International English Language Testing System

EIKEN: Test in Practical English Proficiency administered by STEP (Society for Testing English Proficiency)

Table 5. *A comparison between the text types and the number of items in the reading comprehension part of the operational test and the pretest*

Operational test		Pretest	
Six Passages		Five Passages	
Chart or Graph	2 Qs	Chart or Graph	2 Qs
Business letter	3 Qs	Business letter	3 Qs
News Report	3 Qs		
Expository Essay Approx. 180~200 words	3 Qs	News Report / Expository Essay Approx. 200~250 words	4 Qs
Expository/Narrative Essay Approx. 200~250 words	4 Qs	Argumentative Essay Approx. 300~350 words	5 Qs
Expository Essay Approx. 200~250 words	5 Qs	Expository/Narrative Essay Approx. 350~400 words	6 Qs

Note: The shadowed boxes indicate changes of text types and the number of items in the pretest.

Adding the Argumentative Text Type

According to the CEFR descriptor of reading for information and argument at the B2 level, "readers can understand articles and reports concerned with contemporary problems in which the writers adopt particular stances or viewpoints" (Council of Europe, 2001:70). In addition, the new TOEFL test, designed to test academic reading comprehension at the university level, includes a broader selection of text types such as "exposition, argumentation, and historical biographical/autobiographical narrative" (Cohen & Upton, 2007:213). Thus, the argumentative text type was also used in the pretest for the long passage of 350 words.

METHODOLOGY

Administration of the Pretest

Two pretest forms of parallel difficulty, A and B, were compiled, and the pretest was conducted in October 2008.

Participants

This study involved a sample of examinees of the GEPT-HR who had registered for the operational GEPT-HR to be given two weeks after the pretest. These examinees represented the GEPT operational test taking population since the structure and the distribution of the pretest were conditioned to resemble those of the operational test. A total of 1,201 participants from the northern, central, and southern parts of Taiwan took part in the study. Two groups of 588 and 613 participants, determined by random sampling, took tests A and B, respectively. Their ages ranged from 17 to 22 years with an average of 19, which is two years younger than the average age of GEPT High-Intermediate target examinees. Of the 1,201 pretest participants, approximately 90% were senior high school students (522) and college students (552), while the rest were junior college or technology university students. Nearly 50% of the participants took the test in Taipei, 30% in Taichung, and the rest in Tainan and Kaohsiung.

Materials and Design

Instrument The materials for this study consisted of two parallel high-intermediate reading comprehension tests, A and B, each containing 45 test items, and two post-test questionnaires for the participants. Both tests lasted 50 minutes, the same testing time as the operational GEPT-HR. Following the pretest and the operational test, participants were given questionnaires to complete. Test A and Test B were each anchored by 14 common items

spread among the three parts of the test. As shown in Table 6, Parts I, II, and III included three, seven, and four anchor items, respectively. The anchor items were from the GEPT item bank and possessed satisfactory statistical attributes. The anchor-item design was employed to facilitate the calibration of all test items during data analysis. This helped to equate item difficulty in the two parallel test forms. The distribution of anchor items is presented in Table 6.

Table 6. *The distribution of anchor items in three parts of the pretest*

	Total no. of items	Total no. of anchor items
Part I: Vocabulary & Structure	10	1 Vocabulary+ 2 Structure items: 3
Part II: Cloze:	Two passages: 15 deletions	One anchor passage: 7 deletions
Part III: Reading Comprehension:	Five passages: 20 questions	One anchor expository passage: 4 questions

Post-test questionnaires The questionnaire for the pretest contained 18 questions regarding the reduction of items in Part I and the employment of long passages of over 300 words in Part III. The subsequent questionnaire for the operational test contained 12 questions relevant to the latter test. The purpose of both questionnaires was to elicit participants' opinions about the number of test items and whether or not the test could effectively measure the participants' level of proficiency. Following the operational test, the results of both questionnaires were compared and analyzed statistically in order to assess the feasibility of the new test form.

Investigation of freshman English textbooks used at universities To ensure consistency between the construct of the GEPT-HR and the learning materials used by the target test-takers in college, the contents of university freshman English reading textbooks, such as *Reading Matters* (Wholey, 1999), *Mosaic: A Reading Skills Book* (Wegmann & Knezevic, 2007), *Reading for the Real World* (Craber & Babcock, 2004), *Active: Skills for Reading* (Anderson, 2002), and *Rethinking America* (Sololik, 1999) were reviewed. The textbooks were found to contain a wide range of topics, such as culture and society, business, sports, entertainment, and technology. The majority of the reading passages in the textbooks contain approximately 400-600 words. These findings further supported the revisions of the operational GEPT-HR test.

Interviews with university English teachers Following the pretest in 2008, six university teachers, including lecturers and professors who were experienced in teaching reading, were interviewed to obtain their opinions about the long passages and the argumentative text type. They were asked whether they considered the revisions appropriate. Their opinions are summarized as follows:

With regard to the length of the reading passages, all of the teachers agreed that their students should study long texts at this level. In fact, many of the articles they read in class contained more than 400 words.

With regard to the text types, all of the teachers supported the inclusion of the argumentative texts, which can help students learn how to distinguish the internal logic of a text and the author's stance. In fact, in reading classes, students are often instructed to read such texts as they are considered more challenging and good practice for learning how to debate controversial issues.

In brief, the six teachers interviewed expressed positive support for the revisions of the text length and text type. They believed that such revisions would have a positive effect on students' learning, as students would be more motivated to read extensively.

RESULTS & DISCUSSION

Performance data

SPSS 15.0 for Windows and ITEMAN 3.5 were used for data and item analysis.

Examinees' performances on the pretest forms Two parallel pretest forms were compiled based on the revised test specifications with 14 common items anchored in both tests. Table 7 shows how examinees performed on anchor and non-anchor items in Test A and Test B. The two groups of examinees performed equally well on the anchor items, and the mean P of Test A and Test B was 0.66 and 0.65, respectively, which indicates that the two groups of examinees had equivalent abilities. Moreover, the mean P of the anchor items was consistent with the average mean P (0.65) of the operational GEPT-HR forms used in the past, which indicates that the abilities of these examinees were also an accurate representation of the abilities of GEPT High-Intermediate test takers.

The difficulty level of the non-anchor items of Test A and Test B varied slightly. The mean P of Test A non-anchor items, 0.62, was lower than that of Test B non-anchor items, 0.64. Therefore, Test B had a satisfactory difficulty level, and Test A, though slightly more difficult than Test B, had a difficulty level that was still within an appropriate range.

Both Test A and Test B had satisfactory reliability; their Alpha values were 0.84 and 0.83, respectively. The mean biserial of Test A and Test B was 0.48 and 0.46, respectively, which indicates that both tests provided good discrimination power.

Table 7. *Descriptive statistics of anchor and non-anchor items of Test A and Test B*

	Test A			Test B		
	Anchor	Non-anchor	Total	Anchor	Non-anchor	Total
No. of Items	14	31	45	14	31	45
No. of Examinees	588	588	588	613	613	613
SD	2.58	5.09	7.07	2.55	5.03	6.97
Alpha	0.62	0.78	.84	0.61	0.767	.83
Mean P	0.66	0.62	.63	0.65	0.64	.65
Mean Biserial	0.56	0.49	.48	0.55	0.48	.46

Comparison of examinees' performances on the pretest forms and the GEPT-HR

The pretest examinees took the operational test, GEPT-HR, two weeks later, and their performances are compared in Table 8.

As mentioned earlier, the examinees of Test A and Test B had equivalent abilities, and their performances in the GEPT-HR therefore remained the same, with a mean P of 0.65, and very similar mean scores of 78.53 and 78.24, respectively. Comparing the pretests forms with the GEPT-HR based on the performances of the same population, the results show that the pretest forms were equivalent to the GEPT-HR in terms of their difficulty level and reliability.

Table 8. *Comparison of examinees' performances on the pretest forms and the GEPT-HR*

No. of Examinees	588		613	
	Test A	GEPT-HR	Test B	GEPT-HR
Mean P	0.63	0.65	0.65	0.65
Mean Score	75.73	78.53	77.61	78.24
SD	18.88	18.73	18.58	18.05
Alpha	0.84	0.85	0.83	0.85
Mean Biserial	0.476	0.481	0.458	0.463

Item analysis of the long passages in Part III of the pretest In Part III of Test A and Test B, there were twenty items, eleven of which (number 35 to number 45) corresponded to the two long passages in each test. The item analysis in Table 9 shows that the argumentative text was comparatively more difficult than the other text type, narrative/expository (Mean Dif.: Test A 0.46/ 0.69, Test B 0.58/ 0.67). The examinees may not have been familiar with the argumentative text type, which presents different, usually opposing arguments, about an issue. In addition, two items for the argumentative text in Test A were too difficult. This type of passage and its test items will be revised to an appropriate difficulty level in the future, based on these results. The expository/narrative texts in both pretest forms, which were about 400 words in length, were not as difficult as expected. The examinees performed quite well on the longest passage in each test, but in the post-test questionnaire, 40% thought the passages and items were difficult. These long passages also provided good discrimination power, with mean biserial values of 0.43 and 0.4, respectively.

Table 9. *Item analysis of the long passages in Part III of the pretest*

Item No.	Text Type	Test A		Test B	
		Dif.	Dis.	Dif.	Dis.
35	Argumentative	0.45	0.34	0.48	0.30
36		0.38	0.39	0.80	0.45
37		0.78	0.44	0.73	0.29
38		0.43	0.36	0.26	0.31
39		0.27	0.31	0.65	0.41
Mean		0.46	0.37	0.58	0.35
40	Narrative/ expository	0.75	0.44	0.77	0.36
41		0.74	0.42	0.61	0.41
42		0.54	0.45	0.66	0.48
43		0.72	0.40	0.78	0.44
44		0.63	0.44	0.56	0.37
45		0.78	0.41	0.61	0.34
Mean		0.69	0.43	0.67	0.4

Questionnaires

Following the pretest and the operational test, GEPT-HR, two questionnaires were administered to collect examinees opinions regarding the number of items in each part, the text length, and the text difficulty in the reading tests. A total of 1,201 examinees completed the pretest questionnaire, and a total of 4,186 examinees completed the GEPT-HR questionnaire. According to the results, 95% of the examinees of both tests agreed that the test they took could effectively measure their reading ability.

Questionnaire Results of anchor examinees There were 470 examinees who took both the pretest and the operational test, GEPT-HR. Their responses to the two questionnaires were further analyzed. The results are presented in the appendix and summarized as follows.

Number of items in each part In terms of the number of items in each part of the reading test, more than 80% of the examinees thought the number of items in Part I (vocabulary & structure) and Part II (cloze) of both the pretest and the GEPT-HR were appropriate. However, with regard to the 20 items in Part III, only 43% agreed that the number of items in Part III (reading comprehension) of the pretest was appropriate, and 54% thought the number was a bit high. Approximately 56% thought the number of items in the

GEPT-HR was appropriate, while 40% thought it was a little high.

The difficulty level of each part Regarding the difficulty level of each part of the tests, 50% to 55% thought Part I of both the GEPT-HR and the pretest was of moderate difficulty. As for the difficulty level of Part II, 54% agreed that the GEPT-HR was appropriate, and 71% thought the pretest was appropriate. The item analysis also showed that Part II of the GEPT-HR was more difficult than Part II of the pretest. When it came to the difficulty level of Part III, 60% agreed that that of the GEPT-HR was appropriate, but only 48% agreed that the difficulty level of the pretest was appropriate. The score data supported the results of the questionnaires as well, showing that the difficulty level of Part III of the pretest was higher than that of the GEPT-HR.

The length and difficulty level of the two long passages There were two longer passages in both Test A and Test B. The argumentative text was about 350 words in length, and the expository/narrative text was about 400 words in length. Approximately 49% of the examinees thought the former text was of appropriate length, while 50% thought it was slightly long. Only 33% of the examinees thought the latter text was of appropriate length, while 66% thought it was slightly long. Over 50% of the examinees agreed that the difficulty level of the argumentative text and its corresponding questions was appropriate. As for the expository/narrative text, approximately 55% agreed that the difficulty level was suitable, and 40% agreed that the text and items were difficult. However, the score data showed that the long texts were actually not difficult. As mentioned earlier, the examinees performed quite well on these long passages. In this case, the examinees underestimated their ability and scored higher than they realized.

CONCLUSION

With the aim of bringing about beneficial washback effects on English teaching and learning, and improving test quality, a decision was made to revise the GEPT High-Intermediate level reading test by using longer texts and a greater range of text types. Based on CEFR B2 level descriptors, this study incorporated input on text length and type from different perspectives, including teachers, examinees, textbooks, and other English language tests. The pretest results of the revised test forms showed that the reliability of the pretest forms was comparable to that of the current operational GEPT. The pretest items based on the longer passages were found to have satisfactory difficulty and discrimination indexes. Although the examinees claimed that the long passages were more difficult, their performances did not support that conclusion. The revision is expected to make the construct of the reading test more meaningful and generalizable and thus reflect real-life language use so as to accurately measure the reading ability of High-Intermediate test takers in Taiwan. It is believed that this will also benefit the teaching and learning of English reading in Taiwan.

ACKNOWLEDGEMENTS

We thank the following teachers for their participation in the interview: Yiu Man Ma & Teh-ming Sung of National Taiwan University, Jing-Sha Chen & Chang-Chun Li of National Taiwan Normal University, Chung-Shin Liang of National Pingtung University of Education, and Wen-Hsien Yang of National Kaohsiung Hospitality College. We are also grateful to Cecilia Liao for her statistical assistance, and to Rachel Wu and Jessica Wu for their constructive comments and valuable suggestions.

REFERENCES

Alderson, J. C. and Bachman, L. F. (2000). *Assessing reading*. Cambridge: Cambridge University Press.

- Anderson N. (2002). *Active skills for reading: book 3*. Heinle & Heinle.
- Bachman, L. F. (2005). Building and supporting a case for test use. *Language Assessment Quarterly*, 2(1), 1-34.
- Bachman, L. F. and Palmer, A. F. (1996). *Language testing in practice*. Oxford University Press.
- Cohen A. D. and Upton T. A. (2007) "I want to go back to the text": Response strategies on the reading subtest of the new TOEFL. *Language Testing*, 24 (2), 209-250.
- College Examination Committee of China (1999). *What does CET test: analysis of students' performance on CET papers band 6*. Shanghai Foreign Language Education Publisher.
- College Examination Committee of China (1999). *What does CET test: analysis of students' performance on CET papers band 4*. Shanghai Foreign Language Education Publisher.
- Council of Europe (2001). *Common European framework of reference for languages: learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Craber, B. & Babcock, P. (2004). *Reading for the real world*. Compass Publishing Inc.
- FCE handbook* (2003). University of Cambridge: ESOL Examinations.
- Little, D. (2005). The common European framework and the European language portfolio: involving learners and their judgments in the assessment process. *Language Testing*, 22 (3), 321-336.
- Sololik, M. E. (1999). *Rethinking America: a high intermediate cultural reader*. Heinle & Heinle Publishers.
- Wegmann, B. & Knezevic, M. P. (2007). *A reading skills book: mosaic one*. The McGraw-Hill Companies, Inc.
- Wholey, M. J. (1999). *Reading matters*. Houghton Mifflin Company.
- Wu, J. (2008). Views of Taiwanese students and teachers on English language testing. *Cambridge ESOL: Research Notes: Issue 34, Nov. 2008*, 6-9
- Wu, J. & Wu, R. (2007). *The LTTC GEPT-CEFR calibration project*. The Language Training and Testing Center.

APPENDIX

Results of the post-test questionnaires by anchor examinees (N=470)

3. 本次閱讀能力測驗第一部份(詞彙與結構)的題數(15 題)對我而言

GEPT-HR#03					PretestQ3				
	稍少	適中	稍多	(空白)		稍少	適中	稍多	(空白)
pass	3%	85%	12%	0%	pass	8%	85%	8%	0%
fail	3%	83%	12%	2%	fail	10%	83%	6%	0%
總計	3%	84%	12%	1%	總計	9%	84%	7%	0%

4. 本次閱讀能力測驗第二部份(段落填空)的題數(15 題)對我而言

GEPT-HR#04					PretestQ4				
	稍少	適中	稍多	(空白)		稍少	適中	稍多	(空白)
pass	2%	90%	8%	0%	pass	6%	87%	7%	0%
fail	2%	80%	18%	0%	fail	6%	84%	11%	0%
總計	2%	86%	12%	0%	總計	6%	86%	9%	0%

5. 本次閱讀能力測驗第三部份(閱讀理解)的題數(20 題)對我而言

GEPT-HR#05					PretestQ5				
	稍少	適中	稍多	(空白)		稍少	適中	稍多	(空白)
pass	2%	64%	33%	1%	pass	2%	47%	50%	2%
fail	3%	47%	49%	1%	fail	0%	39%	58%	3%
總計	2%	56%	40%	1%	總計	1%	43%	54%	2%

6. 本次閱讀能力測驗第一部份(詞彙與結構)對我而言

GEPT-HR#06					PretestQ6				
	容易	適中	稍難	(空白)		容易	適中	稍難	(空白)
pass	3%	45%	50%	2%	pass	10%	56%	33%	0%
fail	1%	54%	43%	2%	fail	6%	53%	41%	0%
總計	2%	50%	47%	2%	總計	8%	55%	37%	0%

7. 本次閱讀能力測驗第二部份(段落填空)對我而言

GEPT-HR #07					Pretest Q7				
	容易	適中	稍難	(空白)		容易	適中	稍難	
pass	2%	58%	39%	2%	pass	9%	77%	14%	
fail	0%	49%	49%	1%	fail	4%	65%	31%	
總計	1%	54%	44%	1%	總計	7%	71%	22%	

8. 本次閱讀能力測驗第三部份(閱讀理解)對我而言

GEPT-HR#08					Pretest Q8				
	容易	適中	稍難	(空白)		正式 PF 容易	適中	稍難	(空白)
pass	9%	69%	21%	1%	pass	3%	55%	42%	0%
fail	3%	50%	45%	3%	fail	1%	41%	57%	1%
總計	6%	60%	32%	2%	總計	2%	48%	49%	1%

9. 本次閱讀能力測驗第三部份(閱讀理解)的文章長度對我而言

GEPT-HR					Pretest				
	稍短	適中	稍長	(空白)		稍短	適中	稍長	(空白)
pass	3%	89%	8%	1%	pass	7%	77%	16%	0%
fail	2%	71%	25%	1%	fail	1%	67%	31%	0%
總計	3%	81%	16%	1%	總計	4%	73%	23%	0%

12. 本次閱讀能力測驗第三部份(閱讀理解)第 35 題至第 39 題的文章長度對我而言

15. 本次閱讀能力測驗第三部份(閱讀理解)第 40 題至第 45 題的文章長度對我而言

PretestQ12

	稍短	適中	稍長	(空白)
pass	0%	51%	49%	0%
fail	0%	47%	53%	0%
總計	0%	49%	50%	0%

PretestQ15

	稍短	適中	稍長	(空白)
pass	1%	40%	59%	0%
fail	1%	25%	73%	0%
總計	1%	33%	66%	0%

13. 本次閱讀能力測驗第三部份(閱讀理解)第 35 題至第 39 題的內容對我而言

Test A Q13

	容易	適中	稍難	(空白)
pass	4%	58%	38%	0%
fail	1%	49%	48%	2%
總計	3%	54%	43%	1%

Test B Q13

	容易	適中	稍難
pass	2%	73%	25%
fail	3%	53%	44%
總計	3%	64%	33%

14. 本次閱讀能力測驗第三部份(閱讀理解)第 35 題至第 39 題的問題對我而言

Test A Q14

	容易	適中	稍難	(空白)
pass	3%	51%	46%	0%
fail	0%	59%	39%	2%
總計	2%	55%	43%	1%

Test B Q14

	容易	適中	稍難
pass	3%	71%	26%
fail	2%	50%	48%
總計	3%	62%	36%

16. 本次閱讀能力測驗第三部份(閱讀理解)第 40 題至第 45 題的文章內容對我而言

Test A Q16

	容易	適中	稍難	(空白)
pass	7%	63%	29%	1%
fail	3%	45%	52%	1%
總計	5%	54%	40%	1%

Test B Q16

	容易	適中	稍難
pass	4%	66%	30%
fail	2%	47%	51%
總計	3%	57%	40%

17. 本次閱讀能力測驗第三部份(閱讀理解)第 40 題至第 45 題的問題對我而言

Test A Q17

	容易	適中	稍難	(空白)
pass	7%	60%	33%	0%
fail	2%	51%	46%	1%
總計	4%	56%	40%	0%

Test B Q17

	容易	適中	稍難
pass	5%	63%	32%
fail	2%	48%	50%
總計	3%	56%	40%

12. 我認為本次閱讀能力測驗是否可以考出我實際的閱讀能力

GEPT-HR#12

	勉強可以	可以	完全可以	(空白)
pass	21%	76%	2%	1%
fail	27%	67%	5%	1%
總計	24%	72%	3%	1%

PretestQ18

	勉強可以	可以	完全可以	(空白)
pass	26%	72%	1%	0%
fail	31%	62%	6%	1%
總計	29%	67%	4%	1%

Peer Review Training and Teacher Feedback

Hui-tzu Min (閔慧慈)

National Cheng Kung University
minhuitz@mail.ncku.edu.tw

This study examined if an EFL writing instructor practiced what she preached in providing written commentary on students' compositions. The instructor trained her students how to generate "revision-oriented" commentary by following a four-step procedure: clarifying writers' intentions, identifying potential issues, expounding the nature of problems, and making text-specific suggestions. In order to examine whether the instructor/researcher also followed these steps while providing feedback, two independent raters and the researcher conducted a textual analysis of the form of her 725 written commentary on 18 students' 2nd drafts prior to and post peer review training. The analysis revealed a significant difference in clarifying writers' intentions and identifying problems respectively. The instructor/researcher probed the writers' intentions significantly more frequently after the training. She also identified significantly fewer problems after the training. The instructor/researcher concludes that the peer review training prompted her to consciously monitor her own commenting skills and to consistently adhere to the 4-step feedback procedure.

INTRODUCTION

A wide array of factors may contribute to the success of peer review training in EFL writing class, one of which is whether "what the teacher says in class is consistent with what she or he models in written or oral feedback" (Ferris, 2003, p. 34). Inspired by Liu and Hansen's suggestion (2002, p. 128) that writing instructors should use the same guidance sheet and follow the same procedure to provide teacher commentary, the instructor/researcher intended to examine her own written commentary on students' drafts and revisions before and after an elaborate peer review training. This self-examination is important given that the peer review training reflects the teacher's belief in how written commentary should be rendered. But whether such beliefs are actually translated into classroom performance is an issue worth further examination. Research has reported mismatches between teacher perceptions and performance in adult ESL writing classrooms (Montgomery & Baker, 2007) and secondary EFL writing classes (Lee, 2009). More research is needed to probe this issue.

In response to Goldstein (2004) for ESL/EFL writing teachers' critical examination of their own practices and to see if the current teacher/researcher's belief in the procedure of providing written commentary matches her own practice, this study aims to examine her written commentary, both before and after a lengthy peer review training session. There are two questions addressed in this study: 1) Did the teacher's written commentary follow the same 4 step-procedure she designed before the peer review training? 2) Did the teacher's written commentary follow the same 4 step-procedure she designed after the peer review training?

METHOD

Participant

The study reported in this paper was part of a larger study (Min, 2005; 2006; 2008) investigating the effects of peer review training in one EFL writing class in an urban university in southern Taiwan. The main participants in this study are 18 sophomores in the instructor/researcher's writing class. There were 16 females and 2 males and their age average was 19. All were native speakers of Mandarin Chinese and had passed the Intermediate Level English Test of the General English Proficiency Test (GEPT)

administered by the Language Training & Testing Center in Taiwan before being admitted to the Department of Foreign Languages and Literature at the university. Their English proficiency was approximately between 523-550 on the Test of English as a Foreign Language (TOEFL) exam. The instructor/researcher has taught EFL writing at this university for 6 years prior to the study.

The Writing Class

The instructor adopted a modified “writing cycle” (Tsui & Ng, 2000) in designing her writing class (Figure 1).

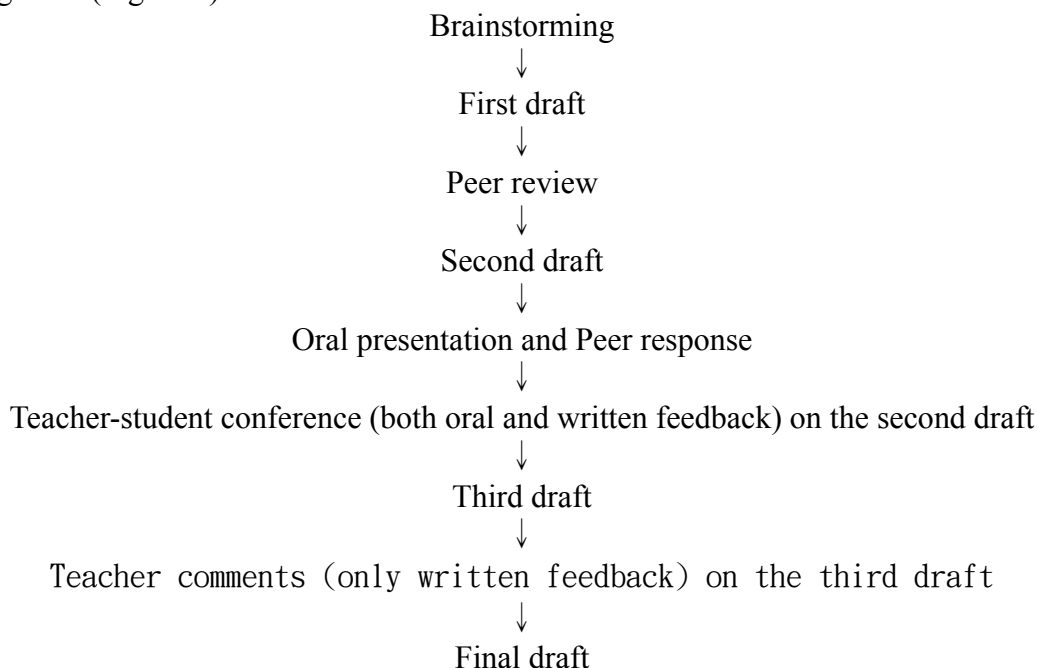


Figure 1 The Writing Cycle

Teacher Belief in Written Feedback Provision

Believing in the effectiveness of specific and relevant feedback (Ferris, 2003; Goldstein, 2004), the instructor/researcher designed a four-step procedure to train her students to make revision-oriented commentary: Clarifying writers’ intentions, identifying problems, explaining the nature of problems and making specific suggestions (see Min, 2005 for details of peer review training). Student reviewers were encouraged to clarify and solicit writers’ intentions first, point to and explain to writers the perceived problems or confusing parts that fail to achieve the intended meaning, and collaborate with them to work on different word, phrase, or idea choices. The combination of four different stances reflects the instructor/researcher’s view of writing as both a process and a product. Confusing ideas are questioned so that more refined ideas or new ones would later emerge. These ideas would then be finalized in accurate forms and acceptable format through collaboration. The step-by-step procedure also reflects the instructor/researcher’s perception of how to offer written feedback.

Data Collection

The instructor/researcher collected the 2nd drafts of students’ 1st (“The advantages/disadvantages of _____ [a new technical invention]”) and 4th (“An Analysis of an Incident of Cross-Cultural Misunderstanding”) essays for analysis. Although these two topics appear to be tapping students’ different knowledge bases, they are comparable in that both were expository essays, requiring students to use their analytical ability to either explain the advantages/disadvantages of a certain technological invention or the sources of cultural misunderstandings as well as to synthesize their sources of information in a coherent manner.

The average number of words produced for the 1st and 4th topics were 263 and 257 words respectively (approximately 1 page).

Data Analysis and Coding

The data analysis is principally a text analysis. Two trained independent raters, both of whom have taught EFL writing for more than 8 years carefully read the instructor/researcher's comments in response to the 2nd drafts of the 1st essay (prior to peer review training) and those to the 2nd drafts of the 4th essay (after peer review training). They first read each sentence in each comment carefully and tried to identify the corresponding constituent step. The unit of analysis for each constituent "step" is a sentence, bound by an initial capital letter and an ending period.

Each sentence was coded according to the four steps in the training procedure: Clarifying writers' intentions [clarification], identifying problems [problem], explaining the nature of the problems [explanation], and making specific suggestions [suggestion]. Examples of applying this coding scheme to the instructor/researcher's comments are presented below.

Comment 1: This sentence doesn't connect well with the previous description.

[problem] You've already described what her house was like and the people she hired to help her out. Then you went back to talk about her moving into the house, which seems a little bit "circular". [explanation] You might want to change it to "after she hired them" instead. [suggestion]

Comment 2: What do you mean by "learning the current situations from her personal investigations"? [clarification]. Do you mean that she "disseminated the first-hand information gleaned from her personal investigations"? [clarification]

The first comment contained three steps—problem identification and explanation of the nature of the problem as well as providing specific suggestions. In the first comment, the instructor/researcher sensed a logical "flow" problem in the sentence and tried to explain the perceived issue to the writer. Then she continued to offer suggestion to remedy this problem. Both raters agreed that problem identification, explanation and making specific suggestion are the main foci of this comment. Both considered the stance of this comment a mixture of "prescriptive", "tutoring", and "collaborative" stances. Comment 2, on the other hand, comprised two questions, with the first one being a genuine question and the second one more like a helping question. The two raters coded the first sentence as "clarification" and the stance "probing" but rated the second sentence as a combination of clarification and suggestion given the instructor/researcher's effort to compose a specific sentence to probe the writer's intentions further.

Interrater Reliability

The two independent raters coded three randomly selected peers' comments for initial coding and then compared the results. Differences were resolved through discussion. The interrater reliability (based on the percentage of agreement) for the steps was .90 and for the stances was .81.

RESULTS & DISCUSSION

Table 1 shows the breakdown of the specific steps that instructor/researcher's comments consisted of. The instructor/researcher generated a total of 412 entries of written commentary before implementing the peer review training. Of these comments, 24.3% were coded as clarifying writers' intentions, 23.5% as identifying problems, 16.7% as explaining the nature of the identified problems, and 35.4 % as making specific suggestions

After the peer review training, the instructor/researcher made 363 comments, 32.5% of which belong to clarification of writers' intentions, 17.9% to problem identification, 15.2% to

problem explanation, and 34.4% to provision of specific suggestions. A general picture of the instructor/researcher's adoption of various stances is the same before and after the peer review training. She adopted the collaborative stance more frequently, followed by the probing stance, the prescriptive stance and the tutoring stance.

Table 1. *The Percentage Breakdown of Instructor Comments Comprising each Step and Corresponding Stance*

Step Stance	CWI Pro	IP Pre	EP Tut	MSS Col	Total
prior to training	100	97	69	146	412
%	24.3	23.5	16.7	35.4	99.9*
after training	118	65	55	125	363
%	32.5	17.9	15.2	34.4	100

Note. CWI: Clarifying writers' intentions IP: Identifying problems
 EP: Explaining problems MSS: Making specific suggestions
 Pro: Probing Pre: Prescriptive
 Tut: Tutoring Col: Collaborative

*Percentages may not add up to 100 due to rounding.

A closer look, however, reveals that the instructor/researcher inquired about writers' intentions more (24.3% versus 32.5%) but identified fewer problems after the peer review training (23.5% versus 17.9%). In fact, a paired difference analysis of the instructor/researcher's stance reveals a significant difference in these two stances before and after the peer review training (Table 2).

Table 2. *Teacher's CWI and PI before and after Training*

Paired Differences						
	M	Std. D	Std. Error M	t	df	Sig.
CWI before training	-1.00	1.41	.33	-3.00	17	.008*
CWI after training						
PI before training	1.78	2.94	.69	2.56	17	.02*
PI after training						

After the peer review training, the teacher asked significantly more questions to probe students' intentions and identified significantly fewer problems. Both reflected a shift in her general stance toward a more probing than a prescriptive one, rendering her more like an active reader than a fault-finding teacher, corroborating her perception of feedback provision.

CONCLUSION

Contrary to previous research which reporting mismatches between ESL/EFL writing teachers' perceptions of and actual performances on written feedback provision (Lee, 2009;

Montgomery & Baker, 2007), the instructor/researcher's feedback practice before the peer review training matches her perception of how written feedback should look like given that most are specific and relevant to the students' writing. The implementation of peer review training might have prompted the instructor/researcher to self-monitor her own teaching practice (either subconsciously or consciously) and to fine-tune her stances (again either subconsciously or consciously) toward students and their texts. She assumed the probing stance more frequently and the prescriptive stance less frequently after the peer review training, shifting her role to that of a collaborating peer.

The pedagogical implication of this study appears to point to a need for an explicit procedure for EFL/ESL writing teachers to follow and self-monitor their own feedback practice to avoid mismatches between their beliefs and practices (Montgomery & Baker, 2007). Frequent self-examination of one's own feedback practices is perhaps more helpful to our writing students in the long term than sporadic evaluation by researchers.

REFERENCES

- Ferris, D. (2003). *Response to student writing: Implications for second language students*. NJ: Lawrence Erlbaum Ass.
- Goldstein, L. M. (2004). Questions and answers about teacher written commentary and student revision: Teachers and students working together. *Journal of Second language Writing, 13*, 63-80.
- Lee, I. (2009). Ten mismatches between teachers' beliefs and written feedback practice. *ELT Journal, 63*(1), 13-22.
- Lee, I. (2008). Understanding teachers' written feedback practices in Hong Kong secondary classrooms. *Journal of Second language Writing, 17*, 69-85.
- Liu, J., & Hansen, J. (2002). *Peer response in second language writing classrooms*. Ann Arbor: The University of Michigan press.
- Min, H. T. (2005). Training students to become successful peer reviewers. *System, 33*(2), 298-308.
- Min, H. T. (2006). The effects of trained peer review on EFL students' revision types and writing quality. *Journal of Second Language Writing, 15*(2), 118-141.
- Min, H. T. (2008). Reviewer stances and writer perceptions in EFL peer review training. *English for Specific Purposes, 27* (3), 285-305.
- Montgomery, J. L., Baker, W. (2007). Teacher-written feedback: Student perceptions, teacher self-assessment, and actual teacher performance. *Journal of Second language Writing, 16*, 82-99.
- Tsui, A. B. M., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments? *Journal of Second Language Writing, 9*(2), 147-170.

Comparing Synchronous and Asynchronous Computer-mediated Peer Review of EFL College Students

Chia-wei Pai (白家瑋), Hsien-chin Liou (劉顯親)

National Tsing Hua University

irispai0420@gmail.com

hsienchin@gmail.com

The current study aimed to compare asynchronous versus synchronous computer-mediated peer review (CMPR) of 18 EFL English-major freshmen using *Google Docs* and *Google Talk* as the review tools in two writing cycles. Copies of asynchronous comments and transcripts of synchronous talk were analyzed in terms of the function as well as the content of peer feedback. In addition, students' first and revised drafts were also analyzed to examine whether different review modes would shape revision changes. Finally, students' responses were collected to reveal their perspectives on experiencing CMPR. Findings of the present study can contribute to a better understanding of how media types affect the nature of peer review activities.

INTRODUCTION

Peer review, ever since studies on L1 process writing has enthusiastically advocated its pedagogical advantages for collaborative writing (Ferris, 2003), has received wide attention in the field of L2 writing. Over the decades, research on peer review in L2 writing has yielded fruitful, though mixed, findings regarding its nature, characteristics, and impacts on students' revisions. Meanwhile, as the rapid growth of technology and increased use of computers in second language writing classrooms have become a popular, or more or less inevitable trend, the "multiplying points of contact" (Ware & Warschauer, 2006) between computers and L2 writing married the concept of CMC with peer review, resulting in a new fashion of review activities: computer-mediated peer review (CMPR), also commonly termed as electronic, online, or virtual peer review. CMPR, according to the CMC media that are chosen, can be implemented in two modes: synchronous versus asynchronous peer review, depending on either instant or delayed response time required of the participant/recipient of the online message. Although researchers have not yet reached a consensus on whether CMPR integrated in ESL/EFL writing instructions always outshines the traditional face-to-face mode of peer review, impacts as well as some pedagogical advantages of CMPR have been reported in previous studies. For example, Ho and Savignon (2007) indicated that CMPR offered the flexibility for learners to carry out peer review whenever and wherever they felt like to. Several studies (DiGiovanni & Nagaswami, 2001; Huang, 1998, as cited in Ho & Savignon, 2007; Kern, 1995; Skinner & Austin, 1999) also considered CMPR to be motivationally encouraging for students to provide feedback to each other in that it may lessen the anxiety during peer communication.

A large proportion of research on CMPR sheds light on comparing the characteristics of peer review between computer-mediated versus traditional face-to-face settings. Face-to-face peer review was said to advance CMPR on maintaining a more fluent flow of

negotiation, whereas CMPR was found to be facilitative and effective to the review process in that they made drawing students' attention to the writing tasks at hand easier (DiGiovanni & Nagaswami; 2001; Ho & Savignon, 2007). In addition, several studies further examined the influence of these two review modes on students' revision changes (Hewett, 2000; Liu & Sadler, 2003; Shultz, 2000; Tuzi, 2004) and found that, in general, revision generated under CMPR tended to include local changes, with a commonly noticed text-based feature, and more revisions took place in CMPR mode compared with face-to-face environments. The findings contributed to the pedagogical insights on integrating CMPR into L2 writing instructions, especially when writing teachers consider moving peer review activities to an online environment.

While most of the CMPR studies compared task characteristics or students' perspectives on peer feedback generated under face-to-face versus computer-mediated mode, little research has been conducted regarding the different features in synchronous versus asynchronous peer review. Honeycutt (2001) found that in asynchronous peer conferencing, students used greater and more specific reference to the texts than that in e-mail. Chien and Liou (2005) examined CMPR in synchronous *MOO* while Peng and Liou (2008) looked at asynchronous comments in blogs; their findings regarding content of peer feedback and the quality of revision to some extent differed. It may be the case that synchronous and asynchronous media tools would result in different outcomes in peer review activities. In fact, Herring (2002a, cited in Ho & Savignon, 2007) suggested that the medium used for communication affect language and language use. It was also claimed that synchronous and asynchronous peer review may be helpful for different purposes during the writing process (Breuch, 2004), or resulted in different types of suggestions that students provided (Honeycutt, 2001). Such claims, though methodologically insightful for writing instructors to implement CMPR into L2 writing classes, needs empirical supports. Thus, the purpose of the present study is to explore peer review activities under synchronous and asynchronous environments by examining features of negotiations and their impacts on subsequent revisions. Three main research questions are formed:

1. Do synchronous and asynchronous peer reviews differ in terms of the function and content of feedback?
2. Do synchronous and asynchronous peer reviews differ with regard to the nature of revision changes?
3. What are student's perceptions toward the synchronous/asynchronous peer reviews?

Findings of this study can help researchers and L2 writing teachers obtain a better understanding of different natures about peer reviews on different CMC modes and of how to adopt the two modes in their writing classes.

METHODOLOGY

Participants and setting

Participants were 18 English-major freshmen in a writing class. Other participants included the researcher serving as a teaching assistant and the course instructor, who is an experienced professor in teaching writing and has been teaching this class for many years.

The class met two times per week, two hours each time. The course was divided into two sessions: (1) a classroom session in which students received instructions on writing with the process approach as well as different genre and rhetorical knowledge, and (2) a computer lab session in which students carried out online peer review. Students were paired up into 9 dyads randomly and remained the partnership for the whole semester.

Writing cycles and writing tasks

2 writing cycles were formed in the present study, with each cycle containing 2 comparable writing topics sharing the same genre. In other words, students needed to write 4

essays in total. The researcher also conducted a pilot¹ study assessing the topic comparability under each genre. The Wilcoxon Matched-pairs signed-ranks test didn't show significant difference between the 2 topic in both cycles in terms of essay length and difficulty.

Peer review tools: *Google Docs* and *Google Talk*

The online peer review tools were *Google Docs* and *Google Talk*, two commercial services offered by the *Google*. They are free and available for anyone with a *Gmail* account.

Google Docs offers a user-friendly interface for creating or editing documents. Users can upload their documents to it or compose online with most of the basic functions commonly used in *Microsoft Word*, and it saves document files periodically and automatically so that the work will not be lost. Through the "Share" function in *Google Docs*, users can invite others who also have *Gmail* accounts as collaborators (viewing plus editing) or viewers (viewing only) to work on the same document. In the present study, for the purpose of monitoring students' writing process and the ease of collecting data, students were asked to put the course instructor and the researcher as collaborators on each writing task as well.

Another feature of *Google Docs* used in the present study is the "revision history." It allows authors and viewers who have shared a document to see a list of revision record, including the date, the person who have edited it, and the changes that have been made. By selecting two revision versions, users can have them compared to locate the changes which are marked with a different color. As for *Google Talk*, it is a synchronous chat program which allows instant information exchange.

One of the functions, "comment," in *Google Docs* was chosen as the tool of asynchronous peer review. This function allows reviewers to leave comments on any area in the document; they can be placed next to a word or a sentence, or at the end of the essay. Reviewers can also delete or edit the comments and change their background color to highlight them. Once a comment is left by a reviewer, his/her user name and the exact time are recorded and displayed next to the comment so that it is clear for not only students, but the instructor and the researcher as well to reference who leave this comment and its date.

The asynchronous peer reviews went through the followings procedures: student A in a peer dyad invited student B as a collaborator for Topic 1 in cycle 1 (student A also invited the course instructor and researcher); student B then used the "comment" function in *Google Docs* to provide feedback to student A's essay. After that, they switched roles. The same procedures were applied to Topic 4 in cycle 2.

As for synchronous peer review, student A and B invited each other as viewers with *Google Docs* for Topic 2 in cycle 1. After they read each other's essay, they logged in *Google Talk*, which allows users to chat through a small window at the bottom right corner on the screen. Student A and B then, through online chatting, provided peer feedback to each other. In cycle 2, the same steps proceeded with Topic 3.

Evaluation questionnaire

An evaluation questionnaire containing four parts was designed to investigate participants' responses to online peer reviews. The first part of the questionnaire contained four items on participants' personal information and previous peer review experiences. The second part of the questionnaire included eight items about students' perceptions on peer reviews while the third part contained twelve items on technology facilitation. In part two and three, each response was presented with a five-point Likert-scale (from five points for "strongly agree" to one point for "strongly disagree"). The final part of the questionnaire asked two open-ended questions which required participants to state their preference for

¹ The pilot study recruited 8 participants who were also English-major freshmen but outside the target class. The 8 participants were divided into 2 groups and each group was assigned one type of genre randomly; each participant composed 2 essays for the topics of the assigned genre. The researcher tested the topic comparability by comparing (1) the average number of words for the two topics in each genre, and (2) the score of the essays with the rating scales by Jacobs et al. (1981). A second rater was involved in the second part of the pilot. The inter-rater reliability was 0.94 and 0.97 for the 2 topics in genre A, and 0.95 and 0.98 in for the 2 topics in genre B (Spearman-Brown).

either online peer review mode as well as their reasons.

Procedures

The entire treatment lasted for fourteen weeks. In the first week in the lab, the researcher introduced *Google Docs* and *Google talk* to the students, and asked them to practice using *Google Docs* by composing a 100-word short paragraph in class. The purpose for this introduction was to help students familiarize themselves with these tools so as to minimize technical problems in peer review activities later. They were also encouraged to try out the “comment” function in *Google Docs* and chat through *Google Talk* with their classmates.

The second and the third weeks were the preparation weeks when students received peer review training² from the researcher and practice asynchronous comments and synchronous chats with two writing assignments in class. Training in peer reviews has been considered as an essential and necessary step in the process approach of writing, and the importance as well as effectiveness of training has been recognized in previous studies (Berg, 1999; Min, 2006; Paulus, 1999; Stanley, 1992). The students were trained based on the guidelines and suggestions of Ferris (2003) with peer review sheets designed by the researcher.

The actual writing cycles started since week 6. Each writing topic required two weeks to complete the whole writing process. Students were assigned a topic a week ahead; they composed a draft at home and brought that draft to class the next week for peer reviews. At the end of cycle 2, the evaluation questionnaire was given to each student.

Data collection and data analysis

Data of the present study came from five sources: (1) copies of asynchronous comments, (2) transcripts of synchronous chats, (3) copies of students’ first drafts of each writing task, (4) copies of students’ revised drafts from each writing task and (5) participants’ response to the evaluation questionnaire.

To answer the first research question, data from sources (1) and (2) were analyzed to examine the “function” of peer feedback using a modification (see Appendix A) from Mendonça and Johnson’s (1994) type of negotiations found in face-to-face peer reviews. The original categories have been used and proved effective in previous peer review research (Chien, 2005; DiGiovanni & Nagaswami, 2001), but due to the fact that the present study considered both synchronous chats and asynchronous comments, Mendonça and Johnson’s original categories were modified into “question, explanation, suggestion, evaluation, grammar and chatting.” In addition to the function, the present study also examined the “content” of peer feedback, using a modification (see Appendix B) from Suzuki’s (2008) coding scheme of type of negotiation episodes³ in peer reviews, and two types of episodes were considered suitable by the researcher to be adopted for the present study: LREs (Language-Related Episodes) and TREs (Text-Related Episodes).

To answer the second research question, data from sources (3) and (4) were compared to identify text changes in revision, and were analyzed by calculating the frequency of each type of text change using Suzuki’s (2008) coding scheme of LRCs (Language-Related Changes), which categorized changes into word, sentence, and discourse levels. Finally,

² There were total five peer review training sessions in the present design. The purpose of the first training section was to teach students how to give effective feedback, and in the following four sections, actual examples from students’ feedback were showed and discussed in class to further enhance the efficacy of training.

³ Suzuki’s (2008) original negotiation episodes contained three main categories: (1) language-related episodes (LREs), (2) text-related episodes (TREs), and (3) revision/writing task-related episodes (RREs). These categories followed the taxonomies of Swain and Lapkin’s (1995, 1998) units of negotiation of language form, Lockhart and Ng’s (1995) nonlinguistic negotiation, and analyses of revision changes in previous research (Faigley & Witte, 1981, 1984; Yagelski, 1995).

responses to evaluation questionnaires were analyzed, and the mean of each Liker-scale item was calculated to understand students' perspectives on online peer reviews.

FINDINGS

Research question 1:

Do synchronous and asynchronous peer reviews differ in terms of the types of feedback?

In asynchronous peer review, the total number of feedback made was 138, while the in synchronous mode, it was 256. The unit of feedback is defined as a complete elaboration of an idea that may or may not be segmented by punctuations or spacing in comments/chat logs. The percentage of each function of feedback is presented in Table 1. The first thing to be noticed is that chatting contributed to a large amount of proportion (31.6%) in synchronous peer review. In asynchronous mode, suggestion and evaluation were the two most frequent types of feedback, with suggestion (37.7%) higher than evaluation (30%). As for the synchronous mode, suggestion and evaluation also occurred more frequently than other types of feedback. Question and explanation were not salient in both modes, and students addressed more grammatical problems in asynchronous comments.

As for the content of feedback as seen in Table 2, the amount of LREs (Language-Related Episodes) as well as TREs (Text-Related Episodes) in both peer review modes weighted nearly equally, with TREs took up only slightly greater proportion in asynchronous comments (39.8%) compared with that in synchronous chats (40.5%). In addition, within LREs, it was found that feedback addressing word-level problems was the most frequent one in both modes.

Table 1 *Function of feedback in both modes of peer reviews*

		Question	Explanation	Suggestion	Evaluation	Chat	Grammar
Asyn.	percentage	7.2%	8%	37.7%	30%	2.2%	15.9%
Syn.	percentage	7.8%	6.6%	28.5%	17.1%	31.6%	8.2%

Table 2 *Content of feedback in both modes of peer review*

		Language-Related Episodes (LREs)			Topic-Related Episodes (TREs)
		Word level	Sentence level	Discourse level	
Asyn.	percentage	33.9%	18.8%	8.5%	39.8%
Syn.	percentage	39.2%	9.8%	10.5%	40.5%

Research questions 2:

Do synchronous and asynchronous peer reviews differ with regard to the nature of revision changes?

Students' first and revised drafts in the 2 writing cycles were collected to analyze the revision changes made during peer review. Students made 147 revision in asynchronous peer review and 261 revisions in synchronous mode. Results as seen in Table 3 showed that in both modes, sentence-level changes took up the largest proportion among the three levels, while discourse-level changes was the least frequent one.

Table 3 *Revision changes in both modes of peer reviews*

		Word level	Sentence level	Discourse level
Asyn.	percentage	31.3%	57.1%	11.6%
Syn.	percentage	34.1%	57.1%	8.8%

Research questions 3:

What are students' perceptions toward the synchronous/asynchronous peer reviews?

The evaluation questionnaire contained four parts. The first part collected participants' personal information, the second part was about their perceptions on peer reviews, the third part on technology facilitation in peer reviews, and the final part asked students' preference for either peer review mode (asynchronous comments or synchronous chats). Questionnaire items were presented with Likert-scales, ranging from "strongly disagree" (one point) to "strongly agree" (five points). The following section presents the second, third, and the last part of the questionnaire results.

DISCUSSION

The present study aims to examine computer-mediated peer review in asynchronous versus synchronous CMC environments. With regard to the first research question, findings suggested that there was a large difference on the number of chatting between the two modes, with synchronous peer review containing approximately one-third of the total feedback received. Although such finding is somehow predictable due to the inherent textual difference of the two CMC modes, the students' chat scripts nevertheless revealed certain interesting facts. Looking at what students were chatting during synchronous peer review, it was found that their chats could be roughly categorized into four types: task-managements, technical problems, on-topic chatting, and off-topic chatting.

Almost every pair of review partners spent a little amount of time on task-managements, deciding whose essays to be reviewed first and who would send the chat logs to the instructor. Sometimes when technical problems occurred such as disconnection from the Internet or mis-clicks on certain icons, students would stop reviewing and tried to tell each other how to solve the problems. The majority of chats belonged to the on-topic one, in which students' talks were often inspired by a certain idea, either in the essays or during the chats, and they would then exchange opinions or share experiences.

In addition to chats, it was also found that students used the function of evaluation and suggestion the most, especially in asynchronous comments. Such findings were similar to Peng's (2008) study, in which suggestion and evaluation also took up the majority of peer feedback types (other categories including clarification and chatting). A closer look at students' feedback revealed that students tended to give an evaluation first, usually a compliment on the essay or a paragraph, followed by suggestions on what could be improved. Such pattern was quite common in the present study.

As for the content of feedback, the results showed that students paid almost equal attention to language and writing topic during peer review, since in both modes, LREs and TREs roughly weighed equally. One possible explanation is that peer review training influenced the way students gave feedback; in training sessions they were taught to look at not only the grammatical mistakes but also the content as well as the organization of their partners' essay. Within LREs, students most frequently gave feedback on word-level problems, which might be due to the fact that such problems were easier for the students to identify. The ability to see the problems beyond sentences or paragraphs and to provide

constructive feedback for improvement may need more practice and experience in writing.

Research question 2 asks if there is a difference on revision text changes between synchronous and asynchronous online peer review, and the results showed that there wasn't the case. In both modes of peer review, sentence-level changes covered the majority proportion of text changes, and such changes mostly belonged to self-initiated editing instead of revisions based on peer's feedback. Interestingly, in the feedback students received, word-level feedback was the most frequent one, yet, some students tended to self-edit on sentences a lot more during revision. Such behavior might be explained as those students' idiosyncratic writing habits which resulted in the inconsistent findings that peer provided word-level feedback the most while authors had more sentence-level changes than word-level ones.

As for students' perceptions on synchronous and asynchronous peer reviews, the questionnaire results suggested that students generally held a positive attitude toward the use of computers in the writing class, including online peer review activities and peer-editing. Although none of the students had used *Google Docs*, a majority of them enjoyed using *Google Docs* to compose in the course, and found both the "comment" function in *Google Docs* as well as *Google Talk* useful peer review tools. In addition, students seemed to prefer receiving than giving feedback, which accorded with the report that they valued their peers' feedback and thought it was helpful for revisions. What's more, most students agreed that peer reviews helped them to read their own essays critically. It is possible that accumulated experiences of peer reviews enabled them to apply the same criteria to evaluate their own writing. Finally, when asked which peer review mode they preferred, students overwhelmingly (15 out of 18) replied that they liked asynchronous comments better. Common reasons were that they didn't have to arrange time for peer reviews, that they could have more time to think about what to comment, and that it was more clear and convenient for them to revise because those comments were displayed right under the texts.

CONCLUSION

The present study compares computer-mediated peer review in two different CMC modes, using asynchronous comments and synchronous chats as peer review tools. Results from the first writing cycle revealed that, apart from the inherent difference that synchronous peer reviews allowed more chats to take place, students tended to evaluate peers' work and give suggestions for revisions most frequently in both review modes, and they preferred using comments since it permitted delayed response which allowed them to think more critically about what to comment. In addition, observed from the four participants' revision changes, feedback generated from asynchronous comments versus synchronous chats didn't seem to cause much of a difference on subsequent revision changes. More data on text changes as well as students' feedback adoption rate are needed for analysis to reach to a conclusion for this part.

Two pedagogical implications can be drawn from the results so far. First, using comments may be a more efficient way for peer reviews and revisions. On one hand, it gives students more time to think and reflect while giving feedback; on the other hand, it directs students' attention to the text immediately because when revising with synchronous peer reviews, students may need to retrieve chat logs to look for useful feedback between lines. The second implication is that peer review training as well as orientation on online environments for peer reviews is recommended, especially when students are not familiar with such innovations in their previous learning experiences. In peer review training, teachers can guide students to various rhetorical patterns, introduce principles for reviewing peers' writing, and provide actual examples of what effective feedback is like. As for developing students' basic skills for specific technology use in the classroom, it is always necessary to

give a clear and step-by-step demonstration and allow a period of preparation time in which students can get familiar with the technology.

REFERENCES

- Berg, E. C. (1999). The effects of trained peer response on ESL students' revision types and writing quality. *Journal of Second Language Writing*, 8(3), 215-241.
- Brech, L. K. (2004). *Virtual peer review: Teaching and learning about writing in online environment*. Albany, NY: State University of New York Press.
- Chien, C. W. & Liou, H. C. (2005). Effects of online peer response on EFL college writing. Unpublished master thesis, National Tsing Hua University, Taiwan.
- Digiovanni, E., & Nagaswami, G. (2001). Online peer review: An alternative to face-to-face? *ELT Journal*, 55(3), 263-272.
- Faigley, L., & Witte, S. (1981). Analyzing revision. *College Composition and Communication*, 32, 400-414.
- Faigley, L., & Witte, S. (1984). Measuring the effects of revisions on text structure. In L. Bridwell (Ed.), *New directions in composition research* (pp. 95-108). New York: Guildford Press.
- Ferris, D. R. (2003). *Response to student writing*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Hewett, B. L. (2000). Characteristics of interactive oral and computer-mediated peer group talk and its influence on revision. *Computers and Composition* 17, 265—288.
- Ho, M. C. & Savignon, S. J. (2007) Face-to-face and computer-mediated peer review in EFL writing. *CALICO Journal*, 24(2), 269-290.
- Honeycutt, B. L. (2000). Comparing E-mail and synchronous conferring in online peer response. *Written Communication*, 18(1), 26-60.
- Huang, S. (1998). A comparison between Chinese EFL students peer response sessions held on networked computers and those held in a face-to-face setting. Bielefeld: The Annual Meeting of the International NELLE Conference. (ERIC Document Reproduction Service No. ED 423685)
- Jacobs, H., Zinkgraf, S., Wormuth, D., Hartfiel, V. and Hughey, J. (1981). *Testing ESL composition: A practical approach*. Rowley, MA: Newbury House.
- Kern, R. G. (1995). Restructuring classroom interaction with networked computers: Effects on quantity and characteristics of language production. *The Modern Language Journal*, 79(4), 457-476.
- Liu, J. & Sadler, R. W. (2003). The effect and affect of peer review in electronic versus traditional modes on L2 writing. *Journal of English for Academic Purposes*, 2, 193-227.
- Lockhart, C., & Ng, P. (1995). Analyzing talk in ESL peer response groups: Stances, functions, and content. *Language Learning*, 45, 605-655.
- Mendonça, C. O., & Johnson, K. E. (1994). Peer review negotiations: Revision activities in ESL writing instruction. *TESOL Quarterly*, 28(4), 745-769.
- Min, H. T. (2006). The effects of trained peer review on EFL students' revision types and writing quality. *Journal of Second Language Writing*, 15, 118-141.
- Paulus, T. M. (1999). The effect of peer and teacher feedback on student writing. *Journal of Second Language Writing*, 8(3), 265-289.
- Peng, Z. Y. & Liou, H. C. (2008). A study of blogging for enhancement of EFL college students' writing. Unpublished master thesis, National Tsing Hua University, Taiwan.
- Shultz, J. M. (2000). Computers and collaborative writing in the foreign language curriculum. In M. Warschauer & R. Kern (Eds.), *Network-based language learning: Concepts and practice* (pp. 121-150). Cambridge, UK: Cambridge University Press.

- Skinner, B., & Austin, R. (1999). Computer conferencing—Does it motivate EFL students? *ELT Journal*, 53(4), 270-279.
- Stanley, J. (1992). Coaching students writers to be effective peer evaluators. *Journal of Second Language Writing*, 1(3), 217-233.
- Suzuki, M. (2008). Japanese learners' self revisions and peer revisions of their written compositions in English. *TESOL Quarterly*, 42(2), 209-233.
- Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16, 371-391.
- Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal*, 82, 320-338.
- Tuzi, F. (2004). The impact of e-feedback on the revisions of L2 writers in an academic writing course. *Computers and Composition*, 21, 217-235.
- Ware, P. D., & Warschauer, M. (2006). Electronic feedback and second language writing. In K. Hyland & F. Hyland (Eds), *Feedback in second language writing* (pp. 105-122). Cambridge, UK: Cambridge University Press.
- Yagelski, R. (1995). The role of classroom context in the revision strategies of student writers. *Research in the Teaching of English*, 29, 216-238.

APPENDIX A

Coding scheme of function peer feedback (modified from Mendonça & Johnson, 1994)

Types of peer feedback	Definitions	Examples
Question	Peers request for further information or clarification on certain points. This request can be either an explicit question or a statement saying that something is unclear.	“...develop much more coordinately, what does coordinately mean? I could not find it in Yahoo dictionary” (from chat) “What profits? Is profit equal to advantage?” (from comment)
Explanation	Peers explain an unclear point, an opinion, or the content of the essay	“If you talk about the merits and demerits, then you need to have a very strong stand point if you support one side (or think it’s better). Because the views may agree with your opinions about the disadvantages, you need to think about good reasons to persuade them, not just talking about ambiguous theories.” (from comment) “if I just look quickly, I will be very confused” (from chat)
Suggestion	Peers suggest ways to change words, sentences, content, and organization of essays	“...you can say from those experiences I had learned in the school, I may use them in the future” (from chat) “I think you can make the third paragraph more connected to the others” (from comment)
Evaluation	Peers provide subjective opinions toward the essays, including complementing or criticizing	“the third paragraph is based on you high school life, but the example seems too less to support your opinions” (from chat) “I like your opinion in the last paragraph; it is very important to find your own learning method and attitude.” (from comment)
Grammar	Peers identify or correct a grammatical structure in essays	“it supply should be it supplies?” (from chat) “on the first section, each “department and supervisor” should be singular” (from comment)
Chatting	Peers chat on things unrelated to the essay or revision, such as greeting or off-task chats.	“it’s crazy to get up in 6 A.M. !!!!” (from chat) “And...I am sorry for doing my homework late” (from comment)

APPENDIX B**Coding scheme of content of peer feedback (modified from Suzuki, 2008)**

Types of peer feedback		Descriptions	Examples
LREs (Language-Related Episodes)	a. Word level	Feedback related to punctuation (a comma, a period, etc.), spelling (including typographical errors), capitalization, word form corrections, and vocabulary/word choice	“line 2 you should you gender quality” (from chat) “I think “arouse” may be more ideal than “promote.”” (from comment)
	b. Sentence level	Feedback related to sentence types or length of sentence	“I think you can make your sentence shorter, it’s hard to read when it’s too long” (from chat) “in your last paragraph, there should be a conjunction connecting the first sentence and the second sentence” (from comment)
	c. Discourse level	Feedback related to organization or paragraphing	“I saw there are two merits in ur second paragraph, if you wanna revise the composition u can try to delete the third paragraph and broaden the ideas u mentioned in the second paragraph and add more information on it to increase the content...” (from chat) “I think you should control the proportion of each paragraph” (from comment)
TREs (Text-Related Episodes)	Feedback related to topics, content and ideas of texts		“I like your point of view, “girls should learn how to do things without boys’ help” because I only think of without their help is such a terrible thing haha” (from chat) “I think you should provide more details about your experience in elementary school. They were not interested in the drama. But, they could come up with novel ideas. It’s a bit conflicting.” (from comment)

APPENDIX C

Coding scheme for revision text changes (Suzuki, 2008)

Types of Language-Related Changes (LRCs)	Descriptions	Examples
Word-level changes	Changes related to punctuation (add or delete a comma, a period, etc), spelling (including typographical errors), capitalization, word-form corrections, and vocabulary/word choice	Then, →Nevertheless, to acquire the ability to mutual respect → to acquire the wisdom to mutual respect
Sentence-level changes	Changes related to sentence types or altering the length of sentence	In my opinion, the most important benefit of coed is that the students can... →In my opinion, to develop a well-connected friend net can make one easier to succeed in the future. The most important benefit of coed is that the students can...
Discourse-level changes	Organization: changes within paragraphs or within an essay, or Paragraphing: changes to whole paragraphs; creating new paragraphs from existing ones	One participant combined the ideas in 2 paragraphs and made it into a new one. (due to space limit, examples are not shown here)

APPENDIX D**Questionnaire**

這份問卷擬了解同學們對這門寫作課活動的心得和想法。請選擇最符合實際感受的選項，並在方框內打勾。我們不會對外洩露個人資訊，答案亦不會影響各位同學在這堂寫作課上的成績，請安心作答，謝謝！可以用中文回答。

Part A

1. Personal information

Name (Chinese): _____ Student ID: _____

Age: _____ Sex: male female

Years of learning English so far: _____ years (or since _____)

2. Before taking this course, I _____ used computers to compose.

 Never Seldom Sometimes Often Always

3. Before taking this course, I have done peer review in my previous writing experiences (before I entered Tsing Hua). (if not, skip question 4)

 Yes. No.

4. In my previous peer review experiences, I've tried

 face-to-face (oral) peer review with my classmate written comments (non-oral) peer review with my classmate peer review that involves the use of computers, and how?
_____**Part B**

5. I found peer review training sessions useful for giving my partner feedback.

 strongly disagree disagree average agree strongly agree

6. I give feedback based on what I have learned in peer review training sessions.

 strongly disagree disagree average agree strongly agree

7. My partner gives me feedback based on what has been taught in peer review training sessions.

 strongly disagree disagree average agree strongly agree

8. I like to give feedback to my partner's essay.

 strongly disagree disagree average agree strongly agree

9. I like to receive feedback from my partner.

 strongly disagree disagree average agree strongly agree

10. My partner's feedback is constructive and is helpful for me to revise my essay.

 strongly disagree disagree average agree strongly agree

11. When I revise, I try my best to take my partner's feedback into consideration.

 strongly disagree disagree average agree strongly agree

12. The experience of peer review helps me read my own essays more critically

 strongly disagree disagree average agree strongly agree**Part C**

13. I like the incorporation of computers in this writing class.

 strongly disagree disagree average agree strongly agree14. I've used *Google Docs* before taking this course. Yes. No.15. I like to use *Google Docs* to write essays.

- strongly disagree disagree average agree strongly agree
16. I tried out different functions provided by *Google Docs* when I composed online.
strongly disagree disagree average agree strongly agree
17. I like the experience of using “**comments**” in *Google Docs* for peer review online.
strongly disagree disagree average agree strongly agree
18. The “comment” function in *Google Docs* is helpful for me to give my partner feedback.
strongly disagree disagree average agree strongly agree
19. I like the experience of using ***Google Talk*** for peer review online.
strongly disagree disagree average agree strongly agree
20. ***Google Talk*** is helpful for me to give my partner feedback
strongly disagree disagree average agree strongly agree
21. I like the experience of revising my essays online.
strongly disagree disagree average agree strongly agree
22. In general, technology (*Google Docs* and *Google Talk*) facilitates the writing process from drafting, peer reviewing, to revising.
strongly disagree disagree average agree strongly agree
23. I like doing online writing and online peer review at home, compared with doing them in class.
strongly disagree disagree average agree strongly agree

Because _____

24. At home, which peer review tool do you prefer?

- “Comment” in *Google Docs* *Google Talk*

Because _____

Part D

25. Which peer review tool do you think is better?

- “Comment” in *Google Docs* *Google Talk* Neither one

What are your reasons?

26. Do you have any suggestions for the peer review activities?

運用指示詞與議論文文本結構教導申論題寫作：

以研究所入學考試試題為例

Fu-yuan Shen (沈富源)
National Taitung University
shen.106@osu.edu

英文的指示詞(direction words)如定義(define)，比較(compare)，對比(contrast)，描述(describe)，舉例說明(illustrate)等辭彙，常出現在申論題(essay question)中，指引學生作答。學生在答申論題時，除了要具備學科領域(content area)的知識，也應該理解題目中的指示詞，才不會答非所問，偏離題意。試題如果有比較(compare)，即指示要說明兩項事物的相同點(similarities)。題目如果出現對比(contrast)，即暗示要闡明兩項事物的相異點(differences)。掌握指示詞可以避免作答時，弄不清楚是要解釋相異點，或相同點的窘境。同時，指示詞也可以做為段落發展的依據。換句話，把指示詞擴充成段落主題發展的方法(paragraph development method)，學生答題時就可以有明確的方法發展思緒，有條理地組織段落的內容，形成嚴謹的文本結構。文本結構反映出作者闡述推理的邏輯關係，並且可以精確地把訊息傳遞給讀者。議論文(expository text)中常見的主題發展的方法有描述(description)，舉例(exemplification)，分類(classification)，比較和對比(comparison and contrast)，因果分析(cause and effect)，過程分析(process analysis)，剖析(analysis by division)等。本論文要以英語研究所入學考試試題為例，首先說明如何教導學生辨識試題中的指示詞，或依題意上下文的提示(context clues)，歸納出相對應的指示詞。然後，再根據指示詞，選用適當的主題發展的方法及其文本結構。本論文嘗試以圖表呈現定型化的文本組織，設計填空作業練習單，以引導式作文的方式，訓練學生的申論題寫作能力。

申論題(essay question)是大學教育中最常見的作業形式，也是研究所入學考試的主要題型。如何回答申論題一直是學生要面臨的挑戰。如何有效地教導學生申論題寫作，更是大學英文作文教學中迫切的課題。本論文要以英語研究所入學考試試題為例，說明如何教導學生申論題寫作。首先，介紹指示詞(direction word)的概念，讓學生熟悉常見的英文指示詞。有了指示詞相關的知識，學生就可以輕易地辨識，入學考試試題中的指示詞，或依題意上下文的提示(context clues)，歸納出相對應的指示詞。然後，學生再根據指示詞，選用適當的段落發展方法(paragraph development method)。為了使學生能快速有效地學習文本的組織架構，筆者嘗試利用圖表，結合定型化的內容組織模式(organization pattern)，設計填空作業練習單(fill-in-the-blanks worksheet)。學生依照作業單的提示來寫作，增進對文本結構的認識。本教學策略希望以引導式寫作(guided writing)方式，訓練學生靈活運用各種文本結構回答申論題，提昇英文寫作能力。

作文教學理論

設計作文教學活動，可以參考 Cambourne(1998)所提出的學習模式(model of learning)：示範(demonstration)、沈浸(immersion)、參與(engagement)。所以，在作文教學中，教師要先示範寫作的技巧，然後提供範文使學生沈浸其中，熟稔用字遣詞及文本組織結構。最後，教師要規劃符合真實情境、有意義的作業(authentic task)讓學生參與。只有真實有意義的作業才能引發學生的寫作動機，學生才會全心投入寫作，而不會只是交差了事。這是因為寫作的本質是具有社會性(social)及個人強烈色彩(personal)的行為活

動。寫作是一種文化上理解的溝通，和人的生活息息相關，因此學生如果無法和寫作產生個人聯結，是很難提筆作文的(Vygotsky, 1978)。依據上述的理論所發展出來的作文教學法，即是「過程法」(a process approach)，強調課堂上學生參與實際寫作的真實過程：學生自己選訂有興趣的題目來抒發情感思想，接著擬定(planning)主題、撰寫(drafting)、修改(revision)草稿、校對(proofread)文稿(e.g., Atwell, 1998; Casey & Hemenway, 2001; Graves, 1983; Murray, 1985)。

另一派的教學法是「技巧-成果法」(a skill-product approach)，以教導學生寫作的技巧(writing skills)為主(Bromley, 2003)。此類傳統的教學法以教師傳授技能為中心，把寫作視為獨立的認知(cognitive)能力。教學的目標是傳授語文讀寫能力(literacy skill)，教學的過程中，較不重視學生個人的寫作動機目的。理想完整的寫作教程應該設法整合「過程法」和「技巧-成果法」，讓學生在體驗寫作的過程中，學到應有的作文技巧(Fletcher & Portalupi, 1998; Harwayne, 2001; Portalupi & Flecher, 2001; Shanahan, 1997)。綜合上述的教學理論，Bromley(2003)認為完善的作文教學應俱備五項特點，分別是：(1)師生有明確的教學與學習目標；(2)教學環境中提供充份的寫作工具、時間、範例；(3)直接明確地教導作文的技巧、文章的規則；(4)彈性選擇寫作文類，並且指派的作業要有真實性(authenticity)；(5)跨學科領域的寫作，練習各種形式文類的作文。本教學研究即依據上述的教學相關理論，結合「過程法」和「技巧-成果法」，試圖利用指示詞來教導申論題寫作，設計出有效的教學活動。

指示詞的功用

學生在回答申論題時，除了要具備學科領域(content area)的知識，也應該理解題目中的指示詞。申論題的命題者常會用指示詞，來要求學生對某一議題做討論，指引學生作答。Langan (2005)整理出英文中常見的指示詞：比較(compare)，對比(contrast)，批評(criticize)，定義(define)，描述(describe)，繪圖(diagram)，討論(discuss)，依序列出(enumerate)，評估(evaluate)，舉例說明(illustrate)，詮釋(interpret)，說明理由(justify)，列出明細(list)，列出大綱(outline)，證明(prove)，重點說明(related)，摘要(summarize)，追源溯始(trace)。此外，筆者以最近十年外文/英語研究所碩士班入學考試試題為資料¹，篩選出常用的指示詞有：解釋(explain)，詳細說明(elaborate)，設計(design)。這些指示詞的用法，依字母排序，詳見表 1。

表 1 指示詞		
比較	Compare	表明事物間的相同點
對比	Contrast	表明事物間的不同點
批評	Criticize	對某一主題提出正面及負面的觀點，並提出證據來佐證正面、反面的立場
定義	Define	為一個名詞訂下正式的意義
描述	Describe	說明某事的細節
設計	Design	為某特定的目標，而發展出的計畫
繪圖	Diagram	畫圖並加上註明
討論	Discuss	說明細節，如果相關的話，應對某一主題提出正面及負面的觀點，並提出證據來佐證正面、反面的立場

¹ 本論文主要以台大、政大、師大、清大、交大、成大等校的研究所入學考試試題為研究資料。

詳細說明	Elaborate	整合相關的概念深入探討
依序列出	Enumerate	依 1, 2, 3 號次, 羅列觀點
評估	Evaluate	對某一主題提出正面及負面的觀點, 並且說明你的評價以及為何做如此評價的理由
舉例說明	Illustrate	舉例解釋說明
詮釋	Interpret	解釋某事的意義
說明理由	Justify	說明某事的原委理由
列出明細	List	列出一系列的觀點, 再加以標號 1, 2, 3
列出大綱	Outline	列出主要和次要的觀點
證明	Prove	提出事實或理由證明某事是正確的
重點說明	Relate	說明重點
摘要	Summarize	簡要地說明重點
追源溯始	Trace	詳細說明某事的發展及歷史

表 1 的指示詞牽涉到理解、批評、分析、重整等高層次的思考(high-order thinking), 包涵各種不同層面的認知。在教學時, 可利用 Bloom (1956)的分類原則(taxonomy), 把這些指示詞細分為知識(knowledge), 理解(comprehension), 應用(application), 分析(analysis), 整合(synthesis), 評量(evaluation)等六種認知的層次, 以便學生學習。單純記憶所學過的即定事實、專有名詞、理論概念, 屬於低層次的知識背誦, 例如定義(define), 依序列出(enumerate), 列出明細(list)。但是, 如果要測試學生是否真正理解, 只是記誦知識是不夠的, 而是必須要求學生做比較(compare), 對比(contrast), 描述(describe), 舉例說明(illustrate), 詮釋(interpret), 列出大綱(outline), 重點說明(relate), 摘要(summarize), 追源溯始(trace)等。另外, 繪圖(diagram), 設計(design), 詳細說明(elaborate)則是屬於整合所學的原理、概念、規則等提出新構想或解決問題的方案。證明(prove), 說明理由(justify)屬於分析的範疇。評量是指批判觀點的正確性, 或評比作品優劣, 以陳述自己的意見主張, 捍衛自己的立場。屬於評量的認知能力有: 批評(criticize), 評估(evaluate)。援用 Bloom 的認知能力分類, 來介紹指示詞, 可幫助學生把看似繁多的指示詞, 簡化成六種認知層次, 比較容易學習。

雖然指示詞林林總總, 有些卻可視為類似詞, 例如批評(criticize)和評估(evaluate), 均是要對某事物從其正/反, 優/劣, 得/失, 做客觀的判斷。相似的指示詞還有, 依序列出(enumerate)和列出明細(list), 都是要依序有條理地列舉數個要點。另外, 有些指示詞看似相同, 其實是不一樣的, 例如比較(compare)和對比(contrast)。比較(compare)是要說明兩項事物的相同點(similarities); 對比(contrast)是要闡明兩項事物的相異點(differences)。學生常會把比較與對比混為一談, 誤解題意而答非所問。所以, 熟悉並且能清楚地辨識指示詞相關的辭彙, 是回答申論題的第一步。

研究所入學考試試題分析

向學生介紹指示詞的基本定義後, 可以拿研究所碩士班入學考試中的申論題為範例, 分析其中的指示詞。教師示範分析試題, 可以加深學生對指示詞的認識。試題中有明顯的指示詞, 是最容易下手的。例如: Compare and contrast Beowulf and Chaucer's works with other texts of your choice to define "English literature" in the Middle Ages (台大外文所 96 學年度英國文學史試題)。本試題同時出現比較(compare)和對比(contrast), 考生必須找出《貝奧武夫》、喬叟的作品, 及其他中世紀英國文學作品之間的相同點和差

異處。類似的命題結構，還有：Compare and contrast the communicative approach and the oral (audio-lingual) approach(師大英語研究所 91 學年度英語教學試題)。此題要考生找出「溝通式教學法」和「口語(視聽)教學法」之間的異同。比較/對比的命題橫跨各領域，似乎是出題者的最愛，語言學概論也有類似的題型：Compare and contrast the consonants in English and Mandarin Chinese in terms of place of articulation, manner of articulation and voicing (師大英語研究所 92 學年度試題)。其他常出現在研究所碩士班入學考試申論題中的指示詞及其例子，請參見表 2。

表 2 研究所入學考試申論題常用的指示詞

比較	Compare	Compare Community Language Learning and Cooperative Language Learning in terms of the role of teacher, the role of learner, goals of language learning, characteristics of the teaching/learning process, and how evaluation is accomplished. (師大英研所 92 學年度英語教學試題)
對比	Contrast	Both Homer's <i>Odyssey</i> and Dante's <i>Divine Comedy</i> use journey to structure the works. Contrast the usage of journey and its significances in both works. (清大外語所 91 學年度歐洲文學史試題)
批評	Criticize	Earlier researchers had a lot of focus on good language learners and strategy categories. Please provide a critique of the assumption and intention of such research. (政大英文系碩士班 97 學年度英語教學理論與實務試題)
定義	Define	Define holistic and analytic scoring of writing. (師大英研所 92 學年度英語教學試題)
描述	Describe	Describe the role of context and background knowledge in the comprehension process. (師大英研所 96 學年度英語教學試題)
設計	Design	... Design a reading activity or task for second graders in senior high school that addressed all of the six learning styles. (師大英研所 95 學年度英語教學試題)
繪圖	Diagram	... Draw two tree diagrams for the original sentence, one for each meaning of the ambiguity. (師大英研所 92 學年度語言學概論試題)
討論	Discuss	Discuss William Makepeace Thackeray's classic novel <i>Vanity Fair</i> in its adaptation into a Hollywood film. (台大外文所 96 學年度英國文學史試題)
詳細說明	Elaborate	In what sense can romanticism be described as "humanistic"? Elaborate in detail and illustrate it with specific literary texts. (台大外文所 96 學年度英國文學史試題)
評估	Evaluate	Criticize and evaluate one of the following novels in terms of its narrative technique, theme and characterization. (清大外語所 91 學年度文類-小說-試題)
舉例說明	Illustrate	Define and illustrate the phenomenon known as "expressionism" in [a] variety of twentieth-century plays. (清大外語所 90 學年度西洋文學-戲劇-試題)

詮釋	Interpret	Read the following two texts carefully and then write a short critical essay on each given text. That is, try to interpret the text in terms of its form, content or any critical perspectives or theories. (成大外文所 95 學年英文閱讀與評析試題)
說明理由	Justify	Can the canon of sixteenth- and seventeenth-century English literature do without Shakespeare and Milton? What will we gain or lose if we exclude the two writers? Justify your argument with specific examples. (台大外文所 91 學年度英國文學史試題)
列出明細	List	... List the phonetic transcription of four different ways in which Aaron would say "cat", when he obeys all his constraints. (交大外國語文學系碩士班 97 學年語言學概論試題)
列出大綱	Outline	In a cogent and comprehensive essay, critically outline and assess the tradition of American poetry from the seventeenth century to the nineteenth. (台大外文所 90 學年度美國文學史試題)
摘要	Summarize	Summarize the following critical text in your own words and elaborate the major arguments. (清大外語所 97 學年度文學批評與理論試題)
追源溯始	Trace	... Use these words of [Coleridge] as a lead to trace a significant motif that appears in the works of Blake, Wordsworth, Byron and Shelley, and substantiate it with discussion of specific references. (台大外文所 93 學年度英國文學史試題)

上述的例子是試題中有明確的指示詞，可供學生依循；但是，有時命題者在試題的文句中並沒有用指示詞。這時候就必須根據題意，找相對應的指示詞。通常這種情形，試題是以 WH-問句的形式，典型的例子是：What are morphemes, words, compounds and phrases? How are they different? Give an example for each of them (師大英研所 91 學年度語言學概論試題)²。本試題要求考生，舉出實際的例子來定義詞素、字彙、複合名詞、片語，並且說明他們之間的差異。試題雖然沒有指示詞，但依照題意，相對應的指示詞應該是：定義(define)、對比(contrast)、舉例說明(illustrate)。學生可以嘗試利用他們所學到的指示詞，來改寫此題目為：Define morphemes, words, compounds, and phrases. Illustrate the contrast among them。另外一個 WH-問句試題：What are some of the advantages and limitations of applying the communicative language teaching approach in Taiwan? (交大外國語言與文化研究所 93 學年度英語教學概論試題)。本試題雖然沒有明確的指示詞，但依題意考生必須探討，在台灣運用「溝通式教學法」的優劣，所以學生可以用指示詞「評估」(evaluate)來改寫此題目為：Evaluate the application of communicative approach to language teaching in Taiwan。練習利用指示詞來改寫 WH-問句試題，可以增進學生對指示詞的認識。

²類似的 WH-問句的例子還有：What is bilingualism? What are the implications of bilingualism in language teaching and learning in Taiwan? (交大英語教學研究所 97 學年度英語教學概論試題)，學生可以試著改寫本試題為：Define the term bilingualism and discuss its implications in language teaching and learning in Taiwan。

如果試題中所出現的指示詞，例如 analyze, comment, explain 等，沒有列在指示詞 1 中，學生還是可以依據題意上下文，找出對應的指示詞³。例如：Comment on the following poem, entitled "O Moon, when I gaze on thy beautiful face" (台大外文所 94 學年度英國文學史試題)。試題要考生對標題為"月亮呀！當我凝視妳的美貌"這一首詩作評論，實則要學生詮釋(interpret)此詩詩句的涵義。「評論」在此等同指示詞「詮釋」。所以，學生只要熟記理解表 1 中的指示詞，日後在參加考試，無論試題中有無出現明確的指示詞，學生都會判斷題意，在腦海中找出相對應的指示詞。

認識指示詞，對釐清試題的題意有很大的幫助。例如：What are the characteristics of standardized language tests? In what ways are they different from teacher-made tests? Please compare two types of language tests in terms of their purposes, approaches, influences on classroom teaching and language learning (師大英研所 96 學年度英語教學試題)。此試題的前半部，要考生回答「標準化測驗」的特徵為何，雖然試題中沒有明確的指示詞出現，學生依題意仍可找出相對應的指示詞，例如依序列出(enumerate)和列出明細(list)「標準化測驗」的特點。此試題的後半部，要考生詳細申論「標準化測驗」和「教師自行命題的測驗」在教學和學習上會造成那些不同的效果。但是，試題中卻用比較(compare)似乎是要說明「標準化測驗」和「教師自行命題的測驗」的相同點，用辭顯然不妥，因為題目中已經提示考生，「標準化測驗」和「教師自行命題的測驗」是截然不同。所以本試題應使用對比(contrast)比較恰當⁴。雖然有時候，考題在選用指示詞有瑕疵，學生如果對指示詞有充份了解，還是可以根據題意判斷，辨識出正確相對應的指示詞。

在教導學生指示詞時除了利用表 1 介紹各種指示詞的意義，筆者也會蒐集多份文學、語言學、英語教學各領域的試題供學生閱覽，讓學生相互討論試題中所出現的指示詞及其涵義。教師的說明和同學間相互討論，都可以加深學生對指示詞的理解。

主題發展的方法

掌握指示詞可以避免作答時，搞不清楚是要說明相異點，或討論相同點的窘境。同時，指示詞也可以做為段落發展(paragraph development)的依據。換句話，把指示詞擴充成議論文(expository essay)的文本結構(text structures)(e.g., Rubin & Hansen, 1986)，學生答題時就可以有明確的方法，發展思緒申論主題，組織段落的內容。筆者在教導寫作時，一向主張：先決定要採用何種申論主題的方法，然後再思索要表達的內容，即先結構後內容。這是因為有了清楚明確的申論主題的方法，就可以形成良好的文本結構。文本的組織結構反映出作者闡述推理的邏輯關係，並且可以幫助作者精確地把訊息傳遞給讀者(Ogle & Blachowicz, 2002)。所以，教導主題發展的方法，有助於學生理解非文學

³ 一般來說，沒有列在表 1 中的提示詞，例如分析(analyze)，解釋(explain)，評論(comment)，都屬於廣義詞(umbrella term)可以涵蓋多個指示詞。例如，「分析」可以探討因果關係，也可以比較相同或差異，或做追源溯始的探究。「解釋」可以是指說明澄清，類似「詮釋」(interpret)，也可以是指提出理由，類似「說明理由」(justify)。遇到這類廣義的提示詞，學生必須依題意來判斷，等同何種指示詞，確定發展主題的方向，較容易答題。

⁴ 另一個指示詞比較/對比，使用上混淆不清的例子是：Select a Victorian novel and a modernist one to form a comparable pair that is capable of illustrating the main differences between Victorian and modernist aesthetics. Discuss specific aspects of the two novels for such a comparison (台大外文所 91 學年度英國文學史試題)。試題要求學生舉兩部作品，來討論維多利亞時期的美學與現代主義的美學有何主要差異。題目中的 comparison (表明事物間的相同點)應該改成 contrast (表明事物間的不同點)較合適。雖然 Brandon (2006) 認為：「貪圖方便時，比較(compare)這個詞也經常適用於比較(comparison)與對比(contrast)」(p. 128)，但筆者參考各版本的作文專書，都有明確分辨「比較」與「對比」，鮮有混為一談。事實上，Brandon 自己編著的教科書，也有區別比較/對比。

類文本的段落組織，增進寫作技巧(Garner, Alexander, Slater, Hare, Smith, & Reis, 1986)。

議論文中常用來申論主題的方法有：敘述(narration)，描述(description)，舉例(exemplification)，分類(classification)，比較和對比(comparison and contrast)，因果分析(cause and effect)，過程分析(process analysis)，剖析(analysis by division)，定義(definition)，論證(argumentation)等方法(e.g., Langan, 2004；Picolo, 1987)。各種發展主題的方法，詳細說明參見表 3。每一種主題發展的方法，都有其獨特的內容組織型態，形成其專屬文本結構。

方法	英文名稱	說明
敘述法	Narration	用說故事的方式來闡明主旨
描寫法	Description	描述人、事、物的感知細節(sensory details)
舉例法	Exemplification	用實際的案例來申論主題
剖析法	Analysis by division	分析構成事物、概念的組成要素
過程分析	Process analysis	說明如何做某事的步驟過程
因果關係法	Cause and Effect	說明原因和結果
分類法	Classification	把事物概念分門別類
比較對比法	Comparison and Contrast	展示相同與差異
定義法	Definition	澄清特定名詞的意義
論證法	Argumentation	據理表達意見、主張、立場

議論文中每一個段落(paragraph)的基本組織，當然是主題句，支持句，結論句。在這個基本架構上，依照主題發展的方法，演變出獨特的文本結構。例如，比較/對比法的文本結構有：區塊組織(block organization)(e.g., Oshima and Hogue, 2006)，主體組織(subject by subject) (e.g., Brandon, 2006)，逐項組織(point-by-point organization) (e.g., Brandon, 2006; Langan, 2004; Oshima and Hogue, 2006)三種型態。「區塊組織型」是指比較主體 A 與主體 B 兩者的異同時，首先整段只說明 A、B 之間有哪些的相同點(similarities)，然後再用另外一整段列舉 A、B 的不同之處(differences)。先後的二段，分別申論 A、B 的相同點及不同點，形成兩大區塊。「主體組織型」是指用兩段分別敘述主體 A 和主體 B，在 X、Y、Z 比較項(point of comparison)的特點。「主體組織型」即是「一次只說明一邊」(one side at a time)的文本結構(Langan, 2004)。相反地，「逐項組織型」則是從比較項 X、Y、Z，一項一項的觀點(point by point)，同時交錯分析 A、B 兩者之間的異同。

因果分析法也有其特別的文本結構，有區塊組織(block organization)和連鎖性組織(chain organization)⁵。利用一整段只敘述原因或結果，即是「區塊組織型」。如果說明原因 A 會導致結果 B，再說明原因 C 會引起結果 D，因果循環交錯敘述，就是「連鎖性組織型」。一篇優良的議論文一定有嚴謹的組織架構，學習各種主題發展的方法及其文本結構，是申論題寫作必備的技巧。

指示詞、主題發展的方法、文本結構

⁵ 依照其他的主題發展方法而衍生出的文本結構，詳細說明可參考 Brandon (2006)，Langan (2005)，Oshima and Hogue (2006)的專書。

學生在了解指示詞及申論主題的方法後，就可以把兩者結合在一起，形成完整的文本結構概念。整個認知的流程應是：先辨認試題中的指示詞，根據指示詞對應到主題發展的方法；依照主題發展的方法，選擇適當即定的(fixed)文本結構。相關的指示詞對應到主題發展的方法，詳列在表 4。從表中可得知，類似的指示詞對應到一種主題發展的方法。例如，批評(criticize)，評估(evaluate)都是要對某事物討論其正/反、優/劣、得/失，是論證法用來表達意見、主張、立場時的辯證策略。此外，指示詞中的定義(define)和追源溯始(trace)皆可歸屬定義法，因為定義法可以是單純解釋名詞的意義，也可以擴大到從歷史發展的脈絡，做溯源式斷定其意義。設計(design)對應到過程分析，因為設計可以用過程分析的模式，規劃按部就班完成某項工作的流程。典型的例子如: Design a simple version of the lesson plan to teach the reading text in a 45-minute period using an approach that aims to enhance students' four skills in English (師大英研所 91 學年度英語教學試題)。本試題要考生設計閱讀教學的計劃，學生可以利用過程分析的方法，以時間流程的(chronological)概念，逐項說明教學步驟。

按照主題發展的方法，就可選定文本結構。有了明確的文本結構做依循，接下來的工作只要填入適當的內容，作文就可大功告成。文本結構可免除學生剛開始寫作，不知如何下筆的苦惱。文本結構幫助學生作文，應該有事半功倍之效。

表 4 指示詞對應到主題發展的方法	
指示詞	主題發展的方法
比較(compare)，對比(contrast)	比較對比法(comparison and contrast)
批評(criticize)，評估(evaluate)	論證法(argumentation)
定義(define)，追源溯始(trace)	定義法(definition)
設計(design)	過程分析(process analysis)
舉例說明(illustrate)	舉例法(exemplification)
說明理由(justify)，證明(prove)	因果關係法(cause and effect)

引導式作業練習單

為了使學生能快速有效地學習文本組織架構，筆者嘗試利用圖表呈現定型化的文本結構。以圖表為藍本，設計填空式的作業練習單，供學生練習文本結構。圖 1 是比較/對比法中，區塊組織的作業練習單。

圖 1 比較/對比法的段落發展：區塊組織 Paragraph Development Method – Comparison and Contrast: Block Organization
<ul style="list-style-type: none"> • Topic Sentence (central idea) _____ and _____ have several (certain) features (characteristics) in common. • Supporting Sentences (showing similarities) <ol style="list-style-type: none"> (1) _____ is similar to _____ in _____. (2) _____ resembles _____ in _____. (3) Just as _____, so _____. • Concluding Sentence On these grounds, I have come to the conclusion that _____.

It seems reasonable to conclude that _____.

圖 2 是比較/對比法中逐項組織的作業練習單。在圖中偶數的支持句是表達主體 A 和主體 B，在 X、Y、Z 比較項上的相同點；奇數的支持句則是指主體 A 和 B 的不同點。

相關的研究文獻指出學生需要範例來學習寫作(Calkins, 1994; Harwayne, 2000; McElveen & Dierking, 2001)，因為範例可供學生「讀寫借用」(literary borrowing)的模仿(Lancia, 1997)。所以，本教學中的作業練習單，除可以當作文本結構的範例，也可以引導學生練習文本組織架構。作業練習單，提供各種主題發展方法常用的句型、詞彙、文本結構，讓學生模仿⁶。學生有現成的句型、詞彙、文本結構可套用，不用再擔憂基礎的英文能力不足。學生也就可以有更多的時間來思索主題充實內容。

圖 2 比較/對比法的段落發展：逐項組織	
Paragraph Development Method – Comparison and Contrast: Point by Point	
<ul style="list-style-type: none"> • Topic Sentence (central idea) There are certain similarities and differences between _____A_____ and _____B_____. • Supporting Sentences (showing similarities and differences) <ol style="list-style-type: none"> (1) Unlike (In contrast to, As opposite to) _____, _____. (2) In _____X_____, _____A_____ and _____B_____ are the same. (3) _____ contrasts with (differs from, is different from) _____ in regard to (in respect to) _____. (4) _____A_____ can be compared with _____B_____ in _____Y_____. (5) _____; however (in contrast, on the other hand), _____. (6) _____A_____ can do _____Z_____, just as _____B_____ can. • Concluding Sentence These similar and different points lead to the conclusion that _____. 	

圖 3 是利用練習比較/對比法的作業單(圖 1)為文本架構，回答申論題: Compare and contrast Beowulf and Chaucer's works with other texts of your choice to define "English literature" in the Middle Ages，所寫的範例。圖 3 中粗體字的部份表示是完全借用圖 1 的句型，非粗體字的部份是另行補充說明，使文義更加完備詳盡。

⁶ 本作業練習單中沒有標示轉折語(transitions)或連接詞(connecting words)，教師在使用時要提醒學生在實際寫作時，句子與句子之間要添加適當的轉折語或連接詞，使文句有凝聚連慣性(coherence)。

圖 3 段落發展的範例：比較法

A Model of Paragraph Development: Comparison

The term “English literature” in the Middle Ages can not easily be defined without reference to *Beowulf* and *Sir Gawain and Green Knight*. ***Beowulf* and *SGGK* have certain characteristics in common. *Beowulf* is similar to *SGGK* in that the poets use the medieval source materials.** Both works are typical of the medieval romance tradition that a hero is engaged in an adventure. ***SGGK* resembles *Beowulf* in illustrating themes of bravery and loyalty.** *Beowulf* is more loyal than ambitious to fight for King Hrothgar. Gawain is proud of his devoted and faithful services for King Arthur. **Just as in *Beowulf* there is symbolic employment of objects and actions, so a variety of literary symbols are used in *SGGK*.** Grendel symbolizes the monstrous force of Nature. Gawain’s pentangle is obviously a biblical symbol. **It seems reasonable to conclude that *Beowulf* and *SGGK* exemplify some of the most characteristic features of medieval literature.**

結論

本教學計劃利用指示詞來教導申論題寫作的過程，大致符合作文教學成效自我評量(Marino, 1997)的原則⁷，具有下列幾項特點：(1) 教導學生認識議論文文本結構；(2) 提供範例輔助學習；(3) 自由創作和規範性的寫作取得平衡；(4) 學生親身體驗寫作的過程；(5) 融合過程法和技巧-成果法。認識指示詞可以幫助學生清楚理解申論題的題意。依循指示詞，學生就會選用適當的主題發展的方法，來闡述議論主旨。作業練習單有明確的文本結構，引導學生寫作，依指示填入內容。作業練習單雖然提供了句型、詞彙、文本結構，但是內容還是需要學生發揮自己的創意，陳述自己的觀點。在筆者課堂教學中，用來練習申論題的題目是開放，不硬性指派題目。學生可以依自己的興趣，在文學、語言學、英語教學等領域，自己選擇題目。利用研究所入學考試的申論題來練習寫作，可以使看似平淡的作文練習變成是一項具挑戰性及真實性的任務(a challenging and authentic task)。作業練習單可以調和自由創作和規範性的寫作。學生自己選擇題目，根據題目中的指示詞，自己再判斷該用哪種主題發展的方法。這一連串的寫作過程，可以使學生親身體驗寫作中構思的過程。整體而言，本教學除了傳授寫作技巧給學生，並且同時利用作業單提供寫作機會，讓學生練習所學到的寫作技巧。

參考書目

- Atwell, N. (1998). *In the middle: New understanding about writing, reading and learning* (2nd ed.). Portsmouth, NH: Boynton Cook.
- Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals*. New York: Longman.
- Brandon, L. (2006). *At a glance: Paragraph*. New York: Houghton Mifflin.
- Bromley, K. (2003). Building a sound writing program. In L. M. Morrow, L. B. Gambrell &

⁷ 依據 Marino(1997)，關於作文教學的成效，教師自我評量的項目分別是：(1) 學生如何成為作文好手？；(2) 學生寫作的目的？；(3) 誰會讀學生的作文？；(4) 有沒有練習到各種形式的作文？；(5) 教師是否有給學生學習的選擇權？；(6) 教師如何運用寫作的過程？；(7) 教師提供什麼樣的教導？；(8) 教師如何運用文學當作範文？；(9) 教師如何幫助學生理解作文中的常規會影響到意義？

- M. Pressley (Eds.), *Best practices in literacy instruction* (pp. 143-165). New York: Guilford.
- Calkins, L. (1994). *The art of teaching writing*. Portsmouth, NH: Heinemann.
- Cambourne, B. (1998). *The whole story: Natural learning and the acquisition of literacy in the classroom*. Auckland, NZ: Scholastic.
- Casey, M., & Hemenway, S. (2001). Structure and freedom: Achieving a balanced writing curriculum. *English Journal*, 90(6), 68-75.
- Fletcher, R., & Portalupi, J. (1998). *Craft lessons: Teaching writing K-8*. York, ME: Stenhouse.
- Garner, R., Alexander, P., Slater, W., Hare, V. C., Smith, T., & Reis, R. (1986). Children's knowledge of structural properties of expository text. *Journal of Educational Psychology*, 78, 411-416.
- Graves, D. (1983). *Writing: Teachers and children at work*. Portsmouth, NH: Heinemann.
- Harwayne, S. (2000). *Lifetime guarantees: Toward ambitious literacy teaching*. Portsmouth, NH: Heinemann.
- Harwayne, S. (2001). *Writing through childhood: Rethinking process and product*. Portsmouth, NH: Heinemann.
- Lancia, P. J. (1997). Literacy borrowing: The effects of literature on children's writing. *The Reading Teacher*, 50(6), 470-475.
- Langan, J. (2005). *College Writing Skills* (6th ed.). New York: McGraw-Hill.
- Marino, J. (1997). *Presentation to area educators*. Binghamton, NY: Board of Cooperative Educational Services.
- McElveen, S. A., & Dierking, C. C. (2001). Children's books as models to teach writing. *The Reading Teacher*, 54(5), 362-364.
- Murray, D. M. (1985). *A writer teaches writing* (2nd ed.). Boston: Houghton Mifflin.
- Ogle, D., & Blachowicz, C. L. Z. (2002). Beyond literature circles. In C. C. Block & M. Pressley (Eds.), *Comprehension instruction: Research-based best practices* (pp. 259-274). New York: Guilford.
- Oshima, A., & Hogue, A. (2006). *Writing academic English*. New York: Pearson Education.
- Piccolo, J. A. (1987). Expository text structure: Teaching and learning strategies. *The Reading Teacher*, 40, 838-847.
- Portalupi, J., & Fletcher, R. (2001). *Nonfiction craft lesson: Teaching information writing K-8*. York, ME: Stenhouse.
- Rubin, A., & Hansen, J. (1986). Reading and writing: How are the first two "R's" related? In J. Orasanu (Ed.), *Reading comprehension: From research to practice* (pp. 163-170). Hillsdale, NJ: Erlbaum.
- Shanahan, T. (1997). Reading-writing relationships, thematic units, inquiry learning . . . in pursuit of effective integrated literacy instruction. *The Reading Teacher*, 51(1), 12-19.
- Vygotsky, L. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.

Perceptions of English Language Use at Tourist Hotels in Taiwan

Wen-yuh Shieh (解文玉)

Minghsin University of Science and Technology

wyshieh@must.edu.tw

A good command and fluent use of English is an essential key to assuring the success of the hospitality service. To cultivate more qualified hospitality personnel into the workforce for the future, most of the college and university Hospitality Management Departments, the cradle of future hospitality personnel, require students to take a half-year to one-year off-school internship to gain field experiential learning experiences. This study examines the perceptions of university seniors on the hotel English use upon completing their required one-year internship. Seventy-five senior students having worked at different hotels filled out the survey questionnaires concerning their perceived workplace English needs and self-evaluated performance. The results reveal a unanimous agreement that the English language is perceived as very important at international tourist/ tourist hotels in Taiwan. Yet, most students indicate their dissatisfaction toward their workplace English performance. Among the four English skills, listening and speaking are considered more important and more frequently used than reading and writing. In regard to the needs and performance, speaking skill is perceived by these seniors as of high importance but low performance. Implications and suggestions from the results are also provided.

INTRODUCTION

English, the “lingua franca” of the world, is the most widely used language for information technology, international commerce, communication and travel. The rapid expansion in the use of English has aroused a fast increase of the English speaking population in the world today. As Taiwan is continuously developing its competitiveness in high-tech industries, and accelerating its economic activities in global business for the 21st century, and as the demand for people traveling to Taiwan has been ever-increasing, service-oriented industries, hospitality services in particular, play a vital role in providing accommodation for visitors to Taiwan. A good command and fluent use of English will be the key to assuring the success the hotel service to international visitors.

The prevailing recognition of the importance of the employees’ English language ability in relating to service quality of the hotels in Taiwan have been brought up by many researchers and educators (e.g. Horng & Lu, 2006; Lafrenz, 1991; Wang, 2001). A unanimous agreement among them is that English proficiency is a required job qualification for people working at the hotel sectors. To meet the increasing demand for hospitality personnel, and to provide hospitality majors to be ready to enter the workforce, many hospitality related departments at the tertiary education offer elective or required hospitality English courses to cultivate students’ language ability prior to their off-school internship. Through the internship field work experience, students will have opportunities to reflect what they have learned from school, experience the real-life work situation, become familiar with the industry, and increase their professional competency for the future (Fletcher, 1995). Any gap between their perceived field job needs and their self-assessed performance will manifest as a guide to improve their professional competencies and benefit future academic course design.

This study therefore focuses on the English language needs and the English language performance that students perceived during their internship year. These questions guided this

study: (1) What are the English language needs that students perceived during their internship? (2) How well do the students perceive concerning their English language performance during their internship? (3) What workplace English that the students perceive as high in needs yet low in their performance? In the following sections a brief review of the benefits of experiential internship learning in international hotels, workplace English needs and performance and the framework of the study will be presented.

LITERATURE REVIEW

The collegiate hospitality off-school internship system in Taiwan dates back to 1982 (Lee *et al.*, 2006) based on the idea that the field experiential learning integrates classroom study with practical work experience (Breiter *et al.*, 1995). It not only provides students with a practical understanding of the real-working professional skills (Wu, 2003), but also provides faculty with updated hospitality trends and demands (Dresh and Gregory, 1989). As such, the off-school internship program becomes a critically important component of hospitality program's curriculum.

Most hotel majors when serving their internship at international tourist/tourist hotels work at the front-of-house where customers/guests meet such staff as front desk at Front Office, wait people at Food and Beverage Division, and house attendants at the Housekeeping Department. Still a few interns work at the back of house which is not usually seen by guests, i.e., staff-only areas (Wagon, 2003). All the activities that guests experience in the hospitality service cycle may involve interactive communication including business transactions with front-of-house personnel in all the major departments of the hotel. Therefore, through the frequent encounters with hotel guests/customers in their daily work experience and observations, interns will be provided with ample opportunities to practice and sharpen their professional knowledge, communication skills and language proficiency as well (Lee *et al.*, 2006).

As noted earlier, proficiency in hospitality English is a required qualification for the job, but such proficiency cannot be promoted without first analyzing carefully what the needs for workplace language really are. Needs analysis (also termed needs assessment) is an information gathering process which has been viewed from more than one perspective (Richterich, 1983). For instance, Hutchinson and Waters (1987) say that for learners needs imply the three considerations, namely necessities, wants and lacks. The things that have to know to function effectively in the target situation are the necessities; the things that they think they need to know are the wants; the things that lie between the present level of proficiency and the targeted level are the lacks. Another, Berwick (1989), says that needs analysis is "the discrepancy between a current state of affairs and a desired future state" (Berwick 1989, p.52).

An important component of needs analysis is the assessment of learners' present language performance so as to measure the gap between present performance and target needs. Performance is the language user's actual production and comprehension of the language. It is as an observable construct (Savignon, 1983) serving as an index to reflect competency, the implicit or explicit knowledge a language user has of the system of language. Performance assessment thus refers to assessing how well employees use language knowledge and skills in the specific workplace of the hotel. Learners are assessed on "what they can do in situations similar to 'real life'". A self-administered performance assessment help learners to raise their awareness of metacognition and also enhances learning motivation (El-Knoumy, 2003). Several formats have been suggested for conducting performance assessment such as direct observation (Fisher, 1995; Tompkins, 2000), introspective and retrospective verbal responses to test tasks (Gibson, 1997; Wu, 1998), and questionnaires or interviews.

FRAMEWORK OF THE STUDY

As the aforementioned that language performance assessment is an approach which measures not only the process of language learning but also the products of language proficiency with specific tasks in a specific context. A framework of Hotel English Proficiency Assessment (HEPA) is then proposed to support the content in the workplace communicative performance in English in the special context of the hotel sector. Adapted from Bachman's (1989) framework of communicative language ability and Douglas's (2000) specific language ability, the HEPA contains a four-componential concept of competence, namely language competence, pragmatic competence, strategic competence and psychophysical modes and skills for assessing communicative performance at workplace.

Language competence refers to the ability of applying language knowledge to communication. Language competence in HEPA refers to the ability for use of linguistic components of vocabulary, grammar and phonology, through which language is structured. Pragmatic competence in HEPA comprises of situational/contextual competence, sociolinguistic competence, sociocultural competence and illocutionary competence (Bachman, 1989). In other words, the competence of performing situational appropriate locutions and the competence of interpreting illocutionary utterances are essential components in the pragmatic competence. Strategic competence consists of the components of communication competence and interpersonal competence. The former refers competence of language users when expressing his intention with or without encountering difficulty (Bialystok, 1990); whereas the latter refers to how one interacts with others to facilitate language learning and enhance communicative performance. Borrowed from Bachman's framework of communicative language ability, the HEPA framework also includes the psychophysiological skills. In addition to the realization of language competence in the modes of receptive and productive, and channels of auditory and visual, interactional medium via face-to-face, telephone, correspondence are also included in the psychophysiological skills. The framework of HEPA with its essential components is illustrated in Figure 1.

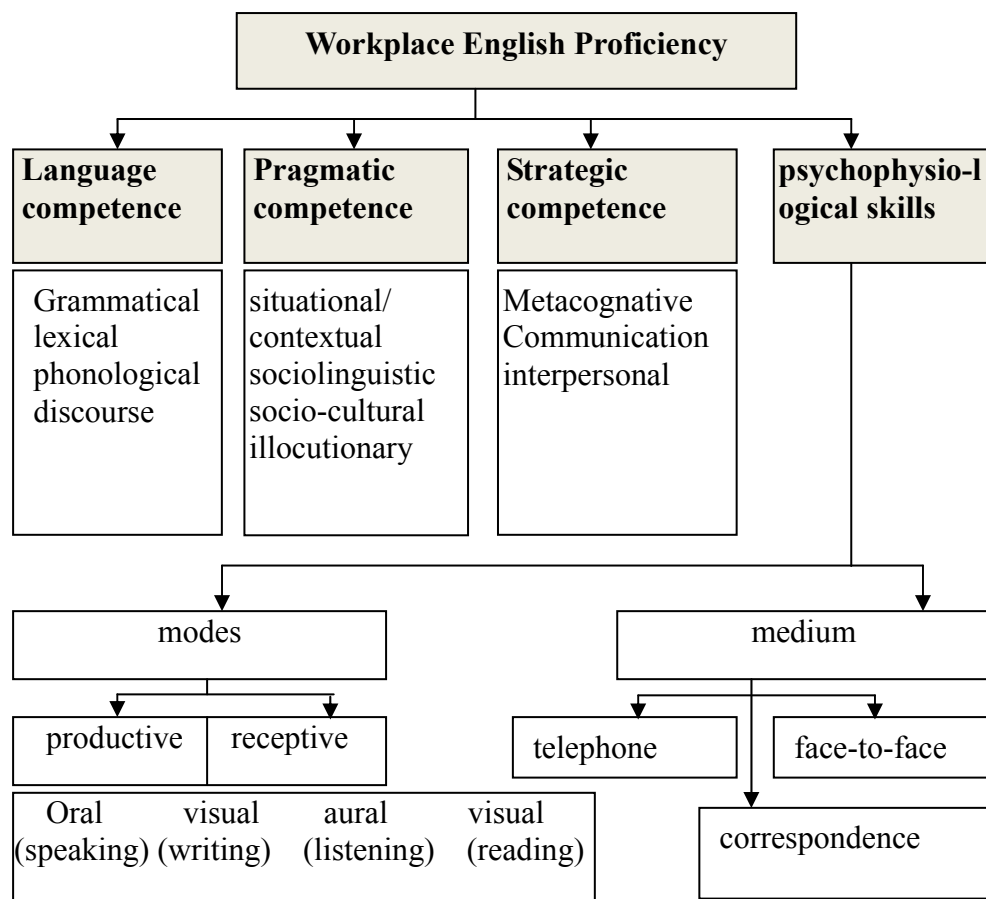


Figure 1: Framework of Hotel English Proficiency Assessment

METHODOLOGY

The participants in this study were 75 university hotel management majors who had just finished their one-year required internship at international hotels in Taiwan. The researcher self-administered the questionnaire survey at the end of a required class. Among the 80 questionnaires distributed, 75 were returned. Two questionnaires filled by respondents who served their internship at back-of-house were excluded, leaving 73 valid questionnaires collected for further data analysis.

Since an established survey of perceptions of hotel English cannot be located through an extensive review of the hotel English literature, the items on the questionnaires were constructed from issues and concerns revealed in the foreign language learning literature specifically related to the hotel context and also themes emerged from interviews with Hotel English instructors and hotel personnel including managers and front-of-house staff.

Constructed with the framework of HEPA, the questionnaire consists of three parts. Part I asks the respondents' demographic information. The items in Part II concern possible situations that the psychophysiological skills are applied to at work and self-assessed language performance of these situations. Part III probes the importance of linguistic skills and language components that respondents perceived and their assessment of how well they performed.

RESULTS AND DISCUSSION

Seventy-five students participated in the survey questionnaires. Since two of the respondents worked at back-of-house, a total of 73 questionnaires ($\alpha = .9644$) were valid for data analysis with Statistical Package for the Social Sciences for Window (SPSS 15.0).

Among these participants, there were 13 males and 60 females working at Front Office (26), Housekeeping Department (16) and Food and Beverage Division (33), respectively. Table One depicts the profile of respondents.

Table 1. *Profiles of respondents*

Features and Descriptions (Total=73)	Frequency	Percentage
Gender		
Male	13	17.81%
Female	60	82.19%
Departments		
Front Office	26	35.62%
Housekeeping	14	19.18%
Food and Beverage	33	45.21%
English Certificate		
No	63	86.30%
Yes	10	13.70%
GEPT Beginner's	6	8.22%
GEPT Intermediate	2	2.74%
Communication in English at work		
Never	0	0%
Seldom	14	19.18%
Sometimes	28	38.36%
Often	22	30.14%
Always	9	12.33%
English Language training course offered		
No	42	57.53%
Yes	31	42.47%
Workplace English phrases provided		
Yes	62	84.93%
No	9	12.33%
Self-assessed Overall performance		
No comments	3	4.11%
Very unsatisfied	10	13.7%
Not satisfied	47	64.38%
Satisfied	12	16.44%
Very satisfied	1	1.37%

As indicated in Table One, there were only eight students who had English language certificate equivalent to GEPT beginner's level and Intermediate level. All respondents used English at work at various degrees: Seldom (19.18%), sometimes (38.36%), often (30.14%), and always (12.33%). Half of the participating hotels offered on-the-job English language training, and 84.93% of them provided common workplace English phrases and sentences to the interns. Last but not least as indicated in Table one, one respondent revealed complete satisfaction of his/her self-assessed over all English performance at work, and twelve respondents (16.44%) showed satisfaction, ten (13.70%) not satisfied, and three (4.11%) marked no comments on their self-assessed English performance at work.

When asked how important the four linguistic skills, listening, speaking, reading and writing that they perceived relating to their workplace English use, the respondents indicated on the Likert-type of scale 1-5 from not important at all to very important an average tendency that listening was regarded as the most important skill (\bar{X} =4.29, SD=0.82),

followed by speaking ($\bar{X}=4.12$, $SD=0.93$), reading ($\bar{X}=3.93$, $SD=0.84$), and writing ($\bar{X}=3.68$, $SD=0.93$). When asked how well they perceived their workplace English performance on the Likert-type of scale from 1 to five indicating no comments, very unsatisfied, not satisfied, satisfied and very satisfied, the students revealed that listening received the highest average of 2.79 ($SD=0.74$), with a general tendency of average below 3. In other words, most of the interns were not satisfied with their workplace English performance. Table Two reports the mean of the importance of the four skills and how well the senior perceived their language performance during their internship.

Table 2: *Interns' perceptions of the importance of the four skills and self-assessed performance.*

	Perceived Importance \bar{x} (SD)	Self-Assessed Performance \bar{x} (SD)
Listening	$\bar{x}=4.29$ (0.82)	$\bar{x}=2.79$ (0.74)
Speaking	$\bar{x}=4.12$ (0.93)	$\bar{x}=2.55$ (0.75)
Reading	$\bar{x}=3.93$ (0.84)	$\bar{x}=2.85$ (0.76)
Writing	$\bar{x}=3.68$ (0.93)	$\bar{x}=2.67$ (0.71)

Seen in Figure 2 is the averages plotted on an x-y axis where the x axis presents the performance and y, needs. Speaking stands out as the skill which exceeds other skills with higher priority to be strengthened due to the higher importance but lower performance score when compared with other skills.

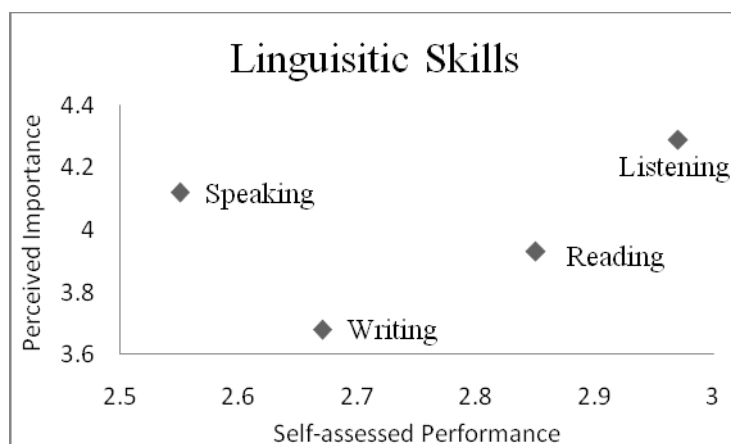


Figure 2: Needs and performance of the four modes of psychophysiological skills

Among the four listening situations, respondents indicated a discerned higher frequency in use with a better self-assessed performance score in the listening to guests' enquiries (Figure 3). The reason for this phenomenon might be that the frame of hospitality services in each stage of the guest cycle (Kasavana, 1993) conforms to certain predictable behavior and with certain consensus of language use, starting with the pre-arrival registration and ends with post-departure correspondence.



Figure 3: Perceived frequency in listening situations and self-assessed performance

In a similar vein, respondents indicated a higher frequency in use in replying to guests' inquiries with a higher score of self-assessed performance (Figure 4). In addition, 84.93% of the respondents revealed that they had received workplace English phrases from the department they worked at. The predictable service content, the possible guests' enquiries and the appropriate responsive language have been noticed and rehearsed, which added to the higher performance score of the interns. Table 3 shows the predictable activities in the guest cycle and possible language use for each stage.



Figure 4: Perceived frequency in speaking situations and self-assessed performance

Table 3. Revised six-stage hospitality services cycle

Stage	Activities	Language used
Pre-arrival reservation	Booking by email, Internet, facsimile or phone.	Standard forms filling out in writing or with guests' oral instruction. Formal language use.
Arrival	Airport picking up (optional); Front door greeting; Porter luggage serving	Routine, rehearsed, and formal language used by greeter, doorman or porter.
Familiarization	Receptionist checking in and registering for newly arriving guests. Giving brief introduction about hotel facilities, meal time, and taking and answering questions from the guests. Porter showing guests to their room. Room facilities introduction.	Rehearsed briefing style and functional, formal language use.
Engagement	Guests' individual use of hotel facilities and services such as restaurant, spa, gift shop, etc. Making transportation or entertainment arrangements.	Mostly formal and impersonal language use. Occasionally semi formal language will be used depending on how long the guests stay in the hotel and what facilities the guests use.
Departure	Checking out and bill settlement. Transportation arranging, luggage transferred, and farewell conversations.	Mostly rehearsed, routine formal language use.
Post-departure correspondence	Response to guests' email or letters concerning their requests or experience while staying in the hotel.	Formal written business letters.

(Adapted from Harun, 1998)

Reading and writing as the respondents revealed as the two less frequently used skills compared with listening and speaking. The various possible reading and writing situations did not reveal much difference in the perceived frequency of use and a low score in self-assessed performance (Table 4). Most respondents indicated that the perceived frequency of reading situation is less than sometimes ($\bar{x} < 2.5$), and the self-assessed performance revealed an average tendency of being not satisfactory ($2.00 < \bar{x} < 3.00$).

Table 4. *Perceived needs and self-assessed performance in reading and writing situations*

		Perceived Frequency of use \bar{x} (SD)	Self-Assessed Performance \bar{x} (SD)
Reading	Professional Articles	\bar{x} =2.07 (0.95)	\bar{x} =2.32 (1.01)
	Business Correspondence	\bar{x} =1.96 (0.98)	\bar{x} =2.40 (1.05)
	Work Manuals	\bar{x} =2.22 (1.03)	\bar{x} =2.51 (0.93)
Writing	Business Correspondence	\bar{x} =1.58 (1.00)	\bar{x} =2.01 (1.05)
	Business Emails	\bar{x} =1.64 (1.05)	\bar{x} =2.10 (1.07)
	Forms	\bar{x} =1.79 (1.03)	\bar{x} =2.21 (1.11)
	Messages	\bar{x} =2.07(1.26)	\bar{x} =2.78 (1.42)

In regard to language competence, vocabulary received a higher score in perceived importance than that of pronunciation and grammar (Figure 5). Second/foreign language learners need vocabulary knowledge to help them function well in the language. Folse (2004) points out that vocabulary knowledge is critical in any communication, and perhaps the most important component in L2 ability. Wilkins (1972) argues that “While without grammar very little can be conveyed, without vocabulary nothing can be conveyed” (p. 111). Lewis (1993) points out that without grammar, little communication may be possible. Yet, without vocabulary, no communication is possible. The results of the students’ perception give supports to the importance of vocabulary knowledge in workplace English.

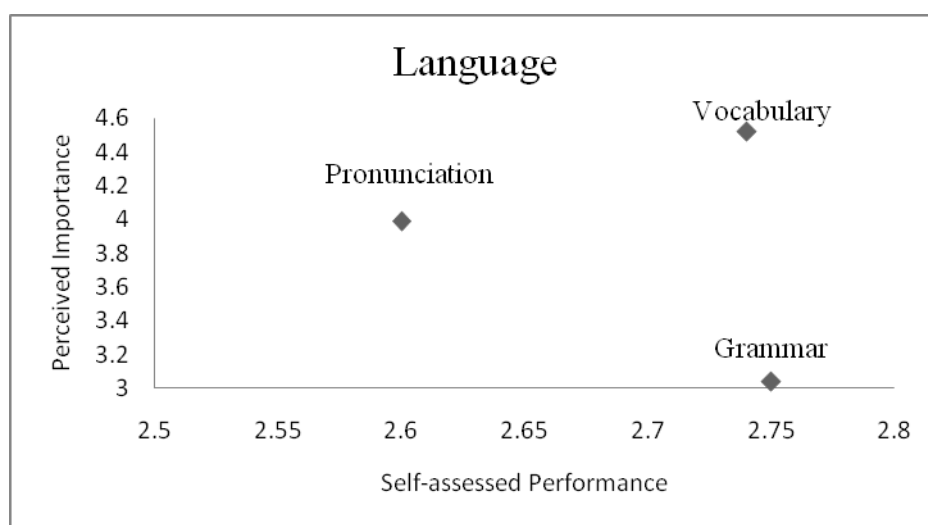


Figure 5: Perceived importance of language components and self-assessed performance

As for the perceived importance of other competence in pragmatic and strategy such as General English, handling emergency situation, the used of body language, using polite English, and an understanding of cross-cultural differences, an average over 4.00 was revealed in each item in the various competence, indicating a general agreement of the perceived importance of these competences. Yet, the perceived performance indicated an average of below 3.00 tendency revealing an unsatisfactory manner to the performance regarding the competence. Table 5 shows the results of the perceived importance of these skills and self-assessed performance.

Table 5. *Perceived Importance and Self-Assessed Performance of other competence.*

Competence	Perceived Importance \bar{x} (SD)	Self-Assessed Performance \bar{x} (SD)
General English	\bar{x} =4.44(0.69)	\bar{x} =2.95 (0.86)
Handling Emergency	\bar{x} =4.42(0.76)	\bar{x} =2.75 (0.88)
Using Body Language	\bar{x} =4.23 (0.83)	\bar{x} =2.92 (0.78)
Using Polite English	\bar{x} =4.48 (0.58)	\bar{x} =3.00(0.82)
Cross-cultural Understanding	\bar{x} =4.45 (0.90)	\bar{x} =2.68 (0.74)

SUMMARY AND CONCLUSION

Competence in English has been regarded as an essential and specific job requirement for existing and new personnel, especially for those holding in front-of-house posts requiring frequent face-to-face encounters with international guests (Wang, 2001). This study investigates university senior' perceptions of their hotel English use and self-assessed English performance after one-year internship at international tourist/tourist hotels. An analysis of the collected seventy-three valid questionnaires revealed that the English language is considered very important in hotel services, yet most students revealed a general tendency of dissatisfaction toward their English performance at work. Among the four language skills, listening and speaking are regarded more important at work than reading and writing. Speaking, however, revealed a higher importance yet lower performance. In addition, among the various listening and speaking situations, these students considered themselves more frequently encountered and performed better in listening to and responding to guests' enquiries.

Based on the results, an understanding of such gaps in the perception of hotel English needs and students' performance can help to draw up the criteria needed in curriculum to strengthen students' language proficiency and to consolidate the language required prior to their off-school internship service with the expectation of the a college education is to prepare students to be ready to enter the workforce.

REFERENCES

- Bachman, L. F. (1989). The development and use of criterion-reference tests of language ability in language program evaluation. In R. K. Johnson (ed.), *the second language curriculum*, pp. 242-258. Cambridge: Cambridge University Press.
- Berwick, R. (1989). Needs assessment in language programming: From theory to practice. In R.K Johnson, (Ed.) *The second language curriculum (pp.48-62)*. Cambridge: Cambridge University Press.
- Bialystock, E. (1990). *Communication strategies: a psychological analysis of second language use*. Oxford: Blackwell.
- Breiter, D., Cargill, c., and Fried-Kline, S. (1995, Spring). An industry view of experiential learning. *FIU Hospitality Review*, 75-80.
- Douglas, D. (2000). *Assessing languages for specific purposes*. Cambridge: Cambridge University Press.
- Dresh, C., and Gregory, S. (1989). Educators must pursue industry relationships. *Nations restaurant news*. April 17, 1989.
- El-Koumy, A. S A. K. (2003). Language performance assessment: *Current Trends in Theory and Research*. Cairo: Dar An-Nahda Al-Arabya.
- Fisher, B. (1995). *Thinking and learning together: Curriculum and Community in a primary classroom*. Portsmouth, NH: Heinemann.
- Fletcher, S. (1995). Hospitality education. In R. Brymer (Ed.) *Hospitality management: An*

- introduction the industry*. Dubuque, IA: Kendall/Hunt Publishing.
- Folse, K. S. (2004). *Vocabulary myths*. Ann Arbor: Michigan University Press.
- Gibson, B. (1997). Talking the test: Using verbal report data in looking at the processing of cloze test. *ERIC Document No. ED 409713*.
- Harun, M. (1998). English at the hotel counter: A discourse analysis of hospitality language. *MA dissertation. University of Southampton*.
- Hong, J-S., & Lu, H-Y. (2006). Needs assessment of professional competencies of F&B/hospitality management students at college and university level. *Journal of Teaching in Travel and Tourism*, 6, 3, 1-26.
- Hutchinson, T., & Waters, A. (1987). *English for Specific Purposes: A Learning-centred approach*. Cambridge: Cambridge University Press.
- Kasavana, M. (1993). Front office operation. In M. A. Khan, M. D. Olsen, & T Var (Eds.), *VNR's encyclopedia of hospitality and tourism*. New York: Van Norstrand Reinhold.
- Lafrenz, L. N. (1991). Performance analysis: Hospitality industry employers' perceptions of their limited English proficient employees. *Doctoral dissertation: Ohio State University*. U.S.A.
- Lee, M-H., Lu, H-T., Jiao, Y-H., and Yeh, P-H. (2006). Research on correlations between off-school internship systems and work performances in hospitality and tourism education. *Journal of Teaching in Travel & Tourism*, 6, 69-87.
- Lewis, M. (1993). *The lexical approach*. Hove, UK: Language Teaching Publications.
- Richterich, R. (1983). *Case studies in identifying language needs*. Pergamon Press.
- Savignon, S, J. (1983). *Communicative Competence: Theory and Practice*. Reading. MA: Addison-Wesley.
- Tompkins, G. (2000). *Teaching Writing: Balancing Process and Product*. New Jersey: Merrill.
- Wagen, L.V. D. (2003). *Professional Hospitality: Core competencies, 2nd ed*. Australia: Hospitality Press.
- Wang, L-L. (2001). Professional competencies required for food and beverage employees working front of the house in international tourist hotels. *Doctoral dissertation: National Normal University*. Taipei, Taiwan.
- Wilkins, D. (1972). *Linguistics in language teaching*. London: Arnold.
- Wu, C. L. (2003). The study of job performance and job satisfaction of part-time employees in Taipei coffee chain stores. *Unpublished master's thesis. Chinese Culture University*, Taiwan.
- Wu, Y. (1998). What do tests of listening comprehension test? A retrospection study of EFL test-takers performing a multiple-choice task. *Language testing*, 15, 21-44.

Effective Teaching in Large Multilevel Classes: Beliefs and Practices of Experienced EFL Elementary School Teachers

Mei-hua Teng (鄧梅華), Yueh-kuei Hsu (許月貴)

National Taiwan Normal University

migo0430@yahoo.com.tw

ykhsu@ntnu.edu.tw

This study aimed to explore two experienced EFL elementary school teachers' beliefs and classroom practices in teaching large multilevel classes. Four research questions were addressed: (1) What are the teachers' beliefs toward teaching English in large multilevel classes? (2) What classroom practices are demonstrated that correspond to these beliefs? (3) What difficulties do the teachers encounter in the process of teaching, and how do these difficulties affect the teachers' beliefs and decision-making procedures? (4) What teaching strategies are identified by the teachers as examples of effective classroom practices which maximize students' learning in large multilevel classes? Various forms of data were collected to enhance the validity of the study namely, questionnaires, interviews, classroom observations, and the researchers' reflective journals. Findings showed that both teachers insisted on following the Nine-year Integrated Curriculum Guidelines and focused their teaching on the needs of the majority. Though students' varied English proficiency and other contextual factors hindered the teachers in practicing effective teaching, both teachers were still striving to conduct meaningful communicative activities such as pair work, group work, and multilevel homework. Results of this study highlight the importance of teachers' beliefs concerning their classroom practices. More specifically, successful implementation of effective teaching practices relies on the teachers' competence and perseverance in balancing their practices between benefiting the majority and facilitating the individual. Pedagogical implications for frontline teachers, parents, school administrators, teacher training programs, and educational authorities are discussed.

INTRODUCTION

As a result of globalization, English has established its steadfast status as the all-purpose, worldwide language (C. -Y. Chen, 2004; Li, 1998). Since English has been acknowledged as the main communication tool worldwide, EFL countries, such as South Korea, Thailand, Singapore, China, and Taiwan have long incorporated English teaching and learning into their formal education (Y. -M. Chiang, 2003; Li, 1998). Similar to other EFL countries, the government of Taiwan also aims to equip our young learners with the ability to communicate in English and to connect with the global village (Y. -M. Chiang, 2003).

According to the Nine-year Integrated Curriculum Guidelines, starting in 2005, all public elementary schools in Taiwan were to include English as a required curriculum from grade three (Ministry of Education, 2006). Nevertheless, a nationwide survey on the English teaching situation in Taiwan conducted by the CitiSuccess Fund in 2003 revealed that 80% of the elementary schools in Taiwan had already started their English programs ahead of the Curriculum Guidelines (Y. -C. Chen, 2004). This finding indicated that, despite the established English policy, most parents in Taiwan found it essential to develop children's

English proficiency at younger ages. As children in Taiwan begin to receive English education at different ages, it is inevitable that these young learners have varied English proficiency when they begin their English learning at school. When these children start to learn English, their teachers would encounter strenuous challenges of developing and implementing effective teaching to meet their diverse English learning experiences and proficiency (Liao, 2007).

Besides the difficulties in meeting students' various needs in large multilevel classes, studies on EFL teaching revealed that such classes caused fortifiable teaching dilemmas for teachers in several aspects, such as lesson planning (Huang, 2003), teaching strategies (Hsu, 2005), classroom management (Y. -M. Chiang, 2003), and learning assessment (Chen, 2003). Besides students' varied proficiency, limited class time and difficulties in classroom management also prohibited teachers from meeting individual student's learning needs (Chen, 2003). In response to the above teaching difficulties, researchers and educators in Taiwan have been active in developing various effective teaching practices or strategies as solutions for teaching large multilevel classes, such as Communicative Language Teaching (CLT) (Liao, 2003), cooperative learning (Liang, 2002), and flexible ability grouping (Y. -M. Chiang, 2003).

Notwithstanding the efforts made by the researchers and educators, successful implementation of effective teaching practices still relies on English teachers themselves. Teachers' beliefs, among other factors, play the crucial role in their decisions for classroom practices (Borg, 2003; Woods, 1996). Thus, to teach English effectively in large multilevel classes, frontline teachers play fundamental roles in implementing effective teaching practices in their classrooms.

Subsequent researches emphasized that teachers' beliefs and practices are context-sensitive, dynamic, and open for change (Golombek, 1998). As Borg (2003) noted, teachers' beliefs and classroom practices "are mutually informing, with contextual factors playing an important role in determining the extent to which teachers are able to implement instruction congruent with their cognitions" (p. 81). In addition to the effect of the contextual factors on teachers' beliefs, other researches showed that teachers will reconstruct and reform their beliefs and practices through teaching experiences (Richards, Tung & Ng, 1992). According to Feikes (1995), teachers are seen as active learners who constantly improve their teaching as they reflect upon their experiences.

In sum, different teachers have different preferences in classroom strategies which tend to be constructed and reconstructed in response to the various contextual needs of the teaching environment and to their reflective thinking (Borg, 2003). In Taiwan, teaching English to large multilevel classes has become one of the major contextual factors which challenge EFL teachers' beliefs in effective teaching and their decision-making in the classroom (Y. -Z. Chiang, 2003; Huang, 2003). The present study, thus, aimed to explore the experienced EFL elementary school teachers' beliefs and classroom practices in teaching large multilevel classes within different school contexts in Taiwan. More specifically, this study intends to address the following research questions:

1. What are the teachers' beliefs toward teaching English in large multilevel classes?
2. What classroom practices are demonstrated that correspond to these beliefs?
3. What difficulties do the teachers encounter in the process of teaching, and how do these difficulties affect the teachers' beliefs and decision-making procedures?
4. What teaching strategies are identified by the teachers as examples of effective classroom practices which maximize students' learning in large multilevel classes?

LITERATURE REVIEW

The issue of teaching large multilevel classes has long been acknowledged and widely discussed by teachers and educators globally (Hess, 2001; Tomlinson, 1999). As defined by

Balliro (1997), “any classroom is really multilevel, because we all bring a range of abilities and aptitudes with us to any learning situation” (p. 6).

Teaching Large Multilevel Classes in Elementary Schools

The Outline Protocol of the Nine-year Integrated Curriculum Guidelines, proposed by the Ministry of Education, listed three curriculum goals for English teaching and learning in the elementary school level (Ministry of Education, 2006): (1) Cultivating students with basic communication ability in English. (2) Arousing students’ interests in learning English. (3) Providing students with the knowledge of both domestic and foreign culture and customs.

Elementary schools are encouraged to organize and accommodate the school-based curriculum upon consideration of individual school conditions, parental expectations, and students’ needs (Ministry of Education, 2006). However, what seems to be a positive curriculum innovation becomes a general concern for English teaching in Taiwan. Research on the implementation of the Curriculum Guidelines showed that large multilevel classes have caused major obstacles for teachers when they try to enforce the new curriculum goals (Chen, 2003). EFL teachers in Taiwan stressed that they were challenged to enhance individual learning outcomes in the classes, especially when the students possess diverse English proficiency and learning priorities (Y. -Z. Chiang, 2003). Commonly reported difficulties teachers have encountered when teaching large multilevel classes were large class size, individual differences, limited class time, ineffective teaching strategies, and inadequate teaching materials. Besides, difficulty in classroom management and deficiencies in employing multiple classroom assessment methods also hindered teachers from implementing effective teaching (Hess, 2001). Apparently, the dilemma of teaching English to large multilevel classes, especially in the elementary schools, has been put in the spotlight (Chung & Hsu, 2005).

Significance of Teachers’ Beliefs on Their Practices

A considerable amount of researches supported that teachers’ beliefs play a crucial role in the decision-making of classroom practices (Bailey et al., 2001). Bailey et al. (2001) noted that teaching is very similar to the concept of “parenting” (p. 15). Parents tend to raise their children the way they were brought up. Likewise, teachers are likely to teach the way they were taught. Since teachers have different learning experiences, they will eventually come up with diverse teaching principles and develop a personalized set of preferred classroom practices (Bell, 2005). In addition to personal learning experiences, internal factors such as teachers’ beliefs and personalities, and external factors such as the education environment would also influence how teachers present their classroom practices (Smith, 2001). Woods (1996) stressed that experienced teachers are inclined to enhance professional development by constantly monitoring, elaborating, and developing their teaching strategies along with their experiences, which is why they tend to develop a more sufficient knowledge of teaching strategies and routines to implement in the classroom (Tsui, 2003).

Teaching large multilevel classes is, without a doubt, an extremely demanding and challenging task. In Taiwan, a number of studies conducted on EFL teachers’ beliefs and practices towards various instructional methods have revealed that contextual limitations such as large classes, individual differences, and limited class time have significant influence in English teachers’ classroom practices. For example, Liao (2003) conducted a study about two Taiwanese senior high school EFL teachers’ beliefs towards Communicative Language Teaching. The two teachers pointed out that exams and contextual factors such as large class size and the twin peak phenomenon caused setbacks in the teacher’s teaching process. Kuo (2005) investigated EFL junior high school teachers’ beliefs and attitudes towards normal class grouping. Findings showed that teachers held negative beliefs towards teaching large multilevel classes due to the difficulty in classroom management, the challenges encountered when planning curriculum instructions, and the predicament in assessing

students' learning achievements. Chung and Hsu's (2005) article on the EFL teaching situation in Taiwan concluded that teachers encountered difficulty in teaching large multilevel classes because most teachers preferred to use whole-class teaching strategies. According to the study, in whole-class teaching, students with different English proficiency had uneven opportunities to participate in classroom interaction. Students who responded to teachers' questions eagerly were usually those with better English proficiency. On the other hand, students with lower English proficiency could only play as onlookers during these teacher-student interactions, making them more afraid or even become indifferent towards learning English.

In the educational environment of Taiwan, besides difficulty in teaching large multilevel classes, the implementation of the Curriculum Guidelines has also led to high demands and expectations for EFL elementary school teachers. "How to teach in accordance with the Curriculum Guidelines and, at the same time, manage to meet the needs of every individual in the large multilevel class?" is still an unsolved problem. Thus, to gain a better understanding of the difficulties EFL teachers encounter, to promote their teaching effectiveness, and most importantly, to help students achieve ideal learning results, a further study is necessary.

METHODOLOGY

The present study aims to investigate experienced EFL elementary school teachers' beliefs and practices in teaching large multilevel classes. To enhance the validity of the study, different forms of data were collected using questionnaires, interviews and classroom observations.

Participants

Two experienced EFL elementary school teachers were invited to participate in this study, Ms. Chou and Ms. Yin. The two teachers were recommended by the researcher's advisor due to their excellent teaching and many years of teaching experience.

Ms. Chou. During the present study, Ms. Chou taught at an elementary school located in a city in the north of Taiwan. She has been an English teacher at the target school for over seven years. Ms. Chou's school was located in a metropolitan area in northern Taiwan. The school was well-known for its impetus to bilingual education. Starting from grade three, students had three English classes per week. Also, to promote students' English proficiency and conform to parental expectations, the school used English textbooks imported from the United States. For the present study, one second-grade class and one fifth-grade class were chosen by Ms. Chou as the target classes to be observed. The second-grade target class, Class A, was composed of thirty-five students. Ms. Chou said that compared with the second-graders from other schools, most of the students had intermediate, if not advanced, English proficiency in all language skills, especially in listening and speaking. Class B, the fifth-grade class, was composed of thirty-two students. Same as Class A, approximately 95% of the students have already learned English in cram schools.

Ms. Yin. Ms. Yin has been an English teacher for about ten years. It was her seventh year of teaching in the current school. Ms. Yin had been teaching the sixth-graders since she started working in the current school. This experience helped her become very familiar with the English proficiency of the six-graders in her school. Though she had been teaching the same textbook for over seven years, Ms. Yin remained active in developing innovative teaching practices and had employed various effective teaching strategies to present various teaching contents in the English textbook. The school Ms. Yin taught in was also located in a city in northern Taiwan. To promote student's English proficiency and to fulfill parental expectations, three versions of English textbooks were used for different grade levels. All

three versions were ESL teaching materials used in the United States, which were more advanced in level and were rich with American cultural contents. When the study was conducted, Ms. Yin taught English to nine sixth-grade classes. To help the researcher observe different teaching situations in large multilevel classes, Ms. Yin was generous in allowing the researcher to observe three six-grade classes, Class C, D, and E. According to Ms. Yin, the three classes chosen were different in their learning atmosphere, students' overall English proficiency, and their discipline in class.

Data Collection

The present study collected data through multiple sources to achieve a more holistic view of the teachers' beliefs and classroom practices in large multilevel classes (Yin, 2003). Four methods were adopted, including teacher questionnaire, teacher interviews and classroom observations.

Teacher Questionnaire. Before teacher interviews and classroom observations, the teacher participants were asked to complete a teacher questionnaire (see Appendix A). The purpose for adopting the teacher questionnaire is to gain the participants' background information (Gillham, 2000). The questionnaire was constructed with reference to the related studies on teacher's beliefs and teaching large multilevel classes¹. Participants' responses in the teacher questionnaire also served as lead-in topics for teacher interviews and classroom observations.

Teacher Interviews. To yield a more comprehensive view of the teachers' beliefs and practices, semi-structured interviews and post-observation interviews were employed in the study (Nien, 2002). Semi-structured interviews consisted of a set of open-ended questions formulated based on the literature related to teachers' beliefs and practices and about teaching English to large multilevel classes (Y. -Z. Chiang, 2003; Nien, 2002). The researcher structured and subsumed the semi-structured interviews into three sections: (1) teachers' beliefs and practices towards teaching English; (2) difficulties the teachers have encountered; and (3) the effective strategies the teachers have adopted when teaching large multilevel classes (see Appendix B). Six semi-structured interviews were conducted for each teacher. Interviews were scheduled at a time convenient to the teacher. The average length of the interviews was about 30 to 40 minutes. After each teaching session, if the teachers had free time, a 10 to 20 minutes post-observation interview was conducted to probe and clarify the participants' beliefs and practices towards the class just observed.

Classroom Observations. According to Yin (2003), video recorded evidence can capture details which might not have been noticed by the researcher during the teaching process. Classroom observations of this study were conducted from April of 2008 to June of 2008. Both teachers were observed and interviewed according to their convenient time. A total of six class periods for each teacher were observed, recorded, transcribed, translated, and analyzed.

Data Collection Procedures

Step One. Select and contact the participants. Before meeting the teachers, the researcher made telephone conversations with each teacher to introduce herself and to explain the focus of the study to the teachers, to gain a brief understanding of the teachers' background, and to schedule time for the interviews and observations.

Step Two. Complete the teacher questionnaire. Before conducting the teacher interviews and the classroom observations, the researcher explained the process of the study to the teachers. After the teachers understood the purpose of the study, a consent form (see Appendix D) was signed by the participants. Then, the participants were asked to complete

¹ The questionnaire is constructed with reference to Y. -C. Chen (2004), Y. H. Chen (2004), Hsu (2003), Huang (2003), Lin (2003), Liu (2004), and Nien (2002).

the teacher questionnaire (see Appendix A). After the questionnaires were completed, the researcher and the participants negotiated and scheduled suitable time for the follow-up interviews and classroom observations.

Step Three. Conduct teacher interviews and classroom observations. Teacher interviews were conducted according to the teachers' timetable. Questions for the interviews were provided to the teachers in advance to grant teachers with sufficient time to reflect on their teaching and prepare for the interviews. Classroom observations were video recorded. Each observation consisted of one instruction period; the researcher acted as a non-participant in the classroom. After each observation, if teachers had extra time, a 10 to 20 minutes post-observation interview was conducted to allow teachers to reflect upon the lessons.

Data Analysis

Miles and Huberman defined the process of data analysis into three synchronal activities: data reduction, data display, and conclusion drawing and verification (Miles & Huberman, 1994, p. 10-11).

Data reduction. Data reduction is the "process of selecting, focusing, simplifying, abstracting, and transforming" (Miles & Huberman, 1994, p. 10) the collected data. Based on data reduction, all the data collected in this study were analyzed and organized to clearly present the teachers' beliefs and practices in teaching large multilevel classes, including the teacher questionnaire, transcripts of interviews and classroom observations.

Data display. Yin (2003) noted that a clear and organized data display facilitates the researcher to better organize and focus the research conclusion drawn from the large amount of data.

Conclusion drawing and verification. According to Miles and Huberman (1994), qualitative data analysis is a continuous enterprise where various preliminary conclusions are drawn and revised before the final conclusions of the research are made. Faced with such a large amount of data collected, the researcher started to review, organize, and pre-analyze the collected evidence from the beginning of the data collection period.

RESULTS AND DISCUSSION

Teachers' Beliefs Toward Teaching English in Large Multilevel Classes

Ms. Chou and Ms. Yin both viewed elementary school education as compulsory education. Ms. Chou insisted that "elementary school education is regarded as compulsory education, not gifted education" (TIC², 20080415). Hence, when teaching large multilevel classes, both teachers believed that they should focus on instructing the majority before accommodating diverse students' needs.

As Tomlinson (1999) suggested, "Flexibility is the hallmark of a differentiated classroom" (p. 15). The findings of this study found that both teachers showed a great amount of flexibility in various teaching aspects such as teacher's role, lesson planning, teaching practices, and assessment approaches in the large multilevel class.

Ms. Chou believed that teachers should be flexible in adjusting their roles when instructing different grade levels.

When I [Ms. Chou] teach them [fifth graders] English or moral principles, I will take the role as a teacher. However, during activities, I wish to return to the role as a friend because they will be more willing to share their thoughts with me.... On the other hand, when teaching lower grade levels, I become more directed and teacher-centered because the kids are too young.... I need to remind them about the little things in daily life [such as monitoring their personal hygiene habits and classroom behaviors]. (TIC, 20080502)

² Data notations: TIC, teacher interview with Ms. Chou; TIY, teacher interview with Ms. Yin.

Ms. Yin stressed that “a different class is a different teaching context” (TIY, 20080424). In order to promote effective teaching, she would adjust her teaching styles and pace in accordance with the learning characteristics of the class.

Findings showed both Ms. Chou and Ms. Yin had no major complaints in implementing the Nine-year Integrated Curriculum Guidelines in their differentiated classrooms. In fact, both teachers have accepted large multilevel classes as a fixed factor and planned their practices in accordance with such teaching contexts. As noted by Tsui (2003), “For expert teachers, the context is very much an integral part of their teaching act” (p. 30). Hence, it can be suggested that after many years of teaching experiences, Ms. Chou and Ms. Yin have already developed a holistic and integral understanding of Taiwan’s education policies, the situation of large multilevel classes, and to what extent they are able to maximize their students’ learning.

Classroom Practices that Corresponded to Teachers’ Beliefs.

According to classroom observations, Ms. Chou and Ms. Yin have been able to include a great amount of flexibility in their lesson planning and teaching practices when instructing large multilevel classes. Besides Communicative Language Teaching (CLT), the two teachers used a mixture of teaching approaches in their practices. For example, when teaching a new sentence pattern, Ms. Chou believed that it was important to leave her slower learners a “silent period” before asking them to speak in front of the class. This practice echoed the aspects of the Natural Approach (NA) in emphasizing the “delay of oral production” (Brown, 2001, p. 33). On the other hand, Ms. Yin used a great deal of the Grammar Translation Method (GTM) in her teaching to help her students comprehend the meaning of the text contents.

The two teachers also balanced between teacher-centered and student-centered instructions in class. For example, if the advanced learners understood the grammar points or sentence patterns, Ms. Yin would encourage their students to take the role of a teacher to share what they know with the class.

(When teaching *A World of Food*, Ms. Yin asked her students the difference between ‘glass’ and ‘glasses.’ An advanced student explained the difference, playing the role as the teacher.)

T: Alright. Look at your books again. What is ‘glass’?

S: glass [in Chinese].

T: glass [in Chinese]. Good. Why can’t we translate it into glasses here [in Chinese]?

S: Because there are two of them in glasses [in Chinese].

T: Good. So, when we talk about glasses, we should add...[in Chinese]?

Ss: ‘es’

T: Why do we have to add ‘es’ in glasses? Because there are two of them. Good! (COYD³, 20080410)

Classroom observations showed that Ms. Chou and Ms. Yin tried to promote students’ involvement by asking extensive questions such as “Why?” and “What else?” and encourage students to contribute further responses in accordance with their proficiency levels. As noted by Tomlinson (2001), “teachers in healthy classrooms continuously invite their students to be part of the teaching” (p. 33).

Difficulties Teachers Encountered in the Process of Teaching

Ms. Chou and Ms. Yin reported several difficulties that have been indicated in previous studies such as large class sizes and limited class time (Y. -Z. Chiang, 2003; Wang, 2006). Besides, Ms. Chou and Ms. Yin also noted that and inadequate teaching materials, students’ individual differences in English competence, and difficulties in teaching different

³ COYD stands for Classroom Observation, Ms. Yin, Class D.

grade levels also hindered their teaching.

Both Ms. Chou and Ms. Yin reported that the textbooks imported from the United States were too challenging for their students to digest. Thus, to achieve effective learning, Ms. Chou planned her activities based on meaningful contexts such as story telling and role plays. On the other hand, Ms. Yin would adopt the practice of whole text translations. According to Ms. Yin, students were able to comprehend the context better through translations and were able to use the vocabulary correctly in sentence pattern practices.

Students' individual differences in English competence also constrained the teachers in conducting effective teaching. In order to promote effective teaching, both teachers made full use of the advanced students' English competence in their instructions. For example, Ms. Yin would ask the advanced students to help her translate the vocabulary and reading passages into Chinese. Ms. Chou would encourage the advanced learners to explain grammatical rules for the class.

Besides the influencing factors mentioned above, Ms. Chou also commented on the difficulties encountered when teaching different grade levels. According to Ms. Chou, large multilevel classes did not cause teaching difficulties in her second grade class because the students had little English background knowledge and were more enthusiastic in participating classroom activities. However, Ms. Chou noticed a clear-cut difference between the two extreme levels in her fifth grade class. Ms. Chou pointed out that some of the slower learners have already lost interest in learning. The same problem was also reported by Ms. Yin when she talked about some of her sixth-graders being indifferent towards classroom activities because of their poor English competence or lack of learning interest. The two teachers reported that it becomes harder for slower learners to regain confidence in English when they reach higher grade levels. This finding was correspondent with Y. H. Chen's (2004) statement that "the numbers of underachievers tended to increase with the students' grade levels" (p. 92). In response to such dilemma, Ms. Yin gave the same suggestions as Chen, which was the urgent need to practice immediate remedial support starting from the lower grade levels.

Teaching Strategies Identified by the Teachers as Effective Classroom Practices

Research in teaching effectiveness have suggested that, to promote effective teaching, teachers should learn to reflect on their own teaching and change their deeply-held beliefs according to their teaching experiences, teaching context, and students' characteristics, (Borg, 2003; Tsui, 2003).

Establish Class Rules and Teaching Routines. In large multilevel classes, it is difficult for teachers to care for each individual at all times. Ms. Yin noted that setting class rules helped her to maintain classroom order and conduct her lessons more smoothly. As noted by Ms. Yin,

Students would never learn well in a noisy classroom environment.... No matter how well I [Ms. Yin] teach, no matter how many supplementary materials I bring to the class, no matter how enthusiastic I am on the stage, if the students are chatting and playing throughout the lesson, everything would be useless. (TIY, 20080424)

Similarly, Ms. Chou noted the importance for students to take responsibility in their own learning. By setting up rules as routines, students would know what they are expected from the teacher, how they are supposed to behave in the classroom, and what they should do to achieve effective learning outcomes.

Carry Out Guided Practices. Ms. Chou's pointed out that most advanced learners could immediately respond to the teachers' questions without extra explanation from the teacher. On the other hand, the slower learners needed more time to process and digest the new information. Therefore, Ms. Chou practiced a three-step sequence when carrying out communicative activities in her large multilevel classes: teacher demonstration, advanced and

average students practice, and slower learner practice. Ms. Chou believed that after listening to the teacher's demonstration and observing how their classmates participate in the activities, the slower learners would feel less anxious and more confident in speaking out the language. Such strategy echoed to one of the main aspects of the Natural Approach (NA), which is to delay students' oral production time and provide students with sufficient comprehensible input before asking them for feedback (Brown, 2001).

Practice Heterogeneous and Homogeneous Group Work. According to Chung and Hsu (2005), cooperative learning enabled students to learn from one another, accept individualities, and work together as a group. In order to let students practice their language fluency as much as possible during the limited instruction time, Ms. Chou and Ms. Yin included various heterogeneous and homogeneous group works in their teaching. In fact, Ms. Chou and Ms. Yin suggested some key points when practicing group work. (1) Always give clear instructions before the activity. (2) Group work is only effective when classroom management is assured. (3) During group practices, teachers should walk to each group and assess the students' learning process. (4) Teachers can seize the opportunity to practice one-to-one remedial instruction for slower learners during group work. (5) Teachers should encourage advanced learners to conduct peer tutoring and assist slower learners in completing the task. (6) After group work activities, teachers should ask each group to share their learning and thoughts with the whole class.

Assign Multilevel Homework. Ms. Chou preformed in-class ability grouping in her fifth grade class and divided the students into three colors (red for advanced, blue for average, and green for slower learners) based on their English proficiency. Ms. Chou reported that, due to time limitation, she preferred whole-class teaching or group work during class hours, but she would always give her students multilevel homework assignments.

Now we [Ms. Chou, fifth-grade class] are learning comparatives. If average students are able to produce sentences like 'My mom is younger than my dad.' That's good enough. As for slower learners, if they can write out the words 'bigger, smaller,' I would be really grateful. For top students, I will ask them to write a composition. They have to make three to four sentences and compose a paragraph. They can add words besides the ones listed in the textbook. For example, we've only learned 'bigger, smaller, faster, shorter' in the class, but they are encouraged to use other words they know, like 'cooler.' (TIC, 20080415)

According to Ms. Chou, students had to complete their work in compliance with their ability to receive the same grades. Ms. Chou's teaching strategy harmonized with Tomlinson's suggestion that "Goals of a differentiated classroom are maximum growth and individual success" (p. 15). In large multilevel classes, teachers should provide differentiated activities to enable students to learn at their own pace and to work in accordance with their different proficiency levels (Corley, 2005).

CONCLUSION AND IMPLICATIONS

This study aimed to investigate two experienced EFL elementary school teachers' beliefs and classroom practices in teaching large multilevel classes in hope to promote teaching effectiveness in the elementary school level. Overall speaking, Ms. Chou and Ms. Yin's teaching practices were generally consistent with their beliefs. In terms of the teacher's beliefs toward teaching English to large multilevel classes, both teachers believed that they should teach according to the Curriculum Guidelines and focus their teaching target towards the majority of the class. Moreover, both teachers firmly believed that teachers should demonstrate flexibly in their instructions such as pair work, group work, and multilevel homework assignments.

As for the difficulties in instructing large multilevel classes, Ms. Chou and Ms. Yin's teaching effectiveness were greatly affected by contextual variables such as large class sizes,

limited class time, student's individual differences, inadequate teaching materials, and students' individual differences. Both teachers noted that if the textbooks were appropriate for students' level, they would have more time to design extended activities for the average and advanced learners and provide one-to-one remedial instructions for the slower learners.

In conclusion, results of this study highlighted the importance of teacher's beliefs on their classroom practices. In order to practice effective teaching in multilevel classes, teachers have to adjust their attitudes and respect the fact that large multilevel classes are a natural phenomenon. It is believed that, if teachers conduct effective teaching solutions to minimize the gap between students' English competence, their students would eventually benefit from their practices and achieve optimal learning outcomes.

Pedagogical Implications

The findings of the present study offer several pedagogical implications respectfully made for the frontline teachers, students' parents, the school administrators, teacher training programs, and educational authorities.

First, it is suggested that the frontline teachers need to become more aware of how their beliefs influence their teaching, and how their classroom practices eventually affect their student's learning outcomes. Teachers should learn to reflect on their teaching and change their deeply-held beliefs according to their teaching experiences, teaching context, and students' characteristics in order to minimize the gap between teachers' beliefs and their practices. Also, Ms. Chou stressed that teachers should maintain a communicative network with students' homeroom teachers and their parents in order to keep track of students' learning progress and offer appropriate practices to meet each student's needs.

Second, as mentioned by Ms. Yin, teachers, parents, and students form an effective learning triangle, each member has to contribute in the learning process in order to maximize their learning. More specifically, teachers should try their best to conduct effective teaching practices even within the limited amount of class time; students should pay attention and enthusiastically participate in classroom activities; and parents should help their children review what they have learned at school. In this way, it would be assured that students would gain solid English competence.

Third, school administrators are suggested to provide a mutual space and time for English teachers to exchange their thoughts and problems in the process of instruction. Thus, teachers could seek practical and effective instruction strategies together to overcome the difficulties in teaching. Also, school administrators should constantly listen to teachers' suggestions and try their utmost to provide necessary assistance for teachers to practice effective classroom practices. Besides the schools, teacher training programs should address to the teaching dilemmas caused by large multilevel classes and provide more in-service training programs, seminars, and workshops related to such issues.

Finally, if teachers, parents, school administrators, and educational authorities could all contribute to meet students' learning needs, it is hoped that large multilevel classes would no longer be interpreted as an insolvable teaching dilemma for EFL teachers in Taiwan.

REFERENCES

- Bailey, K. M., Bergthold, B., Braunstein, B., Jagodzinski Fleischman, N., Holbrook, M. P., Tuman, J., Waissbluth, X., & Zambo, L. J. (2001). The language learner's autobiography: Examining the "apprenticeship of observation." In D. Freeman, & J. C. Richards (3rd Eds.), *Teacher Learning in Language Teaching* (pp. 11-29). New York: Cambridge University Press.
- Balliro, L. (1997). *Introduction: The multilevel question*. (ERIC Document Reproduction Service No. ED 411458).
- Bell, T. R. (2005). Behaviors and attitudes of effective foreign language teachers: Results of

- a questionnaire study. *Foreign Language Annals*, 38(2), 259-269.
- Borg, S. (1999). Studying teacher cognition in second language grammar teaching. *System*, 27, 19-31.
- Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching*, 36, 81-109.
- Brown, D. (2001). *Teaching by principles: An interactive approach to language pedagogy* (2nd ed.). White Plains, New York: Longman.
- Chen, C. -Y. (陳佳韻) (2004). <台中市國小學童英語學習態度之調查研究。> [A Study of English Learning Attitudes of Elementary School Students in Taichung City]. Unpublished master's thesis, National Taichung University, Taichung, Taiwan.
- Chen, H-m. (陳秀美) (2003). A study of primary school English teachers' beliefs and practices in multiple assessments: A case study in Taipei city. Unpublished master's thesis, National Taipei University of Education, Taipei, Taiwan.
- Chen, Y. -C. (陳艷秋) (2004). <國民小學實施校本英語教學之個案研究。> [A Case Study of the School-Based Curriculum Development of English Instruction in Elementary School]. Unpublished master's thesis, National Taitung University, Taitung, Taiwan.
- Chen, Y. H. (2004). Elementary and junior high school English teachers' perceptions and implementation of remedial instruction for underachievers. Unpublished master's thesis, National Taiwan Normal University, Taipei, Taiwan.
- Chiang, Y. -M. (蔣月美) (2003). <彈性能力分班在國小英語教學之實施：北市龍安國小個案研究> [The implementation of flexible ability-grouping in English teaching in elementary schools: A case study of Long An Elementary School in Taipei]. Unpublished master's thesis, National Taipei University of Education, Taipei, Taiwan.
- Chiang, Y. -Z. (江吟梓) (2003). A case study of English classroom management in relation to children's individual proficiency difference issues in Taipei primary schools. Unpublished master's thesis, National Tsing Hua University, Hsinchu, Taiwan.
- Chung, S. Y., & Hsu, S. Y. (鐘樹掾、許淑燕) (2005). 當前英語教學的迷思語因應對策—小組合作與資訊科技的應用。教師之友, 46(3), 57-64。
- Corley, M. A. (2005, March). Differentiation Instruction: Adjusting to the needs of all learners. *Focus on Basics*, 7(C). Retrieved November 10, 2007, from www.ncsall.net/index.php?id=444
- Feikes, D. (1995). *One teacher's learning: A case study of an elementary school teacher's beliefs and practice*. (ERIC Document Reproduction Service No. ED 389605).
- Gillham, B. (2000). *Case study research methods*. London, UK: CONTINUUM
- Golombek, P. R. (1998). A study of language teachers' personal practical knowledge. *TESOL Quarterly*, 32, 447-464.
- Hess, N. (2001). *Teaching large multilevel classes*. Cambridge, UK: Cambridge University Press.
- Hsieh, H-J. (2002). Teachers' beliefs about EFL learning: A study of elementary school English teachers in Taipei County. Unpublished master's thesis, National Taipei Teachers' College, Taipei, Taiwan.
- Hsu, M. Y. (許銘裕) (2005). 經營一個英語特色班級。師友月刊, 467, 88-91。
- Hsu, Y-k. (2003). Researching multilevel assessment: Evidence from EFL elementary school classrooms. *English Teaching & Learning*, 28(1), 29-60.
- Huang, Y-W. (2003). The influence of different grouping methods on junior high school students' oral performance in mixed-ability Classroom. Unpublished master's thesis, TamKang University, Taipei, Taiwan.
- Kuo, Y-T. (郭原再) (2005). <國民中學教師對常態編班的支持度及其信念與態度之相關研究。> [A study of the relationship of the support of junior high school teachers toward

- normal class grouping with teachers' belief and teachers' attitude]. Unpublished master's thesis, National Cheng Kung University, Tainan, Taiwan.
- Li, D. (1998). "It's always more difficult than you plan and imagine": Teachers' perceived difficulties in introducing CLT in South Korea. *TESOL Quarterly*, 32(4), 677-703.
- Liang, T-l. (2002). Implementing cooperative learning in EFL teaching: Process and effects. Unpublished master's thesis, National Taiwan Normal University, Taipei, Taiwan.
- Liao, P. (2007). Teachers' beliefs about teaching English to elementary school children. *English Teaching & Learning*, 31 (1), 43-76.
- Liao, W-w. (2003). Senior high school English teachers' beliefs towards communicative language teaching and their classroom practice. Unpublished master's thesis, National Taiwan Normal University, Taipei, Taiwan.
- Liu, S. -F. (劉淑芬) (2004). <國中常態編班的教學困境與教學策略之研究—以八位教學優良教師的成功經驗為例。> [Teaching difficulty and effective teaching strategies in junior high school normal class grouping: A case study of eight effective teachers]. Unpublished master's thesis, National Chengchi University, Taipei, Taiwan.
- Ministry of Education. (教育部) (2006). <九年一貫課程> [The Nine-year Curriculum Guidelines]. Retrieved 2006, January 26, 2006, from <http://teach.eje.edu.tw/9CC/index.php>
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd Ed.). Newbury Park, California: Sage.
- Nien, Y. -H. (2002). Teacher beliefs and their influence on classroom practices: A case study of a senior high school English teacher. Unpublished master's thesis, National Taiwan Normal University, Taipei, Taiwan.
- Richards, J. C. (1998). *Beyond training: Perspectives on language teacher education*. New York: Cambridge University Press.
- Richards, J. C., Tung, P., & Ng, P. (1992). The culture of the English language teacher: A Hong Kong example. *RELC Journal*, 23(1), 81-103.
- Smith, D. B. (2001) Teacher decision making in the adult ESL learning classroom. In D. Freeman, & J. C. Richards (3rd Eds.), *Teacher Learning in Language Teaching* (pp. 197-216). New York: Cambridge University Press.
- Tsui, A. B. M. (2003). *Understanding expertise in teaching: Case studies of second language teachers*. Cambridge, U. K.: Cambridge University Press.
- Tomlinson, C. A. (1999). *The differentiated classroom: Responding to the needs of all learners*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Tomlinson, C. A. (2003). *Fulfilling the promise of the differentiated classroom: Strategies and tools for expositive teaching*. Alexandria, VA, USA: Association for Supervision and Curriculum Development.
- Wang, P-H. (王伯雄) (2006). <以有效教學策略促進課程銜接之個案研究—以七年級英語科為例。> [Utilizing effective teaching strategies to promote curriculum articulation: A case study of 7th grade English teachers]. Unpublished master's thesis, National Taiwan Normal University, Taipei, Taiwan.
- Woods, D. (1996). *Teacher cognition in language teaching: Beliefs, decision-making and classroom practice*. New York: Cambridge University Press.
- Yin, R. K. (2003). *Case study research: Design and methods*. (3rd ed.) London, U.K.: Sage.

Examining Metacognitive Performance Between Skilled and Unskilled Writers in an Integrated EFL Writing Class

Han-min Tsai (蔡漢民)
Chungyu Institute of Technology
hmntsai@yahoo.com.tw

There has been an increasing call for taking an integrated position and strengthening learners' metacognitive models in EFL writing instruction. However, little research has been conducted to try to examine EFL students' metacognitive performance in an integrated writing setting. This study attempted to examine skilled and unskilled EFL writers' metacognitive performance in such a setting. The data were collected from the students' self-monitoring annotations and reflective journal entries and analyzed qualitatively. The results show that the unskilled writers' main concern in their writing process was with low order aspects and their experiences in strategy use were less selective and less productive. In contrast, the skilled writers were able to focus most of their attention on high order concerns, more concerned about how to make their texts impressive to readers, and more flexible, self-conscious and effective in strategy use. This study further suggests that scaffolded instruction and activities be constantly provided to promote EFL writers' metacognitive awareness, especially the less proficient ones.

INTRODUCTION

Writing is a complex and holistic enterprise, involving not only cognitive and affective but also social and metacognitive domains. This complexity becomes especially apparent in learning and teaching how to write in EFL contexts since EFL learners, as Kamimura (2000) observes, are from cultural and linguistic backgrounds where discourse conventions and audience expectations are different from those of the target language community. Adopting an integrated approach, according to Badger and White (2000), Kamimura (2000) and Xudong (2005), thus can better suit learners' needs in various aspects and be more effective in writing instruction.

Taking an integrative approach enables the teacher to play multiple roles not only as a knowledge provider but also as a consultant, facilitator and sensitive responder. The learners are not passive receptacles of knowledge but active participants, assuming certain responsibilities and autonomy to act as readers, collaborators and reflectors. Reports by Blanton (1987) and Crandall (1999) have claimed that making learning collaborative can reduce anxiety in EFL language classrooms. Cotterall's (2000) study indicates that fostering learner autonomy helps to develop learners' language proficiency, which in turn will increase their writing confidence.

Taking an integrated approach also means that writing in EFL classes is a structured and gradual process. The teacher first provides appropriate input based on the target context and the understanding of learners. Writers like Kamimura (2000) and Raimes (1985) have emphasized the importance of instruction on both content and form in EFL writing classes because, as Grabe and Kaplan (1996: 142) state, L2 students do not have the "implicit knowledge of rhetorical plans, organizational logic, and genre form" in the target language. After this, the students engage in multiple and recursive creation of texts by getting feedback from various sources. The first draft is not the final product and writing is not about getting instant perfection. De-emphasizing perfection, in line with Leki's (1999) stand, can help

reduce writing apprehension. Breaking down the writing process into manageable pieces and providing appropriate feedback at different steps can make learners aware of their own progress (Blanton, 1987) and release them from the pressure of time (Tsui, 1996). Errors in such an approach are treated hierarchically; global aspects such as idea organization, clarity and coherence are emphasized first and local errors such as those in grammar and mechanics are delayed until the editing stage. Doing this can prevent learners from premature editing, one of the factors Rose (1984) attributes writing block to.

To summarize, taking an integrated approach mostly concurs with the five criteria for effective scaffolding in writing advocated by Applebee (1986): student ownership of the learning event, appropriateness of the instructional task, a structured learning environment, shared responsibility, and transfer of control.

Emphasis on the process writing in a synthetic approach can further facilitate and promote the EFL instructor's and learners' metacognitive awareness. The instructor can better understand the composing process learners undergo and the strategies and skills they adopt and develop to deal with writing problems and can therefore help learners to perform their writing tasks effectively. Learners are given opportunities to explore their writing tasks, reflect on their cognitive enterprise, monitor and revise their written work, assess the effectiveness of the strategies used, and take regulatory strategies to deal with the problems emerging, thus making writing more natural and effective.

Metacognition is defined as an awareness of one's own cognitive processing and an ability to manage and regulate one's strategies appropriately to meet different situations (Williams & Burden, 2001). Flavell (1979) divides metacognition into metacognitive knowledge and metacognitive experiences. The former consists of personal, task and strategy variables, while the latter refers to any kind of conscious experience involved in cognitive enterprise. In Baker and Brown's (1984) notion of metacognition, not only knowledge but also regulatory activities, such as planning, monitoring, revising and evaluating, are emphasized. More recently, Schraw (2001) also defines metacognition in terms of knowledge and regulation of cognition. Anderson's (2001) model of metacognition is strategy-oriented, including the components of preparing and planning for effective learning, selecting and using particular strategies, monitoring strategy use, orchestrating various strategies, and evaluating strategy use and learning. In a nutshell, metacognition includes not only knowledge of cognition but also a procedural experience to plan, monitor, revise, evaluate and regulate one's cognitive activities, and all of these stages can be interactive and recursive.

To date, most studies on metacognition focus on its correlation with language proficiency in general rather than with specific language aspects. Writers like Wenden (2001), however, have claimed a need to explore the relationship between metacognition and specific language tasks. Compared with the availability and maturity of studies investigating metacognition in FL reading (Mokhtari & Reichard, 2002), metacognition research into FL writing is still very limited. Nevertheless, the role of metacognition in FL writing is drawing growing attention as its advantages and importance keep standing out. For example, research by Raphael, Kirschnet, and Englert (1986) shows that enhancing students' metacognitive awareness facilitates their performance in FL writing. In recent empirical studies, writers like Kasper (1997), Victori (1999) and Angelova (2001) have verified a significant positive correlation between metacognition and ESL writing performance. Sasaki's (2000) investigation into EFL learners' writing processes indicates that differences in strategy use between EFL expert and novice student writers can partly explain their written proficiency. Metacognition is also viewed as an essential discriminator between skilled and unskilled EFL writers (Omaggio, 1986; You, 2003). Magno's (2008) research further demonstrates that knowledge of cognition and regulation of cognition can be used as significant predictors of English written proficiency.

Based on the belief that an integrated approach can better suit EFL student writers' needs and that metacognition plays an important role in writing process, the present study adopted a process-product approach in a college-level EFL writing class in Taiwan and aimed to examine metacognitive differences, in terms of self-monitoring, self-assessment and self-management, between skilled and unskilled student writers in such a class.

METHOD

Subjects

33 students were involved in this study. They were all English majors at a private two-year institute of technology in Northern Taiwan and were in the second semester of the course English Composition when this study was undertaken. In the first semester the class was mostly product-oriented because about one-third of the students did not have any experience in writing an English composition. The students ranged in age from 21 to 28, but the majority (78.8%) were in their early twenties. Their average length of previous English study was 7 years. In addition, 30.3% of the participants had short-term overseas study experience in English-speaking countries. The current study spanned 18 weeks and three modes of written discourse, *viz.* narrative, expository and argumentative, were involved.

An Instructional Model for the Present Study

Based on the beliefs of an integrative writing approach, an instructional model was devised for the current study. Four key components were included: *input*, *writing process*, *evaluation*, and *environment*. The process and interaction of these components are illustrated diagrammatically in Figure 1. The use of dashes is to indicate that based on the problems detected, it may be necessary to revert to the former stage(s).

Depending on the need of learners, different kinds of input will be provided, including knowledge of the writing task, knowledge of global and local aspects, and training and demonstration of the skills like planning, peer revision, as well as self-monitoring annotations. Such input is conducted in terms of teacher instruction, model text analysis and group discussion. In the writing process, this model adopts the three sub-components in Flower and Hayes' (1981) cognitive process theory of writing, namely, planning, translating, and reviewing. Planning entails the operations of goal setting, idea exploration, and idea organization. Translating means formulating texts according to what has been attained in planning and reviewing stages. Reviewing is aimed at improving the text produced so far from different sources of feedback, including self-monitoring annotations after the completion of the first draft, peer revision prior to writing draft 2, and written teacher comments after draft 2. These sub-components are interactive and recursive throughout composing.

Another component in this model is evaluation. In order to reduce the students' writing anxiety, assessment is formative rather than summative. Different dimensions in the writing process, including writing performance, activity participation and metacognitive tasks, such as outlining, peer revision, self-monitoring annotations and journal writing, are counted for evaluation. In addition to these components, the environment in which writing takes place also plays an important part in this model. This includes the redefined role of the teacher as facilitator and knowledge provider and students as active, responsible and reflective participants; dynamic syllabi, *viz.* keeping refining the designed syllabi in responding to the students' reflection and writing performance; and the interactive and collaborative climate. Such an environment encourages the teacher and students to reflect on the created environment and regulate themselves, thus making a writing class more effective and less threatening.

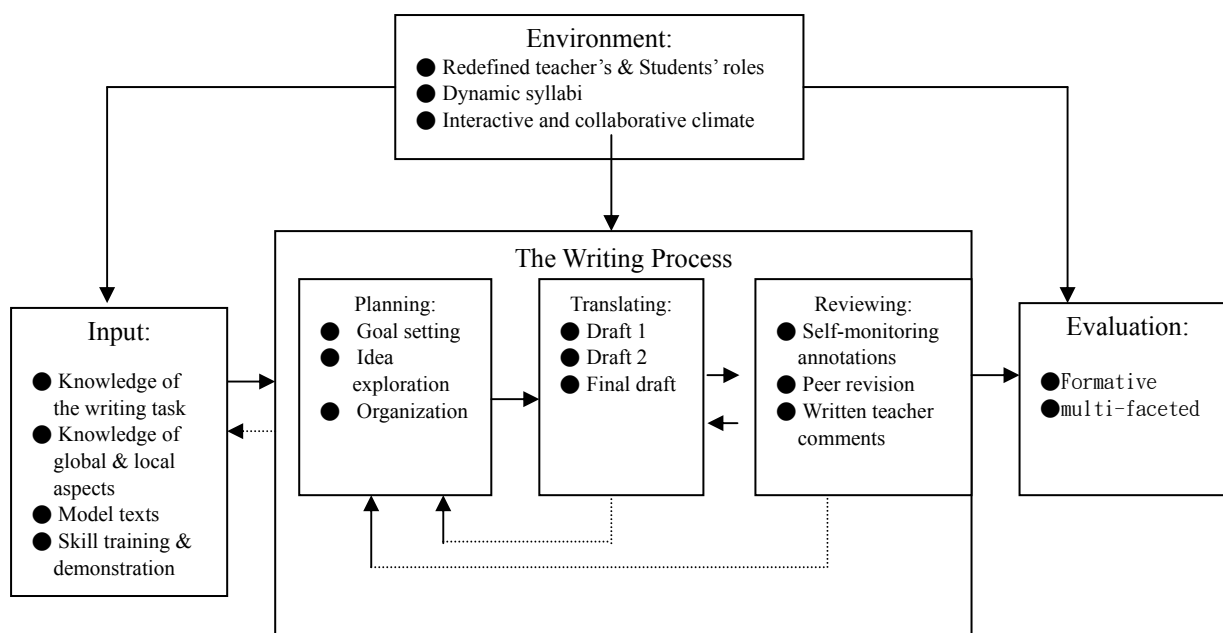


Figure 1 An Instructional Model for the Present Study

Data Collection and Analysis

The means of data collection included the students' self-monitoring annotations and reflective journal entries. The students' self-monitoring annotations were based on Charles' (1990) procedure of self-monitoring; the students note down any queries or difficulties of their first draft that they want their teacher to help with, and then the teacher replies to the questions raised. The students kept a journal once a week to reflect on and evaluate the teaching and learning. The students' final journal entries served to elicit data to examine their progress made in the semester, problems still hanging on, and strategy use in dealing with a writing problem. For data analysis, the students were classified into skilled and unskilled groups. The former refer to the students ranking in the top one-third of the final course grades, and the latter to those ranking in the bottom one-third.

The data from the students' final journal entries and self-monitoring annotations were analyzed qualitatively, following the procedures from description to analysis and then to interpretation. Statistics on frequency were used where appropriate. The data collected were read and re-read four times. The first reading attempted to look at the overall character of the corpus of data. In the second reading, themes in each data source were sorted out. The third reading was aimed at rechecking the character of each theme and labeling the theme. In the fourth reading, convincing quotations to substantiate the themes attained were selected and coded. The data collected from the students' self-monitoring annotations were further categorized in terms of high order and low order concerns (HOCs and LOCs). The former refer to global aspects, such as content, organization and coherence, and the latter to features like grammar, word usage, and spelling.

RESULTS

The Students' Self-Monitoring Performance. The students performed self-monitoring annotations twice, one for the first draft of expository writing, the other for persuasive writing. There was one-month lapse and to improve the students' performance, more instruction and samples, focusing on high order concerns, were provided for the second annotation. The results of these two annotations show that in the first annotation the students' primary focus was on LOCs (93%); only 7% pertained to HOCs. However, after appropriate training, their focus on LOCs was reduced to 57.8%, whereas 42.2% of the focus was devoted to high order concerns such as global content, organization, coherence and unity.

To investigate differences between the skilled and unskilled groups in this aspect, data from the students' second self-monitoring annotations were analyzed in terms of sub-categories belonging to HOCs and LOCs. The results (see Table 1) indicate that after appropriate instruction and demonstration on global aspects, the skilled subjects attended to HOCs (77.3%) more than to LOCs (22.7%). Their main foci were on global content, unity and organization. Some representative examples from the students' annotations are:

“In my conclusion (paragraph 5), I did not summary the main idea very clearly. Can you give me some suggestions?”

“Does my first sentence of the first paragraph relate to the topic? Do I need to add more to my 3rd paragraph?”

“Lines 33-36, I tried to solve the problem of smoking? Do you think these sentences support my main ideas appropriately? Is there any unnecessary sentence or word?”

“Should I combine the second and third paragraph into one paragraph because my less details in the third paragraph.”

“On line 28, I am not quite sure if the transition ‘therefore’ is appropriate.”

However, what concerned the unskilled subjects remained in LOCs, in which a significant proportion centered on asking for direct translations from Chinese to English (42.9%) and word use (23.8%). They did not manifest any awareness in organization, coherence or unity. The following representative annotations illustrate this:

Table 1. *Percentage/Number of the Skilled and Unskilled Groups in the Second Self-Monitoring Annotations*

Category	Skilled (AN =22)	Unskilled (AN = 21)
HOCs		
1. Global content	6 (27.3%)	1 (4.8%)
2. Organization	4 (18.2%)	0 (0%)
3. Topic sentences	0 (0%)	3 (14.3%)
4. Coherence	2 (9.1%)	0 (0%)
5. Unity	5 (22.7%)	0 (0%)
Total	17 (77.3%)	4 (19%)
LOCs		
1. Grammar	1 (4.5%)	3 (14.3%)
2. Word Use	2 (9.1%)	5 (23.8%)
3. Translation	2 (9.1%)	9 (42.9%)
Total	5 (22.7%)	17 (81%)

Note. AN refers to the number of annotations each group presented.

“How to say ‘不孕症’ in English?”

“Can you translate ‘不抽煙者’ in English?”

“Is it OK to use ‘skin aging’ in line 17?”

“ash of cigarettes 的 ash 可以加 es 嗎?”

The Students' Self-Assessment Performance. Data for measuring self-assessment were collected from the students' final reflective journal entries. To observe detailed differences between the skilled and unskilled writers, the journal entries of these two groups were read many times and categorized. In relation to the progress after this study, Table 2 indicates that more than half of the skilled (72.7%) and unskilled (54.5%) subjects felt less anxious about English writing. Around half of the skilled and unskilled groups (63.6% vs. 45.5%) felt clearer about how to organize an essay. In addition, 72.7% of the skilled writers expressed that they could write more content and were able to adopt strategies and skills in their composing process, whereas only 36.4% and 27.3% of the unskilled ones felt so, respectively. Also, 54.5% of the skilled group felt that they could perform critical thinking

through journals and write different modes of written discourse; however, only 18.2% of the unskilled ones were able to perform the former, and 27.3% the latter. Finally, a considerably higher percentage of the skilled writers (45.5%) reported their ability in self-monitoring writing problems than their counterparts (9.1%).

Table 2. *Percentage/Number of the Skilled and Unskilled Groups in Self-Assessment*

		Skilled (N = 11)	Unskilled (N = 11)
Progress	1. Less anxious about English writing	8 (72.7%)	6 (54.5%)
	2. Able to write more content independently	8 (72.7%)	4 (36.4%)
	3. More able to take strategies and skills to deal with a writing topic	8 (72.7%)	3 (27.3%)
	4. Clearer about how to organize an essay	7 (63.6%)	5 (45.5%)
	5. Able to perform critical thinking through journals	6 (54.5%)	2 (18.2%)
	6. Able to write different modes	6 (54.5%)	3 (27.3%)
	7. More able to self-monitor writing problems	5 (45.5%)	1 (9.1%)
Existing Problems	1. Unable to revise one's or peers' work	1 (9.1%)	6 (54.5%)
	2. Still having difficulty in grammar and word usage	5 (45.5%)	7 (63.6%)
	3. Hard to express ideas in English	0 (0%)	4 (36.4%)
	4. Unable to present impressive ideas and details	2 (18.2%)	0 (0%)

With respect to the existing problems after the study, some of the unskilled writers still had trouble revising their own and peers' written work (45.5%) and expressing their ideas in English (36.4%). 18.2% of the skilled group reported that they were unable to present impressive ideas and details in writing. A considerably high percentage of the skilled (45.5%) and unskilled (63.6%) subjects expressed their difficulty in grammar and word usage; however, what mattered to the skilled writers was how to use effective words or sentence patterns to make their writing impressive to the readers. The following observation indicates such a desire:

"I still have some problem on writing. Even though I improve my writing skills, sometimes I can't find right pattern to describe my opinion. Also, word choice is another difficult part for me. I don't know whether my patterns and words can make my composition interesting."

The unskilled writers, however, simply worried about whether the grammar or words used were correct or not. An example of such is, "The problems I still have in English writing is I often use wrong words and grammar."

Furthermore, the ways how these two groups described their self-assessment in the journal entries were examined and analyzed. This brought about more intriguing differences between the two groups. For example, in describing progress in making a writing plan, the unskilled writers simply pointed out the advantages obtained from such a task. A representative student response is, "Now I would write outlining before writing. Writing outlining can help me have more ideas." In contrast, the skilled writers were able to make assessments more convincing by providing psychological variation before and after the study. This is evidenced in the following response:

"At the beginning of the writing class, teacher asked us to do lots of practice which we have never done before. I could not accept and get used to the change at first, and I did not think it will make any improvement. In fact, I was wrong. After doing practice, I gradually know how to make a writing plan and clearly write down a composition. I think I made an improvement on drawing a writing plan."

Another point worth mentioning is that after the study some skilled writers, in contrast to the unskilled ones, expressed their progress in a way that reflects the awareness of themselves as writers. For example, one skilled subject wrote this, “I think in these four months I often reflected myself and observed more carefully about things happened in the writing class. Now I feel that I know what I am doing when I [am] making an essay.”

The Students’ Self-Management Performance. Self-management is thought of as one type of metacognition. It refers to the ability, based on experiences, to manage one’s further cognitive development (Brown & Palinscar, 1982, discussed in Rivers, 2001). The present study examined the students’ self-management by asking the students in the final journal how they dealt with a writing problem when it occurred to them. The data were analyzed in terms of the skilled and unskilled writers by employing Oxford’s (1989) taxonomy of language learning strategies. Five categories were identified: social, metacognitive, affective, cognitive and compensation strategies. An additional category encompassing more than two types of strategy was recognized and labeled as *multiple strategies*. As presented in Figure 2, social strategies, mainly referring to discussing with peers and asking for correction from the teacher, were frequently used by the skilled and unskilled subjects (81.8% vs. 63.6%). However, a considerably higher percentage of the skilled writers (81.8%) claimed to employ metacognitive strategies than their counterparts (18.2%). Some of the unskilled writers would set objectives to reduce their errors, but the objectives were vague, for example, “I would study more vocabularies and grammar to solve my problems.” The skilled writers, however, would engage in more varied strategies on this level, including revising the essay, paying more attention to the task, or seeking practice opportunities. They were also more likely to take immediate actions to tackle the problem. An example is, “After keeping calm and paying more attention to the problem, I often can understand why I couldn’t write this well. Then I would try my best to improve the problem.”

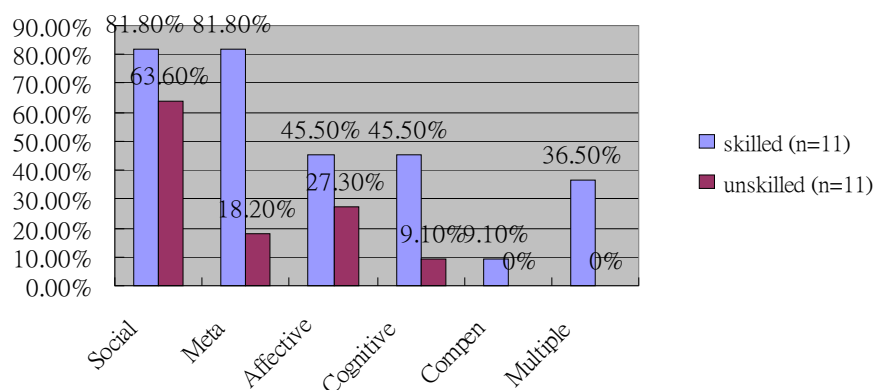


Figure 2 Results of the Strategies Taken to Writing Problems

With respect to the affective level, some of the skilled and unskilled students used relaxation and left their work for a while (45.5% vs. 27.3%). The skilled subjects stated that during their relaxation some wonderful ideas often occurred to them. As one wrote:

“Before writing, I often experienced a period of anxiety, so I left my work and watched TV, listened to radio or did exercise. But strangely enough, this was also the time a lot of ideas kept coming out. So when I sat down again to write, I usually had some ideas in my mind.”

The unskilled writers, nevertheless, seldom experienced such a productive journey. Their mind blankness and worries remained even after relaxation.

At the cognitive level, 45.4% of the skilled writers would consult the dictionary, refer to the model expressions and structures presented in class, or make use of outside class resources. When one skilled subject described how she dealt with her difficulty in persuasive writing, her comments read, "Writing persuasive work is kind of difficult to me. So I had to spend a lot of time collecting information from reading newspapers or articles." However, only one unskilled writer was involved in this level, and the strategy pointed out was resorting to the dictionary. The compensation strategies were least employed; only one skilled writer mentioned that she would switch to another word or pattern with a similar meaning when faced with a writing problem.

As for the multiple strategies, the data show that the unskilled writers did not have any awareness on this level, but 36.4% of the skilled writers would resort to this level. Their approaches might go from metacognitive to social levels, for instance, "I will review my essay first, then trying to find out the correct answer. If I can't find out the answer by myself, I will discuss with classmate. My last choice is discussing with teacher." Also, one skilled subject reported that based on the problems emerging, she would take different strategies. Her response is, "If my words do not support my main ideas, I will read my writing repeatedly. If my English is non-English, I will ask my classmates or teacher. If I have grammar problems, I will check grammar books."

DISCUSSION

In the present study, incorporating elements of a process writing approach into the traditional product approach provided the students with the opportunities to engage in metacognitive activities, such as planning a writing task, monitoring different drafts, taking strategies to cope with problems and keeping journals to assess one's learning as well as the class progress. Being able to think about one's thinking and learning process, according to Anderson (2001), can lead to more effective learning. De-emphasizing perfection and breaking down the writing process into manageable pieces, as Leki's (1999) and Tsui (1996) note, respectively, can help reduce writing apprehension. The learners' self-assessed improvement in writing anxiety, to a certain degree, was attributed to the implementation of the process approach.

The qualitative analyses of the students' self-monitoring annotations and reflective journal entries came up with some intriguing distinctions between the skilled and unskilled writers. First, despite the significant improvement made in the second self-monitoring annotation, the skilled writers obviously focused most of their attention on HOCs, whereas what concerned the unskilled writers remained mainly in LOCs. Such results conform with the perspective of Flower and Hayes (1981) and Zamel (1983) that writing effectiveness can be distinguished by whether or not writers' revision focus is on global aspects. Second, in relation to performing the self-assessment task, these two groups were different in several respects. The skilled writers apparently made more progress in metacognition, such as taking strategies to tackle a writing topic, performing critical thinking in journals, and self-monitoring writing problems. In addition, they were more able to make their assessment convincing by presenting detailed depiction of their change in writing before and after the study. Also, what concerned some of the skilled writers was how to use right language and details to make their work interesting and impressive to readers and while writing a composition, they were more aware of themselves as a writer. Having such reader and self awareness, according to Kasper's (1997) study on assessing ESL writers' metacognitive growth, is a key to becoming successful writers.

Even though using appropriate strategies and performing self-monitoring, peer reviewing as well as critical thinking were difficult to most unskilled writers, it is believed that consistent instruction, demonstration and practice might improve this, as one unskilled writer wrote in her journal entry, "After some practice, I now can examine my English writing, and find some problems." Writers like Usuki (2002) and Vanjee (2003) have confirmed that learning environment and teaching materials are the crucial factors in promoting learner autonomy. The current study therefore favors Xiao's (2007) idea that EFL writing instructors should teach *with* metacognition (equipped with metacognitive knowledge and keeping reflecting upon his/her own teaching) and *for* metacognition (helping learners to build up metacognitive knowledge and experiences). Writing proficiency involves not only linguistic competence but also the buildup of cognitive processing skills (Grabe & Kaplan, 1996). Metacognitive awareness helps add to learners' cognitive knowledge and skills, and this, as Wenden (1999) points out, enhances language learning.

With respect to the self-management level, qualitative analyses of the students' journal entries reveal that the skilled writers were more flexible in using strategies to cope with their writing problems. The strategies they employed were more varied and sometimes not confined to a solitary level. Such results are consistent with those of Bransford, Brown & Cocking (1999, cited in Rivers, 2001), Gardner & MacIntyre (1992) and Lan & Oxford (2003) who demonstrate that advanced learners tend to use a wider selection of strategies in new situations. Keeping flexible in adopting strategies also implies that the skilled writers are more willing to take risks and can be more precise in dealing with writing problems. Even though affective strategies were employed by some of the skilled and unskilled students, the former were more able to find solutions to their problems after relaxation. This indicates that affective strategies might lead to a more positive effect on the skilled writers than on the unskilled ones. Another noticeable difference observed is that the skilled writers tended to solve problems first by themselves and then resorted to other agents if needed. Going through such a conscious effort, as noted by Khaldieh's (2000) study, helps to enhance the process of language acquisition. The unskilled writers, in contrast, seldom experienced such a self problem-solving process. They often asked for help simply from their classmates or teacher.

CONCLUSION AND SUGGESTIONS

The main objective of this study was to examine metacognitive performance between skilled and unskilled student writers, in terms of self-monitoring, self-assessment and self-management, in a process-product EFL writing class. The results show that after the study, the majority of the skilled and unskilled subjects felt that they were less anxious about English writing and were clearer about how to organize an English essay. However, the skilled writers were distinguished from their counterparts in that after appropriate instruction, they were able to focus most of their attention on HOCs in self-monitoring annotations. In addition, the students' self-assessment shows that the skilled writers apparently made more progress in metacognition and in the composing process some of them tended to have reader and self awareness. Although both skilled and unskilled subjects expressed their difficulty in grammar and word usage, the skilled ones were concerned about how to make word choices and sentence patterns more effective and impressive to readers. In self-management, the skilled writers were more flexible, self-conscious and effective than the unskilled ones in using strategies to cope with their writing problems.

The current study seems to demonstrate metacognition as a significant predictor of English written proficiency and supports Rubin and Thompson's (1994) implication that metacognition might help set up optimal learning conditions and thus lower learners' 'affective filter'. However, since implementing metacognitive tasks means transferring some responsibilities to learners, which in turn might increase their pressure, particularly on the

less proficient ones, it is therefore suggested that explicit and direct instruction and modeling, and guided practice be consistently provided. Also, when teaching EFL writing metacognitively, the instructor should be supportive and encouraging to learners, and attend to their voices from different venues to monitor, evaluate and regulate the teaching strategies employed. In addition, the situation of taking metacognition into practice might be affected by cultural factors and need more research in this aspect. Studies by Littlewood (1999) and Palfrey (2004) report that collectivism and de facto acceptance of power and authority in Asian contexts can lead to hindrances in promoting learner autonomy, while those by Usuki (2002) and Vanijdee (2003) support otherwise. The effects of cultural factors on the execution of metacognitive tasks in an Asian writing setting hence need to be explored fully. Research in this area using qualitative methodologies will bring about new insights into the existing body of metacognition research in EFL writing.

Learning writing in an integrative approach makes learners aware that writing is not a one-step product of getting instant perfection but a recursive and social process of meaning exploration and reformulation. Promoting students' metacognitive awareness in an integrative EFL writing class implies that the process and the product approaches are not opposite but complementary (Devine, 1993). It also helps teachers and students to reflect on and regulate their efforts, thus making an EFL writing class more effective. However, it is suggested that as students' knowledge of and familiarity with process-product approaches accrue, genre approaches can also be incorporated as, in the words of Badger & White (2000: 155), "an extension of product approaches" to link writing with different social contexts.

REFERENCES

- Anderson, N. J. (2001). Developing Metacognitive Skills in Foreign Language Learners. *Selected Papers from the Tenth International Symposium on English Teaching* (pp. 1-7). Taipei: Crane.
- Angelova, M. (2001). Metacognitive knowledge in EFL writing. *Academic Exchange Quarterly*, 5 (3), 78-83.
- Applebee, A. N. (1986). Problems in process approaches: Towards a reconceptualization of process instruction. In A.R. Petrosky & D. Bartholomae (Eds.), *The teaching of writing: Eighty-fifth yearbook of the National Society for the Study of Education* (pp. 95-113). Chicago: University of Chicago Press.
- Badger, R. & White, G. (2000). A process genre approach to teaching writing. *ELT Journal*, 54, 153-160.
- Baker, L., & Brown, A. (1984). Cognitive monitoring in reading. In J. Flood (Ed.), *Understanding reading comprehension* (pp. 21-44). Newark, DE: International Reading Association.
- Blanton, L. L. (1987). Reshaping ESL students' perceptions of writing. *ELT Journal*, 41, 112-118.
- Charles, M. (1990). Responding to problems in written English using a student self-monitoring technique. *ELT Journal*, 44, 286-293.
- Cotterall, S. (2000). Promoting learner autonomy through the curriculum: Principles for designing language courses. *ELT Journal*, 54, 109-117.
- Crandall, J. A. (1999). Cooperative language learning and affective factors. In J. Arnold (Ed.), *Affect in language learning* (pp. 226-245). Cambridge: Cambridge University Press.
- Devine, J. (1993). The role of metacognition in second language reading and writing. In J. G. Carson & I. Leki (Eds.), *Reading in the composition classroom: Second language perspectives* (pp. 105-127). Boston: Heinle and Heinle Publishers.
- Flavell, J.H. (1979). Metacognition and cognitive monitoring: A new era of cognitive-development inquiry. *American Psychologist*, 34, 906-911.

- Flower, L. and Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32, 365-387.
- Gardner, R. C. & MacIntyre, P. D. (1992). A student's contributions to second language learning: Part I: Cognitive variables. *Language Teaching*, 25 (4), 211-220.
- Grabe, W. & Kaplan, R. (1996). *Theory and practice of writing*. London: Longman.
- Kamimura, T. (2000). Integration of process and product orientations in EFL writing instruction. *RELC Journal*, 31 (2), 1-28.
- Kasper, L. F. (1997). Assessing the metacognitive growth of ESL student writers. *TESL-EJ*, 3 (1).
- Khaldieh, S. A. (2000). Learning strategies and writing processes of proficient vs. less-proficient learners of Arabic. *Foreign Language Annals*, 33, 522-534.
- Lan R., & Oxford, R. L. (2003). Language learning strategy profiles of elementary school students in Taiwan. *International Review of Applied Linguistics in Language Teaching*, 41, 339-379.
- Leki, I. (1999). Techniques for reducing second language writing anxiety. In D.J. Young (Ed.), *Affect in foreign language and second language learning: A practical guide to creating a low-anxiety classroom atmosphere* (pp. 64-88). Boston: McGraw-Hill.
- Littlewood, W. (1999). Defining and developing autonomy in East Asian contexts. *Applied Linguistics*, 20, 71-94.
- Magno, C. (2008). Reading strategy, amount of writing, metacognition, metamemory, and apprehension as predictors of English written proficiency. *Asian EFL Journal*, 29, 15-48.
- Mokhtari, K., & Reichard, C.A. (2002). Assessing students' metacognitive awareness of reading strategies. *Journal of Educational Psychology*, 94(2), 249-259.
- Omaggio, A. C. (1986). *Teaching language in context*. Boston, MA: Heinle & Heinle.
- Oxford, R. (1989). *Language learning strategies: What every teacher should know*. New York: Newbury House.
- Palfreyman, D. (2004). Introduction. In D. Palfreyman & R. C. Smith (Eds.), *Learner autonomy across cultures*. Great Britain: Antony Rowe Ltd, Chippenham and Eastbourne.
- Raphael, T.E., Kirschner, B.W., & Englert, C.S. (1986). *Students' metacognitive knowledge about writing*. East Lansing, MI: Michigan State University, Institute for Research on Teaching.
- Raimes, A. (1985). What unskilled ESL students do as they write: A classroom study of composing. *TESOL Quarterly*, 19, 229-358.
- Rivers, W. P. (2001). Autonomy at all costs: An ethnography of metacognitive self-assessment and self-management among experienced language learners. *The Modern Language Journal*, 85 (ii), 279-290.
- Rose, M. (1984). *Writer's block: The cognitive dimension*. Carbondale, IL: Southern Illinois University Press.
- Rubin, J. & Thompson, I. (1994). *How to be a more successful language learner*. Boston, MA: Heinle and Heinle Publishers.
- Sasaki, M. (2000). Toward an empirical model of EFL writing processes: An exploratory study. *Journal of Second Language Writing*, 9 (3), 259-291.
- Schraw, G. (2001). Promoting general metacognitive awareness. In H.J. Hartman (Ed.), *Metacognition in learning and instruction* (pp. 3-16). Boston: Kluwer Academic.
- Tsui, A. B. M. (1996). Learning how to teach ESL writing. In D. Freeman & J. C. Richards (Eds.), *Teacher learning in language teaching* (pp. 97-119). Cambridge: Cambridge University Press.
- Usuki, M. (2002). *Learner autonomy: Learning from the student's voice*. Dublin: Center for

- Language and Communication Studies. (ERIC Document Reproduction Service No. ED 478 012).
- Vanijdee, A. (2003). Thai distance English learners and learner autonomy. *Open Learning*, 18(1), 75-84.
- Victori, M. (1999). An analysis of writing knowledge in EFL composing: A case study of two effective and two less effective writers. *System*, 27, 537-555.
- Wenden, A. (1999). Metacognitive knowledge and language learning. *Applied Linguistics*, 19, 515-537.
- Wenden, A. L. (2001). Metacognitive knowledge in SLA: The neglected variable. In M. Breen (Ed.), *Learner contributions of language learning* (pp. 12-24). Essex: Pearson Education Limited.
- Williams, M. & Burden, R. L. (2001). *Psychology for language teachers*. Cambridge: Cambridge University Press.
- Xiao, Y. (2007). Applying metacognition in EFL writing instruction in China. *Reflections on English Language Teaching*, 6(1), 19-33.
- Xudong, D. (2005). Teaching ESL/EFL writing: Approaches and pedagogical practices. *Selected Papers from the Fourteenth International Symposium on English Teaching*. Taipei: Crane.
- You, Y. L. (2003). The differences between L2 mature and immature writers: A metacognitive approach. *Proceedings of 2003 International Conference on English Teaching and Learning in ROC* (pp. 597-612). Taichung: Providence University.
- Zamel, V. (1983). The composing processes of advanced ESL students: Six case studies. *TESOL Quarterly*, 17, 165-187.

Facilitating EFL Reading Strategy Use Through a Course Management System

Yea-ru Tsai^a (蔡雅茹), Chia-pei Wu^a (吳佳蓓), and Cheng Hsu^b (許正)

^aI-Shou University

^bNational Chung Cheng University

yrtsai@isu.edu.tw

cpwu@isu.edu.tw

telch@ccu.edu.tw

This study reports on the design and implementation of a reading program in a blended course. The purpose of this program was to facilitate reading comprehension and strategy use for EFL learners in Taiwan. The reading strategies used in this project were classified into three categories: metacognitive, cognitive and social/affective strategies. Specific reading strategy instruction was embedded in a course management system (CMS), which included exercises on problem identification, monitoring comprehension, inferencing, summarizing, transfer, resourcing and questioning for clarification. The CMS-assisted reading training was integrated into an undergraduate level reading class. The comparison between pre-test and post-test indicated that students had improved their reading comprehension performance after the experimental course. A questionnaire was conducted to collect students' perceived learning progress and strategy use. Before the instruction, there was no difference in strategy use between skilled and less-skilled students. After the instruction, skilled students employed significantly more reading strategies than their less-skilled counterparts. However, no difference in strategy use was found among less-skilled students before and after the instruction. The results reveal that readers utilizing more strategies may have better performance in reading comprehension and show that CMS-assisted instruction is an effective approach to improve reading strategy use of EFL students. Pedagogical implications are briefly discussed based on the findings of this study.

INTRODUCTION

For more than two decades, strategy-based instruction has been regarded as an effective approach to enhance reading comprehension (e.g. Anderson 1999; Carrell, Pharis, & Liberto, 1989; Dole, Duffy, Roehler, & Pearson, 1991; Janzen & Stoller 1998; Souvignier & Mokhlesgerami, 2006; and Sung, Chang, & Huang, 2008). The efficiency of various types of strategies has been reported in previous studies on reading comprehension in different languages. A study conducted by Carrell et al. (1989), indicated that metacognitive strategy training was effective for the experimental group of ESL readers. Another study by Fotovatian and Shokrpour (2007) also confirmed that metacognitive strategies have positive effects on reading comprehension of Iranian students.

Concerning learning strategies with Chinese EFL learners, Zhang (2003) offered an extensive review. His recent study (Zhang, 2008) on the effect of strategic instruction demonstrated that the experimental group of Chinese EFL students had improved their reading performance and increased their strategy use after the experiment, compared to the control group.

In spite of the promising effects of the strategy-based instruction, there are also some implementation difficulties, as reported in various studies, such as strategy training of the

instructors (Duffy, 1993a; 1993b), time and effort on the preparation for the teaching materials (Pressley, Johnson, Symons, McGoldric, & Kurita, 1989). Recognizing these challenges, some researchers have constructed strategy-based instruction with the support of information technology, with satisfactory results (Chang, Sung, & Chen, 2002; Kaniel, Licht, & Bracha, 2000; Sung, Chang, & Huang, 2008). The general rationale of incorporating e-learning and strategy instruction is that such a combination may bring several advantages (Lynch, Fawcett, & Nicolson, 2000; Mathes, Torgeson, & Allor, 2001; Sung, et al., 2008).

In the context of EFL classrooms in Taiwan, the reading classes are offered often as large-scale classes, thus the interaction between instructors and students may be limited. It becomes difficult for instructors to offer feedback to their students immediately. Furthermore, compared to students in Western society, students in Taiwan are relatively shy and lack the confidence to ask questions in a foreign language. These aspects might reduce the efficiency of reading strategy instruction. Therefore the first purpose of this study was to investigate whether strategy instruction supported by information technology may enhance reading comprehension of university EFL learners in Taiwan.

Another controversial issue relates to the strategy use of learners with different levels of proficiency. Current available data reveals that poor readers are less efficient in using strategies in certain areas, compared to learners with better performance in reading (Dole, Brown, & Trathen, 1996; Garner, 1987; Lau & Chan, 2003; Paris, Lipson & Wixson, 1994). To date, there are still few studies which have investigated whether there are differing impacts of computer-assisted strategy instruction on strategy use of learners with different language proficiency. Therefore, the second purpose of the study is to compare the effects of computer-assisted strategy instruction on Taiwanese college EFL learners with different reading abilities. In the following sections, related literature will be reviewed, the methodology will be stated, the results will be explained and discussed, and areas for future studies will also be addressed.

Strategy-Based Reading Instruction

The efficiency of reading strategies identified in various studies has encouraged researchers to design instruction based on reading strategies. Some authors have developed instruction based on their own theoretical frameworks. For example, Allen (2003) compared three models of reading strategy instruction: the Reciprocal Teaching Approach (RTA) by Palinscar and Brown (1984), Transactional Strategy Instruction (TSI) by Pressley and Wharton-McDonald (1997), and the Cognitive Academic Language Learning Approach (CALLA) by Chamot and O'Malley (1996).

Following the taxonomy of Oxford (1990) and O'Malley and Chamot (1990), learning strategies can be divided into three categories: cognitive strategies, metacognitive strategies, and social/affective strategies. According to Chamot and O'Malley (1996), metacognitive strategies refer to strategies that the students use to plan, monitor, and evaluate their learning goals and processes. Cognitive strategies are used to accomplish both language and content tasks, including elaborating on prior knowledge, making inferences, and using imagery or linguistic transfer. Students may also use social and affective strategies to "complete a learning or communication task" (p. 264) through asking questions or cooperating with classmates. The reading strategies selected for the present study were based on the frameworks of Chamot and O'Malley (1996), Oxford (1990), Smidt and Hegelheimer (2004), and Yang (2006). The categories and descriptions of the reading strategies are summarized in Appendix A.

These different categories of strategies have been integrated by various researchers into reading comprehension instruction. The results of the study by Fotovatian and Shokrpour (2007) revealed that metacognitive strategies had positive effects on Iranian EFL readers, whereas no significant effect of cognitive strategies was found. Souvignier and

Mokhlesgerami (2006) used self-regulation as a framework for implementing strategy instruction to foster reading comprehension of German children and found that a strategy-oriented program proved to be effective. In another study by Zhang (2008), strategy-based instruction within a constructivist framework of was conducted using Chinese ESL learners. The findings revealed that the subjects' use of strategies and reading performance had been improved after the intervention. A rich body of literature also demonstrated that proficient readers tend to use more learning strategies more efficiently and successfully than less-proficient readers (Cohen, 1998; Oxford, 2001; Zhang, 2003).

Computer-Assisted Reading Instruction

In order to reduce the difficulties with implementing strategy instruction, one trend of the development is the integration of information technology with reading instruction. Some efforts of researching have been made on the effects of glossary, background information (Chang, Sung & Chen, 2002; MacArthur & Haynes, 1995), visual/auditory or static/animated aids of the text (Elkind, Cohen & Murray, 1993; Montali & Lewandowski, 1996), vocabulary recycling (Tozcu, & Coady, 2004; Johnson & Heffernan, 2006), and advance organizers with animation (Lin & Chen, 2007).

However, none of these computer-assisted strategy instructions has used a course management system to support the instruction. Therefore, this study intends to foster EFL reading strategy instruction through a course management system (CMS) called Knowledge Management and Language Learning System (KM & LLS) within a blended e-learning environment. The blending comes from a CMS and traditional instruction. There are numerous CMSs to date. In this study, MOODLE (Modular Object-Oriented Dynamic Learning Environment, Brandl, 2005; Cole, 2005; Robb, 2004) was selected as the infrastructure to support strategic instruction. MOODLE shares several common advantages with other CMSs.

In this blended class with a CMS curriculum, the learners have opportunities to practice the exercises constructed in the system, specific to individual strategies, without time limits. Hence, the purpose of this article is to address the following research questions:

- 1) What is the relative effect of CMS-assisted strategy instruction on reading comprehension performance of EFL learners with different reading proficiency?
- 2) Are there different effects of CMS-assisted strategy instruction on strategy use of EFL learners with different reading proficiency? If yes, in which concern?

METHOD

Subjects

The participants of the present study were 124 Chinese-speaking EFL learners at a university in Taiwan who were majoring in English (ages ranged from eighteen to twenty-four, with an average of twenty years). All subjects had studied English for at least six years in their high schools, and were enrolled in English reading comprehension courses. In order to investigate whether the instruction had different effects on learners with different proficiency levels, the students were classified into skilled and less-skilled readers by using a median split on their reading comprehension pretest (Media = 23.00, SD = 5.49). Those who performed above the median were deemed to be skilled readers and those who performed below the median were categorized as less-skilled readers.

Materials

KM&LLS was constructed to support ESL/EFL learners/teachers when they read/teach online by using MOODLE (<http://moodle.org>) course management system. The whole package of essential tools for the initial set-up of MOODLE can be found at MOODLE.org; for example, such tools as Apache, the homepage server, or a database system called MySQL, can be used for account management. The MOODLE system contains six basic elements: (1)

a content delivery interface, (2) a user-friendly editing system, (3) communication channels, (4) an administrative system, (5) assessment and evaluation, and (6) the teaching material repository.

Specific exercises were constructed in the system to enhance reading strategies. The students were assigned to practice the exercises each week according to the syllabus. To increase the use of metacognitive strategies, exercises of error identification and monitoring comprehension were constructed. The exercise on error identification is concerned with identifying syntactic structural errors. Ten syntactic categories were compiled, including nouns, articles, pronouns, relative clauses, verbs, infinitives, verbal nouns, participles, verb tenses and subject-verb-agreement. The students should be sensitive to correct syntactic structures in order to recognize the grammatical error appearing in each item (200 items in total).

Several exercises were considered to enhance cognitive strategy use of the students, including inferencing, summarizing, identifying the main idea, transferring and resourcing. Transferring strategy refers to the strategy of transferring L1 linguistic knowledge into L2. For resourcing, it is suggested to use reference materials, such as dictionaries, encyclopedias, or textbooks to support comprehension.

The strategy of summarizing was promoted through exercises on paraphrasing and identifying the main idea. In the exercise on paraphrasing, principles of paraphrasing were first provided step by step. In each item, the learners were asked to paraphrase the sentences in each paragraph given. Possible answers were provided. Identifying the main idea is essential for reading comprehension and summarizing. The exercise for identifying the main idea was multiple-choice. After reading a paragraph, the learners were supposed to identify the main idea of the paragraph by selecting the best answer.

Social/affective strategies can be used to access additional explanations or verification from a teacher, peers, or other experts. As mentioned in the above section, a CMS can provide a learning environment in which learners' anxiety levels may be reduced, because the learners don't have to talk to the instructor face to face. In order to facilitate social/affective strategies, several communication channels were activated, such as discussion forums, the chat room and the message board. The exercises on reading strategies embedded in KM&LLS are listed in Table 1.

Table 1. *List of exercises on reading strategies in KM&LLS.*

Strategy	Description	Units	Items
Metacognitive strategies			
Identifying grammatical errors	Awareness to identifying structural errors	10	200
Monitoring comprehension		10	50
Cognitive strategies			
Lexical inferencing	Using context clues; Analyzing words into various components, such as roots, affixes and suffixes	30	300
Summarizing	Paraphrasing the reading passage	10	10
Identifying main ideas	Identifying major points to train selective reading	15	15
Transferring syntactic knowledge	Using L1 syntactic knowledge to analyze L2 syntactic structure	10	10
Resourcing	On-line dictionaries, websites on English reading	-	-
Social-affective strategies			
Participating in chat-rooms, discussion forums, message boards		-	-
Total		85	585

Two Reading Comprehension Tests

A pre-test served as English reading competence in order to classify the students into two groups, skilled and less-skilled. The pre-test contained three parts: grammar, vocabulary, and reading comprehension. The grammar test instrument included the TOEFL (Institutional Based TOEFL) Section 2 (Structure and Written Expression: (1) Sentence completion task and (2) Error recognition task. Moreover, the vocabulary test was designed based on Test A of Nation's Vocabulary Levels Test (Nation, 1990). Only the three basic levels (2000-word, 3000-word, 5000-word levels) were administered in order to adjust to students' levels within this study. Eighteen items were selected, with six items at each level. Reliability analyses were conducted. The estimated reliability of the reading comprehension was acceptable (Cronbach's alpha 0.80 and 0.81 for the pre-test and post-test respectively).

Questionnaire on English Reading Strategies

The preliminary strategy classification used in the present study is presented in Appendix A. The distinction made in this taxonomy among metacognitive, cognitive and social/affective strategies is similar to that employed in Chamot and O'Malley (1996), Oxford (1990) and Smidt and Hegelheimer (2004).

Procedure

The pre-test of reading comprehension and the questionnaire on reading strategy were conducted during the first week of the course. The introduction to MOODLE system usage was conducted during the second week, followed by the eight-week CMS-assisted strategy instruction. After the blended course, the post-test reading comprehension was administered together with the questionnaire on strategy use after the completion of the program.

DATA ANALYSIS AND RESULTS

This study compared scores of reading comprehension and strategy use before and after the strategy instruction between skilled and less-skilled readers. Different statistical procedures were employed in order to explore the effect of the instruction. The data were

collected from the pre- and post-tests of reading comprehension and the questionnaire on strategy use. Paired sample T-tests were used to determine whether there were differences in the mean scores of the pre-tests and post-tests among skilled and less-skilled readers. ANOVA tests were computed to compare the different responses between two groups of readers.

Reading Comprehension before and after the Instruction

The test instruments used to assess reading comprehension and strategy use were administered to a total of 124 students, but the number of cases finally remitted for analysis was 112. Data from students who did not complete the test instruments were omitted. The mean scores of the pre-test and post-test of skilled ($n = 57$) and less-skilled ($n = 55$) readers are displayed in Table 3. Table 3 shows that after the strategy instruction reading comprehension increased in both groups. The mean performance of the less-skilled readers increased from 18.02 ($SD = 3.83$) on the pre-test to 24.07 ($SD = 4.26$) on the post-test, whereas the mean performance of the skilled readers increased from 25.76 ($SD = 3.97$) on the pre-test to 28.86 ($SD = 2.81$) on the post-test. The repeated-measure ANOVA on the pre-test, post-test and each sub-test shows a significant difference between the skilled and less-skilled readers before and after the strategy instruction, indicating that skilled readers scored higher than less-skilled readers in all variables.

To better investigate the effect of the instruction on each group, paired sample t-tests were conducted to compare the mean scores of the pre-tests, post-tests and each sub-test. Among less-skilled readers, the performance in each sub-test improved significantly after the instruction. Among skilled readers, the performance in the post-test was significantly better than in the pre-test, except in the sub-test of vocabulary ($t = -1.48$, $p = 0.146$, not significant). The results show that less-skilled readers made more progress after the instruction than skilled readers.

Analysis of reading strategy use

To examine the differences between skilled and less-skilled readers in the pre-test and post-test, a one-way ANOVA was conducted. In the pre-test, the results revealed that between-group differences were found only in three items: using context clues (item 10, $F = 5.277$, $P = 0.024$), analyzing word components (item 11, $F = 6.176$, $P = 0.015$) and translating words into Chinese (item 21, $F = 7.757$, $P = 0.006$). These findings indicate that before the instruction the skilled readers and less-skilled readers used reading strategies to a similar extent. Interestingly, the means in Table 5 also show that skilled readers utilized some reading strategies less frequently than less-skilled readers, such as planning the parts and sequence of the ideas (item 4), scanning key information (item 5), figuring out word meaning based on similarity with other words (item 12), analyzing the syntactic structure (item 14), trying to understand the texts without knowing the vocabulary (item 15), making a mental summary (item 20), translating words into Chinese (item 21), transferring Chinese to English (item 22), applying rules to complete the task (item 27), and asking teachers for clarification (item 28).

However, after the instruction, skilled readers performed better than did the less-skilled readers on every item of the post-test. There were significant between-group differences on all items except translating words into Chinese (item 21, $F = 0.842$, $P = 0.361$, not significant) and working with peers (item 29, $F = 2.091$, $p = 0.151$, not significant). Of all the strategies, the most prominent strategies discriminating the two groups were making analogies (item 17, skilled group $M = 3.54$ vs. less-skilled group $M = 2.54$; $F = 36.308$, $p < 0.001$), identifying structural or meaning errors (item 7, skilled group $M = 3.52$ vs. less-skilled group $M = 2.67$; $F = 35.177$, $p < 0.001$), writing down key words in simplified forms (item 26, skilled group $M = 3.74$ vs. less-skilled group $M = 2.94$; $F = 33.858$, $p < 0.001$) and classifying words according to their attributes (item 25, skilled group $M = 3.71$ vs. less-skilled group $M = 2.77$; $F = 31.711$, $p < 0.001$).

The results of paired-sample t-test reveal that no difference was found among less-skilled readers before and after the instruction, whereas there was an overall change in the skilled group's utility of all the listed reading strategies except one (item 13, using information in the text to guess meanings of new items or predict upcoming information, $t = -1.428$, $p > 0.05$, not significant). This might suggest that the subjects had already used guessing strategy prior to the strategy instruction program.

DISSCUSION

The major purpose of the present study was to explore the effects of CMS-assisted strategy instruction of reading comprehension and strategy use of Chinese EFL readings with different reading abilities. The findings of each area of inquiry are discussed in this section.

The results of the study indicated that differential performance was found between pre-test and post-test reading comprehension among skilled and less-skilled readers. The comparison showed that less-skilled readers achieved statistically higher mean scores on the reading comprehension measures in the post-test as compared to the pre-test. For the skilled readers, significant difference was found in all sub-tests except the sub-test evaluating vocabulary. The results reveal that the CMS-assisted strategy instruction has a positive effect on the reading comprehension of the participants.

As far as strategy use is concerned, the findings reveal that there is no significant difference between skilled and less-skilled readers prior to the instruction, thus suggesting that skilled and less-skilled readers possessed the ability of using reading strategies at a similar level before receiving the instruction. However, after the strategy-oriented program (instruction and CMS), skilled readers utilized significantly more strategies than less-skilled readers in all items in this study except two: translating words into Chinese and working with peers.

The results in general support the findings of previous studies. It was found in several studies (Dole, Brown & Trathen, 1996; Garner, 1987; Paris, et al., 1994; Lau & Chan, 2003) that poor readers were less efficient in using some strategies, such as identifying main ideas, decoding unfamiliar words, recognizing text structures, summarizing main ideas, detecting errors and inferring implicit meanings among Chinese subjects, as mentioned in Lau and Chan (2003). Lau and Chan (2003) further argued that the indirect explanation of reading texts instruction in Hong Kong may not offer sufficient help for poor readers to improve their ability of using higher-order cognitive and metacognitive strategies.

CONCLUSION

Many previous studies have shown satisfactory application results of reading strategy instruction programs (Brown, Pressley, Van Meter, & Schuder, 1996; De Corte, Verschaffel, & De Ven., 2001; Dreyer & Nel, 2003; Lau & Chan, 2003; Palincsar & Brown, 1984; Sung et al., 2008; Zhang, 2008), however, strategy-based instruction has long faced the challenges of time-consuming and exhausting implementation. The integration of CMS with reading strategy instruction is beneficial because it reduces the burden of preparation and efforts of the instructors and offers self-paced practice and resource functions for learners. With the support of the CMS, this study demonstrates that both groups of Chinese EFL students benefited from the strategic reading instruction in that skilled and less-skilled students showed significant progress in reading comprehension. Furthermore, the results also reveal that skilled readers employed reading strategies more frequently compared to their less-skilled counterparts, thus indicating that readers utilizing more strategies may have better performance in reading comprehension. Nevertheless, it does not mean that strategy-based instruction may not benefit less-skilled readers, because there is evidence showing the effectiveness of strategy instruction on readers with low reading ability in a computer-based

training program (e.g. Sung, et al., 2008).

Although this study echoes previous studies by emphasizing the positive effects of strategy instruction on EFL reading comprehension, there are some limitations to be noted. First, the sample size was relatively small and the data were collected among skilled and less-skilled readers. It is difficult to generalize the results of this study to or for other Chinese-speaking EFL learners; therefore the need for future studies including data with a larger sample size and a control group should be emphasized. The second limitation is concerned with system improvement. The implementation of strategy instruction was developed within a short period of time. Based on the consultation with some participants in this study, several aspects should be improved in the future. For example, more precise explanations to the mistakes made by the students should be provided. The current system only shows correct and incorrect answers on multiple choice items. Some students responded that the system was good and helpful for improving their reading comprehension, but they also suggested more functions should be added in the future. Last but not least, there are some influencing factors, which are equally important variables, but were not taken into consideration in the present study. Some of these variables include participants' learning style differences, motivational levels, topics of interest, cultural inclinations, difficulty levels of instructional materials, and L1 and L2 reading proficiency. Variables related to computer-assisted instruction include learners' knowledge, habit, interest and time of using the on-line course management system. Future studies involving these fields may provide valuable insight into the interwoven relations among these variables and reading strategy instruction.

ACKNOWLEDGEMENTS

The research and the preparation of the manuscript were supported in part by grants from the National Science Council under Contract No. 96-2411-H-214-003. We appreciate Yu-Kon Chang who has allowed us to link his website with the application of syntactic analysis to support reading comprehension.

REFERENCES

- Allen, S. (2003). An analytic comparison of three models of reading strategy instruction. *International Review of Applied Linguistics in Language Teaching*, 41(4), 319-338.
- Anderson, N. J. (1999). *Exploring second language reading*. Boston, MA: Newbury House.
- Brandl, K. (2005). Are you ready to "MOODLE"? *Language Learning and Technology*, 9(2), 16-23.
- Brown, R., Pressley, M., Van Meter, P., & Schuder, T. (1996). A quasi-experimental validation of transactional strategies instruction with low-achieving second-grade readers. *Journal of Educational Psychology*, 88, 18-37.
- Carrell, P., Pharis, B., & Liberto, J. (1989). Metacognitive strategy training for ESL reading. *TESOL Quarterly*, 23(4), 647-678.
- Chamot, A. U., & O'Malley, J. M. (1996). The Cognitive Academic Language Learning Approach: A model for linguistically diverse classroom. *The Elementary School Journal*, 96, 259-273.
- Chamot, A. U., O'Malley, J. M., Barnhardt, S., El-Dinary, P., & Robbins, B. (1999). *The learning strategies handbook*. New York: Longman.
- Chang, Y. K. (2007). *English for Science and Technology*. Taipei: New Wun Ching Developmental Publishing Co., Ltd.
- Chang, K. E, Sung, Y.T., & Chen, I. D. (2002). The effect of concept mapping to enhance text comprehension and summarization. *Journal of Experimental Education*, 71, 5-24.
- Cohen, A. (1998). *Strategies in Learning and Using a Second Language*. London: Longman.

- Cole, J. (2005). *Using MOODLE*. Sebastopol: O'Reilly community Press.
- De Corte, E., Verschaffel, L., & De Ven, V. (2001). Improving text comprehension strategies in upper primary school children: A design experiment. *British Journal of Educational Psychology*, 71, 531-559.
- Dhieb-Henia, N. (2003). Evaluating the effectiveness of metacognitive strategy training for reading research articles in an ESP context. *English for Specific Purposes*, 22, 387-417.
- Dole, J.A., Brown, K.J., & Trathen, W. (1996). The effects of strategy instruction on the comprehension performance of at-risk students. *Reading Research Quarterly*, 31, 62-88.
- Dole, J.A., Duffy, G. G., Roehler, L.R., & Pearson, P. D. (1991). Moving from the old to the new: research on reading comprehension instruction. *Review of Educational Research*, 61, 239-264.
- Dreyer, C., & Nel, C. (2003). Teaching reading strategies and reading comprehension within a technology-enhanced learning environment. *System*, 31, 349-365.
- Duffy, G.G. (1993a). Teachers' progress toward becoming expert strategy teachers. *The Elementary School Journal*, 94, 109-120.
- Duffy, G.G. (1993b). Rethinking strategy instruction: Four teachers' development and their low achievers' understanding. *The Elementary School Journal*, 93, 231-247.
- Elkind, J., Cohen, K., & Murray, C. (1993). Using computer-based readers to improve reading comprehension of students with disabilities. *Annals of Dyslexia*, 43, 238-259.
- Fotovatian, S. & Shokrpour, N. (2007). Comparison of the efficiency of reading comprehension strategies on Iranian university students' comprehension. *Journal of college reading and learning*, 37(2), 47-63.
- Garner, R. (1987). *Metacognition and Reading Comprehension*. Norwood, NJ: Ablex Publishing Co.
- Janzen, J., & Stoller, F. (1998). Integrating strategic reading in L2 instruction. *Reading in a Foreign Language*, 12(2), 251-269.
- Johnson, A., & Heffernan, N. (2006). The short readings project: A CALL reading activity utilizing vocabulary recycling. *Computer Assisted Language Learning*, 19(1), 63-77.
- Kaniel, S., Licht, P., & Bracha, P. (2000). The influence of metacognitive instruction of reading and writing strategies on positive transfer. *Gifted Education International*, 15, 45-63.
- Koda, K. (1994). Second language reading research: Problems and possibilities. *Applied Psycholinguistics*, 15, 1-28.
- Koda, K. (1999). Development of L2 intraword orthographic sensitivity and decoding skills. *The Modern Language Journal*, 83(1), 51-64.
- Koda, K. (2007). Reading and language learning: Crosslinguistic constraints on second language development. *Language Learning*, 57: Suppl. 1, 1-44.
- Lau, K. L. & Chan, D. W. (2003). Reading strategy use and motivation among Chinese good and poor readers in Hong Kong. *Journal of Research in Reading*, 26(2), 177-190.
- Lin, H.F. & Chen, T.P. (2007). Reading authentic EFL text using visualization and advance organizers in a multimedia learning environment. *Language Learning & Technology*, 11(3), 83-106.
- Lynch, L., Fawcett, A., & Nicolson, R. I. (2000). Computer-assisted reading intervention in a secondary school: An evaluation study. *British Journal of Educational Technology*, 31, 333-348.
- MacArthur, C., & Haynes, J. (1995). Student assistant for learning form text (SALT): A hypermedia reading aid. *Journal of Learning Disabilities*, 28, 150-159.
- Mathes, P. G., Torgesen, J. K., & Allor, J. H. (2001). The effects of peer-assisted literacy strategies for first-grade readers with and without additional computer-assisted

- instruction in phonological awareness. *American Educational Research Journal*, 38, 371-410.
- McWhorter, K. T. (2000). *Guide to college reading*. N.Y.: Addison Wesley Longman.
- Montali, P., & Lewandowski, L. (1996). Bimodal reading: Benefits of a talking computer for average and less skilled readers. *Journal of Learning Disabilities*, 29, 271-279.
- Nation, P. (1990). *Teaching and learning vocabulary*. Boston: Heinle & Heinle Publishers.
- O'Malley, J.M., & Chamot, A.U. (1990). *Learning Strategies in Second Language Acquisition*. Cambridge: Cambridge University Press.
- Oxford, R. L. (1990). *Language Learning Strategies: What Every Teacher Should Know*. Boston, MA: Heinle and Heinle.
- Oxford, R. L. (2001). Language learning strategies. In R. Carter & D. Nunan (Eds.), *The Cambridge guide to teaching English to speakers of other languages* (pp. 166–172). Cambridge, UK: Cambridge University Press.
- Palinscar, A., & Brown, A. (1984). Reciprocal teaching of comprehension and monitoring activities. *Cognition and Instruction*, 1-2, 117-175.
- Paris, S.G., Lipson, M.Y., & Wixson, K.K. (1994). Becoming a strategic reader. In R.B. Ruddell, M.R., Ruddell & H. Singer (Eds.), *Theoretical models and processes of reading* (pp. 788-811). Newark, DE: International Reading Association.
- Paris, S. G., & Paris, A.H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist*, 36, 89-101.
- Phillips, D. (2007). *Longman Preparation Course for the TOEFL Test*. New York: Pearson Education, Inc.
- Pressley, M., Johnson, C. J., Symons, S., McGoldric, J.A., & Kurita, J.A. (1989). Strategies that improve children's memory and comprehension of text. *Elementary School Journal*, 90, 3-32.
- Pressley, M., & Wharton-McDonald, R. (1997). Skilled comprehension and its development through instruction. *School Psychology Review*, 26(3), 448-466.
- Robb, T.N. (2004). Moodle: A Virtual Learning Environment for the Rest of Us. *Teaching English as a Second or Foreign Language*, 8(2), 1-8.
- Salataci, R. & Akyel, A. (2002). Possible effects of strategy instruction on L1 and L2 reading. *Reading in a Foreign Language*, 14(1), 1-16.
- Smidt, E. & Hegelheimer, V. (2004). Effects of online academic lectures on ESL listening comprehension, incidental vocabulary acquisition, and strategy use. *Computer Assisted Language Learning*, 17(5), 517-556.
- Smith, B.D. (2003). *Bridging the Gap*. N.Y.: Addison-Wesley Education Publishers.
- Souvignier, E. & Makhlesgerami, J. (2006). Using self-regulation as a framework for implementing strategy instruction to foster reading comprehension. *Learning and Instruction*, 16, 57-71.
- Sung, Y.T., Chang, K. E., & Huang, J. S. (2008). Improving children's reading comprehension and use of strategies through computer-based strategy training. *Computers in Human Behavior*, 24, 1552-1571.
- Tozcu, A., & Coady, J. (2004). Successful learning of frequent vocabulary through CALL also benefits reading comprehension and speed. *Computer Assisted Language Learning*, 17(5), 473-495.
- Wegmann, B. & Knezevic, M. (2007). *Mosaic 1 - Reading*. Columbus, OH: McGraw-Hill companies.
- Yamashita, J. (2002a). Mutual compensation between L1 reading ability and L2 language proficiency in L2 reading comprehension. *Journal of Research in Reading*, 25(1), 81-95.
- Yamashita, J. (2002b). Reading strategies in L1 and L2: Comparison of four groups of

readers with different reading ability in L1 and L2. *Review of Applied Linguistics*, 135-136, 1-35. Yang, Y. F. (2006). Reading strategies or comprehension monitoring strategies? *Reading Psychology*, 27, 313-343.

Zhang, L. J. (2003). Research into Chinese EFL learning strategies: Methods, findings and instructional issues. *Regional Language Centre Journal*, 34(3), 284-322.

Zhang, L. J. (2008). Constructivist pedagogy in strategic reading instruction: Exploring pathways to learner development in the English as a second language (ESL) classroom. *Instructional Science*, 36, 89-116.

APPENDIX A: Framework of reading comprehension strategies

Strategy	Description	Definition
Metacognitive strategy		
Planning		
Advanced organization	Preview Skim Gist	Previewing the main ideas and concepts of a text; identifying the organizing principle
Organizational planning	Plan what to do Outline	Planning how to accomplish the learning task; planning the parts and sequence of ideas to express
Selective attention	Read selectively Scan Find specific information	Attending to key words, phrases, ideas, linguistic markers, types of information
Self-management	Plan when, where, and how to study	Seeking or arranging the conditions that help one learn
Evaluating	Performance evaluation Strategy evaluation	Evaluating the function of the strategy
Problem identification	Identify errors	Identify structural or meaning errors
Monitoring		
Monitoring comprehension	Think while reading	Checking one's comprehension during reading
Evaluating: self-assessment	Check back Keep a learning log Reflect on what is learned	Judging how well one has accomplished a learning task
Cognitive strategies		
Inferencing	<ul style="list-style-type: none"> ● linguistic inferencing ● lexical inferencing ● kinesic inferencing ● extralinguistic inferencing ● between parts inferencing 	Using context clues; Analyzing words into various components, such as roots, affixes and suffixes; Figuring out word meaning based on its sound or form similarity with other words; Using information in the text to guess meanings of new items or predict upcoming information
Elaboration	<ul style="list-style-type: none"> ● personal elaboration ● world elaboration ● academic elaboration ● questioning elaboration 	Using what you know Using background knowledge Making analogies Relating new to known

	<ul style="list-style-type: none"> ● creative elaboration ● imagery; visualize; make a picture 	information and making personal association; Using mental or real pictures to learn new information or solve a problem
Summarizing	Say or write the main idea	Making a mental, oral, or written summary of information gained from reading
Translation	Translate the text from L2 to L1	Translating the word or sentences in one's own mother tongue

APPENDIX A: Framework of reading comprehension strategies (cont.)

Strategy	Description	Definition
Cognitive strategies		
Transfer	Transfer L1 linguistic knowledge into L2 <ul style="list-style-type: none"> ● phonemic awareness ● phonemic-graphemic knowledge ● syntactic feature recognition 	Transferring L1 linguistic knowledge into L2
Repetition	Repeat the part of the text.	Repeating any portion of the text, including the word, the phrase, or the sentence in which the word has occurred.
Resourcing	Use reference materials	Using reference materials such as dictionaries, encyclopedias, or textbooks
Grouping	Classify Construct graphic organizers	Classifying words, terminology, quantities, or concepts according to their attributes
Note taking	Take notes on idea maps, T-lists, semantic webs, etc.	Writing down key words and concepts in abbreviated verbal, graphic, or numerical form
Detection/induction	Use a rule/make a rule	Applying or figuring out rules to understand a concept or complete a learning task
Social/affective strategies		
Questioning for clarification	Ask questions	Getting additional explanation or verification from a teacher or other experts
Cooperation	Cooperate Work with classmates Coach each other	Working with peers to complete a task, pool information, solve a problem, or get feedback
Lowering anxiety/self-encouragement	Think positive Talk yourself through it	Reducing anxiety by improving one's sense of competence

This framework is based on O'Malley and Chamot (1996), Oxford (1990), Smidt & Hegelheimer (2004) and Yang (2006).

English Teachers' Conceptual Change in CALL: A Reflective Approach

Jun-jie Tseng (曾俊傑)
National Taiwan Normal University
jjtseng@ntnu.edu.tw

It is often seen that English teachers hesitate to use the computer in their instructions after completing CALL teacher training programs. One possible reason for this problem is that the way those English teachers learn can not cause a significant change in their conceptualizations of CALL so that they still hold a conservative attitude towards the integration of technology into teaching practices. To address this problem, the present study developed a 12-week CALL workshop that incorporated reflective practices. Successfully completing the CALL workshop in which they were required to submit weekly reflection entries on the discussion forum, four in-service English teachers constituted the ultimate subjects of the study. All of their reflection entries were evaluated for their critical levels according to the revised Bloom's Taxonomy scale. It was found that the four teachers' reflectivity was overall on the increase. At the beginning of the workshop, their reflections were mostly rated at the lower levels of the scale, which involve understanding pedagogical significance associated with technology and applying technologies to particular contexts. Toward the end of the workshop, their reflections were promoted to the higher levels, which involve viewing CALL practices with analytical thoughts and evaluating them with theoretical concepts. In addition, they acknowledged the reflective practice as a crucial procedure that could facilitate their reflections. The CALL workshop developed in this study seemed to have contributed to the four teachers' professional development in terms of enhancing critical thinking on the connection between technology and pedagogy.

INTRODUCTION

Although computer technology is advocated at various levels of education in Taiwan, computer-enhanced teaching practices have not been widely and substantially embraced by the teachers in Taiwan. In 2000, Lee (2000) pointed out that the percentage of teachers enhancing their teaching activities with computer technology was low. In her survey administered to 2,965 elementary and high school teachers in Taiwan regarding their information literacy, it was found that one out of five teachers used the computer or the Internet in his or her teaching. While the computer was mostly utilized in producing test sheets (62%), preparing for the lessons (57.5%), tallying scores (53.9%) and giving assignments (43.1%), it was less often applied to classroom-based teaching activities (19.9%). This survey showed that the teachers might lack an ability to integrate the computer into their teaching despite the fact that the computer became more available at school. This situation was not considerably improved with years. In 2006, Chang and Wong (2006) conducted a similar questionnaire survey regarding elementary school teachers' competency in integrating the computer into the curriculum in the Great Taipei Metropolitan Area, where nearly one-third of Taiwanese citizens reside. With 233 valid responses returned by 30 schools, their study revealed that many a teacher was devoid of capabilities and experiences in infusing information technologies into teaching activities. Although some of the teachers were very skilled at applying technologies to their teaching practices, their computer-supported teaching was merely seen as a low level of incorporating information technologies into teaching, in which technology is employed mostly for the sake of technology itself. The same happens to English teachers who have attended ICT (information & computer technology) training courses and workshops. After the training programs, many of them still hesitate to integrate

the computer into their instruction (Chang, 2003; Kim, 2001). One possible reason for this problem is that the way they learn can not cause a significant change in their conceptualizations of CALL. To address this problem, the researcher designed a CALL teacher development (CALLTD) workshop that incorporated reflective practices, which are believed to “raise teachers’ awareness about teaching, enable deeper understanding, trigger positive change, and make oneself prone to change” (Liou, 2000, p. 13).

Reflection-based teacher development program has been acknowledged in the literature (e.g. Liou, 2000; Richards, 1998; Richards & Lockhart, 1996; Wallace, 1991). What distinguishes reflection-based teacher development program from general teacher training programs lies in the provision of opportunities for teacher learners to reflect on their classroom practices (e.g. Arechaga, 2001; Lord & Lomicka, 2007). Although studies generally showed that a fairly large number of teachers failed to reflect critically, they revealed that many teachers held a positive attitude toward the reflective procedures utilized in the teacher development programs (e.g., El-Dib, 2007; Farrell, 1999; Kanuka & Anderson, 1998; Lee, 2007; Liou, 2001; Tolmie & Boyle, 2000; Zhu, 2006). Moreover, there were a few studies that acknowledged CMC technology as facilitating teachers’ reflections. They claimed that CMC applications, such as e-mail and discussion forum, could allow teachers to nurture reflections without the constraints of time and space (Carboni, 1999; Edens, 2000; Wade & Fauske, 2004).

Thereby, a reflective approach to CALLTD workshop was adopted to help the participating teachers develop their knowledge and beliefs about CALL. The present study intended to investigate whether their reflections on CALL could be promoted over the course of the study. It was hoped that changes in their reflectivity could be traced through the study.

THE STUDY

A 12-week CALLTD workshop was conducted from April 1 through June 17, 2006 at the English Language Training Center of a national university in Taiwan. During this period, the teachers, who came from various local elementary or high schools, met face to face three hours a week. The on-campus workshop took up thirty-six hours in total. In addition to the weekly 3-hour teacher-student interactions in a physical classroom, the CALLTD workshop also involved online mode, which referred to a web-based learning environment available twenty-four hours a day seven days a week.

At the beginning of the workshop, there were sixteen in-service English teachers who were recruited to enroll in the workshop. At the end of the workshop, only four teachers (under pseudonyms: Shirley, Fanny, Teresa, and Fiona) were considered valid cases in this study because they were the only ones who completed all the tasks required by the study, such as submitting reflection entries and going for an interview with the researcher.

The CALLTD workshop was designed with a focus on helping English teachers not only learn technological skills but also develop proficiency in relating technology to individual teaching contexts through reflective procedures. In particular, the participating teachers were asked to share ideas and thoughts on the discussion forum where the teacher educator (also the researcher of the study) provided discussion prompts specifically related to learning content and activities assigned for each week.

To understand the four teachers’ development of reflectivity, data collected from the discussion forum were examined. The discussion forum provided one source of data on the four teachers’ thoughts about reflective practice employed in the workshop. In addition, interviews were conducted with the four teachers for clarifying unclear statements and points made in their discussion forum entries.

To determine the reflectivity level of the forum entries, the revised Bloom’s *Taxonomy for Educational Objectives* (Anderson, Krathwohl, Airasian, & Bloom, 2001), which has been

used to measure teachers' reflections in some studies (e.g. Christopher, Thomas, & Tallent-Runnels, 2004; Crotty & Allyn, 2001), was used as a coding framework because it provides a fine distinction of reflectivity levels: *Remember*, *Understand*, *Apply*, *Analyze*, *Evaluate*, and *Create*. Accordingly, each meaningful segment out of a forum entry was coded into one of the six categories: *Remember*, *Understand*, *Apply*, *Analyze*, *Evaluate*, and *Create*, with *Remember* representing the lowest level of reflectivity and *Create* the highest level. For the convenience of rating, a numerical score was given to each level to indicate the degree of critical reflection: 1 (*Remember*), 2 (*Understand*), 3 (*Apply*), 4 (*Analyze*), 5 (*Evaluate*), and 6 (*Create*).

RESULTS

Except for Teresa, in the beginning (weeks 1 to 4) of the workshop, the teachers' reflectivity largely fell into the lower levels in Bloom's Taxonomy, i.e., levels 2 (*Understand*) and 3 (*Apply*). That is, they were able to articulate what technologies meant in English teaching contexts and to invent some teaching scenarios in which computer technologies could be utilized. Since the middle of the workshop, specifically the 5th week, all of the four teachers' reflectivity grew to the higher end of the taxonomy, varying between level 4 (*Analyze*) and level 5 (*Evaluate*). This growth sustained to the end of the workshop. That is to say, since the middle of the workshop, they could not only organize their views for and against particular CALL practices but also evaluate them according to some theoretical concepts. Overall, they exhibited a trend of growth in reflectivity level over the course of the workshop.

According to the individual interviews with the four teachers, they generally thought that the interactive reflective practice, i.e. the discussion forum, provided them with an opportunity to communicate their thoughts to one another and the teacher educator, thereby facilitating them to engage in critical reflections. Specifically, the four teachers were facilitated to initiate their reflections by referring to one another's ideas submitted onto the discussion forum. Therefore, the four teachers often did not submit their reflection entries until the last minute of the due time. The reason was that while they nurtured their reflections, they were usually waiting for others' reflection entries to help them generate innovative ideas. Shirley particularly relied on the discussion forum to stimulate her reflections: "*I am a person who has a bit of difficulty expressing thoughts, so I frequently waited for others' completed reflection entries from which inspiration was sought.*" (forum reflection, Shirley, 7/27/2006)

The teacher educator's comments on the discussion forum also played an important role in enhancing the participating teachers' reflectivity. Without the teacher educator's feedback, the four teachers might have focused more on the technical aspect of the computer than on the pedagogical aspect of the computer. For example, Teresa thought that the teacher educator's replies to her reflection entries helped her link technology to pedagogy. She said:

The role of the teacher [educator] was very important. Not just giving a reply, in fact, you (the teacher educator) stimulated everyone to re-log on [to the website] to click on [the forum entries]. With your frequent stimulations, we had to constantly reflect on the relations [of computer technology] to teaching practices ... We couldn't possibly have generated so much thought if you hadn't frequently helped us to hold on to this topic [purpose]. (interview, Teresa, 7/17/2006)

DISCUSSION AND CONCLUSION

As reported in the previous section, the four teachers' reflectivity grew to the critical end of Bloom's Taxonomy scale over the workshop. This result is in agreement with Garvin's (2003) finding that student teachers exhibited reflections of advanced levels via the

discussion board; it also supports previous literature on the effectiveness of technology in promoting teachers' reflections (e.g. Duffy, Dueber, & Hawley, 1998; Strampel & Oliver, 2007).

Asynchronous discussion forum was thought to facilitate the four participants' reflectivity development. This tool might help them overcome the problem of having no ideas for reflection because the discussion forum offered the participants chances to view others' reflection entries and receive input from their peers. As maintained by several researchers of teacher education, the discussion forum could give teachers access to others' pedagogical ideas associated with technology (Ruan & Beach, 2005; Schuck, 2003), "expansive feedback" (Ostorga & Yanes, 2007), and "professional dialogue" that could facilitate teachers' reflections (Wickstrom, 2003).

In addition to the discussion forum, the teacher educator was found to play an important role in the development of the participating teachers' reflectivity. In particular, the teacher educator helped direct the four teacher trainees' reflections on the relation of technology to pedagogy. For example, Teresa reported that the teacher educator had been stimulating her thoughts about the significance of pedagogy in utilizing a particular technology. Without the guidance and advice of the teacher educator, her reflections might end up simply focusing on technical issues rather than pedagogical implications. Therefore, like those pre-service teachers in Garmon's (2001) and Lee's (2007) studies, Teresa showed great appreciation for the teacher educator's advice and comments.

This study has provided empirical evidence on how a reflection-based CALLTD workshop could help in-service teachers promote their reflectivity through the discussion forum. It hopes to inspire teacher educators to help English teachers revolutionize and transform their teaching practices with the computer.

REFERENCES

- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., & Bloom, B. S. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives* (Complete ed.). New York ; Toronto: Longman.
- Arechaga, D. M. D. (2001). Teacher development: Awareness, reflection and sharing. *ELT Newsletter*, Article 69. Retrieved from <http://www.eltnewsletter.com/back/August2001/afl692001.htm>
- Chang, C.-S., & Weng, W.-T. (2006). Investigating the competency of information technology integrated into teaching of elementary school in-service teachers: For example of Taipei County. *Journal of National Taipei University of Education*, 19(2), 129-162.
- Chang, Y. Y. (2003). *A preliminary survey of Taiwan's English teachers' attitudes toward and applications of the Internet*. Unpublished thesis, National Taiwan Normal University, Taipei.
- Christopher, M. M., Thomas, J. A., & Tallent-Runnels, M. K. (2004). Raising the bar: Encouraging high level thinking in online discussion forums. *Roeper Review*, 26(3), 166-171.
- Crotty, T., & Allyn, D. (2001). *Evaluating student reflections*. Wisconsin: University of Wisconsin--River Falls.
- Duffy, T., Dueber, B., & Hawley, C. (1998). Critical thinking in a distributed environment: A pedagogical base for the design of conferencing systems. In C. Bonk & K. King (Eds.), *Electronic collaborators: Learner-centered technologies for literacy, apprenticeship, and discourse* (pp. 25-78). Mahwah, NJ: Lawrence Erlbaum.
- Garmon, M. A. (2001). The benefits of dialogue journals: What prospective teachers say. *Teacher Education Quarterly*, 28(4), 37-50.

- Garvin, T. (2003). *Using threaded discussion to develop reflective practice in preservice teachers*. Paper presented at the Society for Information Technology and Teacher Education International Conference (SITE) 2003, Albuquerque, New Mexico.
- Kim, H. (2001). *Initial efforts of three ESL teachers to effectively utilize network-based language teaching*. Unpublished dissertation, State University of New York at Buffalo.
- Lee, I. (2007). Preparing pre-service English teachers for reflective practice. *ELT Journal*, 61(4), 321-329.
- Lee, S.-L. (2000). A Survey on the Teachers' Abilities to Use Information Network. *Commonwealth Magazine*, 29.
- Lord, G., & Lomicka, L. (2007). Foreign language teacher preparation and asynchronous CMC: Promoting reflective teaching. *Journal of Technology and Teacher Education*, 15(4), 513-532.
- Ostorga, A. N., & Yanes, M. J. (2007). Web-based discussion: Construction of meaning. *The Texas Journal of Distance Learning*, 4(1), 1-12.
- Ruan, J., & Beach, S. A. (2005). Using online peer dialogue journaling to promote reflection in elementary preservice teachers. *Action in Teaching Education*, 27(3), 64-75.
- Schuck, S. (2003). The use of electronic question and answer forums in mathematics teacher education. *Mathematics Education Research Journal*, 5, 19-31.
- Strampel, K., & Oliver, R. (2007). *Using technology to foster reflection in higher education*. Paper presented at the ICT: Providing choices for learners and learning. Proceedings ascilite Singapore 2007.
from <http://www.ascilite.org.au/conferences/singapore07/procs/strampel.pdf>
- Wickstrom, C. D. (2003). A "funny" thing happened on the way to the forum. *Journal of Adolescent & Adult Literacy*, 46(5), 414-423.

Validating a Pictorial Vocabulary Size Test Through Rasch Modeling

Wen-ta Tseng (曾文鍵)
National Taiwan Normal University
wtttseng@ntnu.edu.tw

The paper presented a newly-conceived vocabulary size test based on pictorial cues. The vocabulary measurement was developed in principle in reference to the 1200-word list as prescribed by the Ministry of Education in Taiwan. The preliminary form of the test contained ninety items in total. The Rasch model was adopted as the analytical framework to execute the validation analysis. 242 junior high school students participated in the study. The results showed that the test was productive, reliable and valid. Although some items perform heterogeneously toward the Rasch Model, 79 items (88%) of the pilot items can receive good fit. Furthermore, it was found that the items of the Pictorial Vocabulary Size Test (PVST) are comparatively easy for the test-takers to obtain high scores. The article concluded that the PVST could have positive washback effects on test development, English vocabulary instruction and social equality.

INTRODUCTION

The study of vocabulary has received a great amount of attention over the past decade in the field of second language acquisition. Significant progress has been made in the areas such as vocabulary learning strategies (Tseng, Dörnyei & Schmitt, 2006; Tseng & Schmitt, 2008; Tseng, 2008), measurement of vocabulary knowledge (Qian, 1999, 2001; Meara & Milton, 2003a, 2003b; Fitzpatrick, 2007; Van Hout & Vermeer, 2007), and formulaic sequences (Schmitt, 2004; Wray, 2002). Among these different lines of research on vocabulary, it is noted that the issue of how to measure a learner's vocabulary size accurately and reliably has been a focal and central sphere for vocabulary researchers to investigate.

Vocabulary size is a key indicator of lexical ability and language proficiency. A wide range of research has unanimously indicated that certain and different vocabulary sizes are necessary to complete different language tasks (e.g., Adolphs & Schmitt, 2003; Hazenberg & Hulstijn, 1996; Laufer, 1988; Nation & Waring, 1997). The lexical requirements for English can be summarized as follows:

- 2,000-3,000 word families for basic everyday conversation (chat)
- 3,000 word families to begin reading authentic texts
- 5,000 word families to independently read authentic texts
- 10,000 word families are a wide vocabulary which should allow most language use

(Source: Tseng & Schmitt, 2008, p. 366)

Given the significance of acquiring sufficient English vocabulary size, the Ministry of Education in Taiwan has also paid close attention to the role of vocabulary size as played in English language learning. In the year of 2003, the Ministry of Education in Taiwan published One-Thousand Words List (now One-Thousand and Two-Hundred Words List), the vocabulary items of which are all high-frequency words. The reason for compiling the word list is that the list can serve as the guideline for different textbook publishers to follow at the levels of primary and junior high school (教育部國教司, 2007). The formation of the word list made reference to a number of authoritative publications:

本字彙表的制定係參考多種字彙來源，包括民國八十三年教育部發布的「國民中學課程標準」參考字彙表、韓國小學課程綱要參考字彙、上海九年制課程綱要詞彙表、日本初中英語教材常用字彙、數種國內外兒童英語教材常用字彙、LTTC「全民英檢初級檢定二千字彙參考表」，以及民國八十四、九十、九十一年大學入學考試中心公佈的高中英文參考詞彙表中的最常用字詞，和 Collins COBUILD 最新字典(1995, 2001)所制定的最常用英文字彙表(該字典是根據 The Bank of English 語料庫編制的)，綜合彙整，輸進電腦，建立一詞庫及字頻表後，並參考美國、英國、南非、及日本等國最常用英文字彙，最後再依我國中小學階段學生的認知能力、生活經驗，英語學習目標，及外語學習環境等因素，由編輯、諮詢委員多次討論、研商後，加以篩選調整完成。(教育部國教司, 2006)

The scope of the references is comprehensive, and the procedures for compiling the word list are also rigid. Furthermore, it has been suggested in the curriculum guidelines that primary school students should at least be able to utter 300 words, and spell 180 words from which they can orally produce(教育部國教司, 2006). For junior high school students, the whole word list can be used to prepare for Basic English Competency Test.

Notwithstanding the existence of the word list, it is however noted that thus far there are few attempts made to develop a reliable and valid vocabulary size test on the basis of this word list. Pedagogically, Ryan (1997), on discussing the ability to recognize word form, argues that “failures at the word level can severely hamper reading ability, and reading ability is a key skill in using English for academic or professional purposes” (p.187). Thus, inasmuch as it is necessary to develop the vocabulary size test of the 1200-word list, the purpose of this project is three-fold: first, to construct a preliminary form of a newly-conceived vocabulary size test based on pictorial cues; second, to examine the degree to which the test items of the test can fit into the measurement model, and third, to determine the difficulty level of the whole test. It is argued that the vocabulary size test to be developed can be used to diagnose or identify both primary and junior high school Taiwanese learners’ strengths and weaknesses regarding the words in the list.

FORMAT OF THE TEST

The test format of the Pictorial Vocabulary size Test bears close resemblance to the one in the Vocabulary Levels Test (Nation, 1990, Schmitt, 2000). Six words in the left-hand column - both target words and distractors - are arranged in such a way that the selected words are semantically-linked. This test feature helps enhance the validity of the test because according to Carter (1998), “words do not exist in isolation: their meanings are defined through the sense relations they have with other words” (p. 19). In total, there are 30 clusters included in the test and ninety items are tested. Similarly, the 30 clusters of the test fall into a 3 (noun) : 2 (verb) : 1 (adjective) ratio. Thus, the pilot form of the Pictorial Vocabulary Size Test comprises of 15 noun clusters, 10 verb clusters, and 5 adjective clusters. This 3-2-1 ratio is also used in Schmitt et al.’s Vocabulary Levels Test.

Research Questions

The research questions of the study are as follows:

1. What are the reliability and validity of the pilot form of the Pictorial Vocabulary Size Test?
2. To what extent the 90 test items fit into the Rasch model?
3. Do the participants consider the test difficulty or easy?

Participants

The participants came from two junior high schools in Taoyuan County. 102 were first graders and 140 were 2nd graders. The average years of English learning of this sampling ranged between 0 and 5 years.

Software of the Rasch model

The latest version of Winsteps computer program – Winsteps 3.67 – was used in the pilot study.

RESULTS

In the Rasch model, reliability is divided into two parts: person reliability and item reliability. The former can offer the information of the extent to which the ability ordering of the same sample of persons can be replicated by another set of items measuring the same underlying construct, whereas the latter can offer the information of the extent to which the difficulty ordering of the same items can be replicated by another sample of similar ability level (Bond & Fox, 2001). High person reliability can ensure that both higher and lower ability of persons exist. High item reliability, on the other hand, ensures that both more difficult and easier items can be identified.

The results showed that both the person (Table 1) and item (Table 2) reliability of the Pictorial Vocabulary Size Test (PVST) were .97 and .98, respectively. The two high reliability indices computed by the Rasch model clearly reveal that the PVST can possess fairly satisfactory reliability. Likewise, where the location of the persons relative to the items is concerned, it was found that the mean item measure was .00, whereas the mean person measure was 1.15. This implies that the items of the PVST are comparatively easy for the test-takers to obtain high scores.

Table 1. *Summary of 242 measured persons*

	RAW SCORE	COUNT	MODEL		INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
			MEASURE	ERROR				
MEAN	59.6	90.0	1.15	.33	.98	.0	1.14	.1
S.D.	22.3	.1	1.91	.09	.38	1.2	1.13	1.3
VALID RESPONSES	99.9							
Person Reliability	= .97							
Cronbach Alpha (KR-20) Person Raw Score Reliability	= .98 (approximate due to missing data)							

Table 2. *Summary of 90 Measured Items*

	RAW SCORE	COUNT	MODEL		INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
			MEASURE	ERROR				
MEAN	160.2	242	.00	.19	.98	-.2	1.14	.2
S.D.	44	.2	1.58	.04	.20	1.9	.77	1.9
Item Reliability	= .98							

Two important statistics also need to be considered and reported: Infit mean-squares (Infit MNSQ) and Outfit mean-squares (Outfit MNSQ). Infit MNSQ statistic the two values report the information as to the extent of score variation between observed and predicted response patterns. Outfit mean squares statistic includes all the observations in the data whereas infit mean squares statistic excludes outliers of the data. The results derived from these two fit measures can determine how meaningful and productive the test items are for the whole measurement system (Linacre, 2004). Ideally, both infit and outfit mean squares are expected to have value 1, which exhibits a perfect fit. According to Linacre (2004), he offers guidelines as shown in Table 3 on how to interpret the values of both fit statistics.

Table 3. *The interpretations of fit statistics in Rasch analysis*

>2	Off-variable noise is greater than useful information. Degrades measurement
>1.5	Noticeable off-variable noise. Neither constructs nor degrades measurement
.5 -1.5	Productive measurement.
< .5	Overly predictable. Misleads us into thinking we are measuring better than we really are. (Attenuation paradox.). Misfits <1.0 are only of concern when shortening a test

The mean fit values of both the infit and outfit MNSQ of the PVST as reported in Table 2 were .98 and 1.14. Since the two mean fit values fluctuate around 1, this result suggests that the PVST is a very productive measurement as a whole.

The ICCs of all the 90 items were shown in Figure 1. It was found that the ICCs of all the 90 items functioned in parallel without intersecting with each other. Moreover, Table 4 summarizes the difficulty measure, standard error, infit MNSQ and outfit MNSQ for each item.

Figure 1 The ICCs for All the Ninety Items

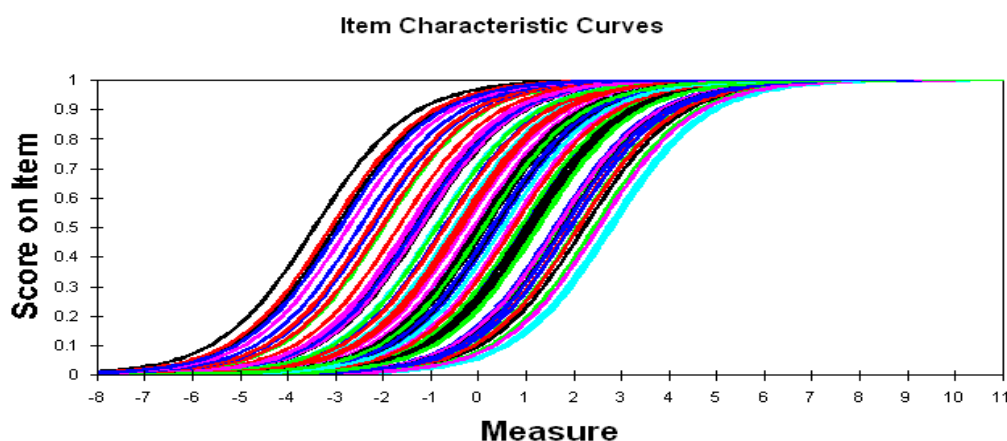


Table 4 showed that the difficulty measures of the 90 items vary between +4.37 and -3.44 logits. Q17 was the most difficult item, whereas Q18 was the easiest one. In the cases where only the expected responses were considered, nearly all of the items (except Q34) could obtain good fit, as indicated by infit MNSQs. However, when the cases of outliers were taken into account, it was found that in total there were 11 items which showed high outfit MNSQs greater than 2.0. These items were: Q2, Q3, Q12, Q17, Q34, Q40, Q42, Q48, Q57, Q64, Q82. An inspection of the misfit items suggested that the pictorial cues were not as clear as the other items of the test and therefore the test-takers could not exactly identify the meanings of misfit pictures.

Finally, regarding the construct validity, the approach – principal components analysis of the standardized residuals – was taken and used in the Winsteps 3.67. The focal interest of this approach is to find out the factors among residuals, not the data itself. So, this approach should not be taken as a usual factor analysis. That is, for a traditional factor analysis such as Principal Axis Factoring, interpretation is based only on positive loadings, whereas in the PCA of Residuals, interpretation has to be based on the contrast between positive and negative loadings. Table 5 showed the results of the PCA of Residuals for the PVST. It was firstly found that the unexplained variance in the 1st contrast was 3.5 eigenvalue units. This

eigenvalue size was slightly above the threshold 3.0 (Linacre, 2008) and therefore may imply the possibility of the 2nd dimension of the data. In the Rasch analysis, each item contributes one eigenvalue unit to the unexplained variance; hence, as the PVST is a 90-item test, the total unexplained variance has a size of 90 eigenvalue units, as indicated in Table 5. As a consequence, the first contrast with 3.5 eigenvalue units has the strength of 3.5 items, occupying only 3.8% of the total 90 eigenvalue units. On balance, the results of construct validity analysis suggested that the PVST was unidimensional and the possibility for a 2nd dimension was low.

Table 4. Summary of the Item Analysis for the Ninety Items

ITEM	MEASURE	COUNT	SCORE	ERROR	IN.MSQ	IN.ZST	OUT.MS	OUT.ZSTD
1	-2.91	242.0	227.0	.30	.91	-.38	.84	-.10
2	-3.10	242.0	229.0	.32	.80	-.86	2.69	2.20
3	-2.73	242.0	225.0	.29	.77	-1.19	2.51	2.06
4	.10	242.0	159.0	.17	.96	-.41	.70	-1.43
5	.51	242.0	145.0	.17	.82	-2.24	.76	-1.27
6	.04	242.0	161.0	.17	.83	-2.07	.63	-1.80
7	-2.50	242.0	222.0	.27	.71	-1.78	.32	-1.54
8	-.46	242.0	177.0	.18	.88	-1.29	.73	-.92
9	-1.45	242.0	203.0	.21	1.06	.56	1.39	.94
10	-.88	242.0	189.0	.19	.77	-2.34	.70	-.80
11	-.92	242.0	190.0	.19	.81	-1.97	.49	-1.63
12	-3.00	242.0	228.0	.31	.97	-.08	4.27	3.40
13	-1.32	242.0	200.0	.21	.94	-.54	.83	-.28
14	-2.30	242.0	219.0	.25	.90	-.58	.39	-1.34
15	.07	242.0	160.0	.17	1.05	.60	.82	-.75
16	-.88	242.0	189.0	.19	.88	-1.19	.54	-1.41
17	4.37	242.0	168.0	.11	1.19	.53	2.16	2.17
18	-3.44	242.0	232.0	.36	.73	-1.04	.20	-1.87
19	-2.50	242.0	222.0	.27	.74	-1.58	.29	-1.65
20	-.70	242.0	184.0	.19	1.07	.70	.85	-.39
21	-1.23	242.0	198.0	.20	1.11	.97	.88	-.15
22	1.23	242.0	119.0	.17	.92	-.92	1.07	.45
23	-1.32	242.0	200.0	.21	1.03	.34	.91	-.08
24	-.02	242.0	163.0	.17	1.03	.37	1.13	.62
25	.28	242.0	153.0	.17	1.00	.00	.91	-.35
26	.73	242.0	137.0	.17	1.03	.35	1.04	.29
27	-.11	242.0	166.0	.18	1.14	1.51	1.01	.11
28	.07	242.0	160.0	.17	.92	-.87	.75	-1.14
29	-.53	242.0	179.0	.18	.99	-.12	1.18	.65
30	.33	242.0	151.0	.17	.95	-.63	.87	-.59
31	.39	242.0	149.0	.17	.81	-2.33	.61	-2.16
32	-.37	242.0	174.0	.18	.78	-2.50	.51	-2.10
33	-.02	242.0	163.0	.17	.95	-.54	.90	-.35
34	1.26	242.0	118.0	.17	2.16	9.90	3.34	8.09
35	-.46	242.0	177.0	.18	.79	-2.38	.52	-1.91
36	.16	242.0	157.0	.17	1.21	2.33	1.25	1.11
37	-2.23	242.0	218.0	.25	1.12	.80	.95	.08
38	-2.05	242.0	215.0	.24	1.09	.67	.92	.01
39	-1.49	242.0	204.0	.21	1.13	1.06	.86	-.17
40	-2.05	242.0	215.0	.24	1.00	.04	3.88	3.51
41	-1.28	242.0	199.0	.20	.97	-.19	.79	-.39
42	-1.36	242.0	201.0	.21	1.25	2.00	2.67	2.84
43	-1.40	242.0	202.0	.21	.99	-.01	1.03	.21
44	-.40	242.0	175.0	.18	1.13	1.38	1.23	.86
45	.65	242.0	140.0	.17	.95	-.57	.82	-.97
46	3.03	242.0	215.0	.23	.82	-1.28	.62	-1.16
47	-.08	242.0	165.0	.18	1.02	.20	1.39	1.53
48	-1.36	242.0	201.0	.21	.98	-.12	2.40	2.49
49	-.74	242.0	185.0	.19	.72	-3.01	.47	-1.87
50	-1.68	242.0	208.0	.22	.75	-2.12	.39	-1.53
51	.01	242.0	162.0	.17	.73	-3.30	.53	-2.37
52	-1.36	242.0	201.0	.21	.93	-.57	.81	-.34
53	.51	242.0	145.0	.17	1.10	1.25	1.15	.82
54	-1.19	242.0	197.0	.20	.81	-1.74	.61	-.97
55	1.79	242.0	99.0	.17	1.14	1.60	1.37	1.63
56	-2.05	242.0	215.0	.24	.88	-.79	.86	-.12
57	2.52	242.0	74.0	.18	1.21	2.14	2.41	3.92
58	.73	242.0	137.0	.17	.89	-1.32	.72	-1.63
59	.39	242.0	149.0	.17	1.24	2.69	1.28	1.34
60	1.76	242.0	100.0	.17	.97	-.38	1.65	2.65
61	.65	241.0	139.0	.17	.87	-1.62	1.00	.06
62	-1.36	242.0	201.0	.21	.97	-.19	.57	-1.04
63	.01	242.0	162.0	.17	.88	-1.38	.72	-1.24
64	2.43	242.0	77.0	.17	1.03	.36	3.66	6.37

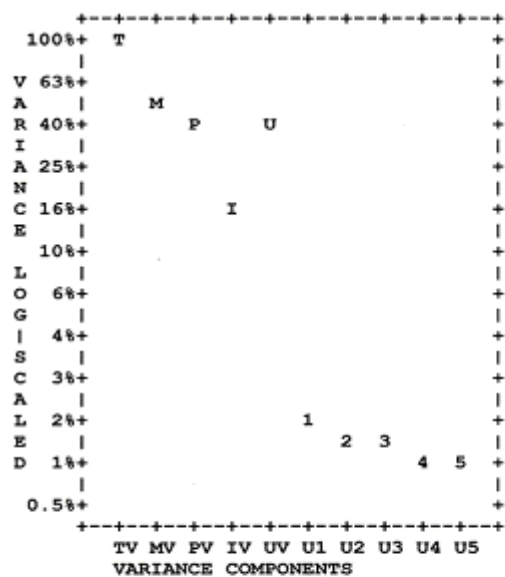
Table 4. (continued) Summary of the Item Analysis for the Ninety Items

ITEM	MEASURE	COUNT	SCORE	ERROR	IN.MSQ	IN.ZST	OUT.MS	OUT.ZSTD
65	2.71	242.0	68.0	.18	.93	-.70	.72	-.99
66	.98	242.0	128.0	.17	1.20	2.37	1.19	1.04
67	.16	242.0	157.0	.17	.97	-.29	.93	-.27
68	-.46	242.0	177.0	.18	.64	-4.26	.42	-2.49
69	-.30	242.0	172.0	.18	.86	-1.52	.77	-.82
70	.87	242.0	132.0	.17	.83	-2.14	.68	-1.95
71	.33	242.0	151.0	.17	1.13	1.50	.99	.03
72	1.01	242.0	127.0	.17	1.07	.85	1.23	1.26
73	1.93	242.0	94.0	.17	1.14	1.56	1.23	1.03
74	2.10	242.0	88.0	.17	1.27	2.85	1.33	1.36
75	.16	242.0	157.0	.17	.83	-2.02	.67	-1.65
76	2.07	242.0	89.0	.17	.95	-.61	1.28	1.21
77	2.78	242.0	66.0	.18	.97	-.30	1.57	1.76
78	2.22	242.0	84.0	.17	.70	-3.69	.55	-2.20
79	.90	242.0	131.0	.17	.89	-1.38	1.09	.55
80	1.90	241.0	95.0	.17	.70	-3.91	.71	-1.41
81	1.64	242.0	104.0	.17	1.12	1.43	1.40	1.81
82	-.08	241.0	164.0	.18	1.17	1.84	2.02	3.33
83	2.81	242.0	65.0	.18	1.13	1.32	1.61	1.83
84	1.17	242.0	121.0	.17	1.22	2.57	1.58	2.76
85	1.59	242.0	106.0	.17	.87	-1.55	.98	-.03
86	.31	242.0	152.0	.17	.86	-1.65	.66	-1.79
87	.95	242.0	129.0	.17	.80	-2.58	.70	-1.81
88	1.67	242.0	103.0	.17	1.35	3.76	1.43	1.90
89	.73	242.0	137.0	.17	1.17	2.03	1.31	1.59
90	1.09	242.0	124.0	.17	1.02	.22	1.02	.19

Table 5. The Principal Components Analysis of Residuals (in Eigenvalue units)

		Empirical		Modeled
Total raw variance in observations	=	215.1	100.0%	100.0%
Raw variance explained by measures	=	125.1	58.2%	58.0%
Raw variance explained by persons	=	88.1	41.0%	40.8%
Raw Variance explained by items	=	37.0	17.2%	17.2%
Raw unexplained variance (total)	=	90.0	41.8%	100.0% 42.0%
Unexplained variance in 1st contrast	=	3.5	1.6%	3.8%
Unexplained variance in 2nd contrast	=	3.0	1.4%	3.4%
Unexplained variance in 3rd contrast	=	3.0	1.4%	3.3%
Unexplained variance in 4th contrast	=	2.7	1.3%	3.0%
Unexplained variance in 5th contrast	=	2.6	1.2%	2.9%

Figure 2 Variance Component of Screenshot Plot

**Note:**

T, TV: Total variance in the observations, always 100%
M, MV: Variance explained by the Rasch measures
P, PV: Variance explained by the person abilities
I, IV: Variance explained by the item difficulties
U, UV: Unexplained variance
1, U1: First contrast (component) in the residuals
2, U2: Second contrast (component) in the residuals, etc.

DISCUSSION

Where the relationship between lexical knowledge and language learning is concerned, Laufer (1986) contends that “the learning of vocabulary lies at the heart of language learning” (p. 69). Meara (1996) also suggests that “vocabulary skills make a significant contribution to almost all aspects of L2 proficiency” (p. 15). Indeed, acquiring sufficient lexical knowledge is a fundamental step towards a full mastery of language learning. However, a reliable and valid vocabulary knowledge measure is necessary in order to pinpoint exactly the level of learners’ development. Otherwise, suitable instructions can not be offered.

The Ministry of Education in Taiwan has compiled a 1200-word list for primary and junior high school learners to follow. It is clear that the language policy makers in Taiwan have considered lexical knowledge an important language element in both primary and secondary English education. However, it is noted that so far few attempts are made to develop a reliable and valid vocabulary size test on the basis of the word list. To be specific, none of the currently available vocabulary size tests referring to the word list are developed under the Rasch model. To bridge the gap, this study marks the first attempt to develop an unbiased, item-oriented and probability-based pictorial vocabulary size test for both the primary and junior high school learners in Taiwan. A successful development of this Rasch-based vocabulary size test has important implications for theory, instruction and the society as a whole. From a theoretical standpoint, it is informative to see the Rasch modeling can be successfully applied to model this newly-conceived vocabulary measurement. Both of the person and item reliability indices are high and fairly satisfactory. The whole test can successfully tap into the trait space that is targeted. Hence, the Pictorial Vocabulary Size Test can be considered reliable and valid in a strict sense.

Pedagogically, because “without adequate lexis there is no proper language competence or performance” (Laufer, 1986, p. 70), the vocabulary size test developed can be used to diagnose or identify both Taiwanese primary and junior high school learners’ strengths and weaknesses regarding the words in the list. Then, appropriate and effective vocabulary

learning programs can be subsequently established to either strengthen or consolidate learners' understanding of the word list.

Effective vocabulary learning programs should have to take learnability into account. English is a language of large vocabulary size (Nation & Meara, 2001). Besides, in addition to knowing the basic meaning of a word, there are still other aspects of word knowledge to be aware of (Nation, 2001). Thus, these characteristics of English vocabulary have inevitably made English vocabulary learning a demanding learning task. Laufer (1997) has pointed out a number of intralexical factors that may affect the learnability of a word: pronounceability, orthography, length, morphology, synformy, part of speech, abstractness, specificity/register restrictions, idiomaticity, and multiplicity of meaning. Facilitating effects can be found while one of the following factors emerges: 1) familiar phonemes, 2) phnontactic regularity, 3) fixed stress, 4) consistency of sound-script relationship, 5) inflexional regularity, 6) derivational regularity, 7) morphological transparency, 8) generality, 9) register neutrality and 10) one form for one meaning. Arguably, Laufer's findings have revealed that not only target language itself but also mother-tongue of language learners may have effects on the learnability of English vocabulary learning. If a word has the properties of phonotactic regularity, fixed stress, consistency of sound-script relationship, inflectional regularity, derivational regularity, morphological transparency, generality, register neutrality, or one form for one meaning, then the learning burden of that word will be light. English instructors should, therefore, have to consider the effects of aforementioned intralingual factors of English on learning the 1200-word list.

More importantly, concerning the social inequality, it is particularly noted that there are many foreign bride mothers in Taiwan. Their social status tends to be relatively minor, and their income appears to be comparatively low and unstable in contrast to the mothers who are not foreign in origin. Their children are a special but now sizable group of members in Taiwan society. It is suggested that more items and parallel test forms of the test should be further established, because the study as such will be no doubt of great importance in laying the groundwork for understanding the extent to which the children of foreign bride mothers acquire the words of the list in comparison to their peers in both cross-sectional and longitudinal study. For instance, if several parallel test forms can be developed as part of the whole test development, this feature will enable teachers as well as vocabulary researchers to track the trajectory of vocabulary size from the beginning of the primary education to the completion of the junior high education. The results of this kind of longitudinal study will be definitely helpful for an even clearer understanding of the way in which the two cohorts develop their vocabulary size knowledge over time. If any significant gaps are detected, then the results are valuable in the sense that a systematic and tailor-made intervention program can be implemented in time.

REFERENCES

中文部分:

教育部國教司 (2007). 「英語基本 1000 字」為國中基測範圍命題.

教育部國教司 (2006, March 10) [國民中小學九年一貫語文科課程綱要](http://teach.eje.edu.tw/9CC/context/03-1-5.html) Retrieved December 30, 2008, from <http://teach.eje.edu.tw/9CC/context/03-1-5.html>

英文部分:

Adolphs, S. & Schmitt, N. (2003). Lexical coverage of spoken discourse. *Applied Linguistics*, 24, 425-438.

Bachman, L. F. (1990). *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.

Fitzpatrick, T. (2007). Productive vocabulary tests and the search for concurrent validity. In H.

- Daller, J. Milton, & J. Treffers-Daller (Eds.). *Modelling and Assessing Vocabulary Knowledge*. Cambridge: Cambridge University Press.
- Hazenbergh, S. & Hulstijn, J. H. (1996). Defining a minimal receptive second language vocabulary for non-native university students: An empirical investigation. *Applied Linguistics*, 17, 145-163.
- Laufer, B. (1986). Possible changes in attitude towards vocabulary acquisition research. *IRAL*, 24, 69-75.
- Laufer, B. (1988). What percentage of text-lexis is essential for comprehension? In C. Laurén & M. Nordmann (Eds.), *Special language: From humans to thinking machines* (pp. 316-323). Clevedon: Multilingual Matters.
- Linacre, J. M. (2008). *A User's Guide to WINSTEPS MINISTEP: Rasch-Model Computer Programs*. Chicago.
- Meara, P. (1983). Word associations in a foreign language. *Nottingham Linguistics Circular* 11, 29-38.
- Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjær & J. Williams (Eds.), *Performance and Competence in Second Language Acquisition*. Cambridge: Cambridge University Press.
- Nation, I. S. P. (1990). *Teaching and Learning Vocabulary*. Boston, Mass: Newbury House.
- Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Nation, P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt, & M. McCarthy. (Eds.), *Vocabulary: Description, acquisition, and pedagogy*. Cambridge: Cambridge University Press.
- Qian, D. D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review*, 56, 282-308.
- Qian, D. D. (2002). Investigating the relationship between knowledge and academic reading performance: An assessment perspective. *Language Learning*, 52, 513-536.
- Wesche, M., & Paribakht, T. S. (1996). Assessing second language vocabulary knowledge: depth vs. breadth. *Canadian modern Language Review*, 53, 13-39.
- Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing* 10, 355-371.
- Read, J. (2000). *Assessing Vocabulary*. Cambridge: Cambridge University Press.
- Ryan, A. (1997). Learning the orthographical form of L2 vocabulary – A receptive and a productive process. In N. Schmitt, & M. McCarthy. (Eds.), *Vocabulary: Description, acquisition, and pedagogy*. Cambridge: Cambridge University Press.
- Schmitt, N. (2004). (Ed.) *Formulaic Sequences: Acquisition, Processing and Use..* Amsterdam: John Benjamins Publishing Company.
- Schmitt, N., & Meara, P. (1997). Researching vocabulary through a word knowledge framework. *Studies in Second Language Acquisition*, 19, 17-36.
- Tseng, W. T. (2008). Optimizing the Response Categories of a SLA Ordered Rating Scale through Rasch-Andrich Rating Scale Model. *Journal of Applied English*, 1, 59-87.
- Tseng, W. T., & Schmitt, N. (2008). Towards a model of motivated vocabulary learning: A structural equation modeling approach, *Language Learning* 58, 357-400.
- Tseng, W. T., Dörnyei, Z., & Schmitt, N. (2006). A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Applied Linguistics*, 27, 78-102.
- Van Hout, R., & Vermeer, A. (2007). Comparing measures of lexical richness. In H. Daller, J. Milton, & J. Treffers-Daller (Eds.). *Modelling and Assessing Vocabulary Knowledge*. Cambridge: Cambridge University Press.

- Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics*, 22, 217-234.
- Wray, A. (2002) *Formulaic Language and Lexicon*. Cambridge: Cambridge University Press.

Advertisement, Critical Thinking and ESP Writing

Wei-fung Tso (左偉芳)
National Taipei University
wftso@mail.ntpu.edu.tw

This paper reports the findings of a study examining the effects of using advertisements to enhance the teaching of critical thinking and ESP writing skills. The pedagogical purpose of incorporating advertisements in an ESP writing course is two-folded: to promote the learning of content knowledge needed for business writing and to help develop critical thinking skills at different levels of sophistication, particularly the skills of synthesis, inference and evaluation. A five-week syllabus was designed and implemented in a college Business Writing class for junior English majors. Students are given a series of writing assignments in which they describe, analyze, and evaluate advertisements in simulated business contexts. Both quantitative and qualitative data was collected via class assignments, a critical thinking assessment instrument and a questionnaire. The findings revealed significant gains in content knowledge mastery and critical thinking skills. Improvement in students' writing skills was also found. The questionnaire results showed positive responses toward this method of instruction. Students reported that advertisements made the writing assignments more interesting and challenging.

INTRODUCTION

Taiwanese students frequently struggle with writing Business English essays as they are expected to evaluate business contexts provided in class and express their thoughts logically and clearly. Essays written for business purposes are often persuasive and evaluative in nature, but the test-oriented instructional process generally adopted in high schools of Taiwan urges students to receive ready-made information without questioning or thinking critically. This is why when students enter college they have problems communicating effectively through writing, especially in Business English writing classes. Their passive learning style is transferred to ESP classes (Dahmeroglu and Vanci-Osam 2005). Another reason that attributes to this problem is related to students' proficiency in content knowledge. Content knowledge is essential to the success of an ESP curriculum (Gatehouse 2001; Brennan and van Naerssen 1989). Without receiving adequate instruction in content knowledge, it is difficult for students to write effectively in specific contexts for specific purposes. To help EFL college English majors to improve their Business English writing performance, this study investigates the effects of implementing advertisement in a college Business English writing class and intends to address the following questions:

1. Does ad analysis help students acquire content knowledge necessary for writing an evaluative business report?
2. Does ad analysis help students employ critical thinking in writing?
3. What aspect(s) of critical thinking is(are) improved?
4. What are the students' attitudes toward the use of advertisement in an ESP writing curriculum?

ADVERTISEMENT AND CRITICAL THINKING

Advertisements of various forms, ranging from printed ads and billboards to radio and TV commercials, lend themselves well to the training of critical thinking. They mobilize multiple semiotic resources, namely verbal, visual, and audio resources and employ complicated interaction between image, sound and text to achieve desired effects (Cook,

1992). Analyzing a complex production such as an advertisement and writing an evaluation report on it can offer ESP students a myriad of opportunities to think critically in the writing process.

Analyzing ads is particularly effective to train critical thinking skills among Asian students. Alagozlu (2007) reported that Asian students do not display critical thoughts in their writing in English either because they are not trained properly or because they lack the confidence of voicing their own ideas critically. Asian students tend to agree with whatever views presented by the textbook and seldom challenge claims made by others. However, advertisements don't usually come with reviews. When students are asked to evaluate an ad, they have no established discussion to rely on but to express their own opinions. By examining the strengths and weaknesses of a multi-facet production students are forced to think about the various sides of an issue critically. Furthermore, advertisements engage students. It works as authentic materials that encourage students to think independently and critically in a low-pressured environment.

METHODOLOGY

Participants

Thirteen junior undergraduate students majoring in English of a public university in northern Taiwan participated in this study. They received two hours of Business English writing instruction each week for five weeks during the time the study took place. None of the students had received any instruction on the topics of Advertisement Analysis or Business Writing before the study.

Instruction of the Course

The instruction of the training session was divided into two stages: ad analysis and simulated business meeting. A description of each stage is presented as follows:

Ad Analysis

The ad analysis stage had two major objectives: to familiarize students with the content knowledge and technical jargons necessary for the task of analyzing advertisements and to help students apply critical thinking skills to writing.

To address the first objective, students were given lectures on how advertising works. Pateman's (1983) principles were adopted as the basis of discussion. Key concepts introduced included identifying the target consumers, describing different parts of an ad, analyzing the sales pitch, detecting the subtext, and understanding the purpose of an advertisement. Both printed ads and T.V. commercials were used as input materials for analysis and discussion.

To foster students' critical thinking as well as writing skills, students were asked to answer essay questions and write short texts that allowed them to practice analyzing facts, organizing ideas, defending opinions, drawing inferences, evaluating input materials, and generating conclusion. The practice of using language necessary for the discussion of logic, deduction and induction was also introduced at this stage.

The analysis stage took three weeks of the total instruction time.

Simulated Business Meeting

Researchers in Business English teaching have long proposed that Business English should be taught in life-like situations via communications that imitate or model genuine business communications (Ellis & Johnson 1994). In light of this, a simulated business meeting is designed for the course. Students were first introduced the important parts of a business meeting: opening a meeting, voicing opinions, responding to questions, closing a meeting, and chairing a meeting. Then, they were asked to conduct an informational meeting in which each student presented an analysis of an ad; whereas other participants raised questions and added comments to the analysis.

Business meetings provide a natural framework for the training of critical thinking because of its inherent dynamics. Unlike presentations, meetings force students to not just focus on the details of a given topic, but also to interact with other students. Interruptions and verbal exchanges in business meetings help students develop arguments supported by logic and evidence; hence, creating an ideal situation for students to practice critical thinking skills.

Data Collection

The data of this study was collected both quantitatively and qualitatively. Quantitative data was gathered from the results of the pre- and post-test essays that students were asked to write. They were given fifty minutes to write an essay discussing the effectiveness of a printed advertisement. The same magazine ad was used for both pre- and post-tests.

Another set of quantitative data was collected from an evaluation questionnaire filled out by the students at the end of the training period. The questionnaire contained four sections. The first section surveyed students' perceived effects on the learning on content knowledge from ads. The second part investigated the effects of the ad analysis instruction on the learning of critical thinking skills. The third section intended to find out whether the instruction helped students to improve writing in general. And the last part of the questionnaire was concerned about students' interest and motivation in doing the ad analysis activities.

Qualitative data was collected mainly from the open-ended questions at the end of the questionnaire, aiming to elicit feedback on the effectiveness of the course and the strengths and weaknesses they perceived regarding the ad-analysis syllabus.

Data Analysis

Analysis of Pre- and Post-test Essays

The effects of the ad analysis instruction on writing ESP essays and learning critical thinking skills were examined according to the results of the pre- and post-tests and the evaluation questionnaire. The pre- and post-tests were scored by two trained raters and the inter-rater reliability score was .76 for the pre-test; .74 for the post-test on Kendall's W. Based on the principles of ICAT Critical Thinking Essay Examination (1996), the raters were asked to evaluate the essays in terms of the content knowledge displayed and the frequency as well as the depth of critical thinking skills applied. The following six items were rated on a 10-point scale: 1) description 2) analysis 3) synthesis 4) evaluation 5) inference and 6) content knowledge. In addition, the content, organization, vocabulary, language use, and grammar of the essays were also graded according to the tenets of ESL Composition Profile (Jacobs, Zingraph, Wormuth, Hartfiel, & Hughey, 1981).

Analysis of the Questionnaire

The questionnaire was analyzed both quantitatively and qualitatively. Except for the last open-ended question, students' answers to the questionnaire were calculated according to the percentage chosen for each survey item. The result of the open ended question was given a qualitative analysis to reveal students' feedbacks and attitudes towards the instruction of the ad-analysis curriculum.

RESULTS AND DISCUSSION

This section presents the findings obtained from both the quantitative and qualitative data and tries to answer the research questions raised in this study: whether the ad analysis training instruction helps students acquire content knowledge in advertising, apply critical thinking into writing, and improve their writing skills in general.

Comparison of Critical Thinking Skills between the Pre-test and the Post-test

Table 1. *Descriptive Statistics and T-Tests on the Group's Pre- and Post Tests Performance of Content Knowledge and Critical Thinking Skills*

(n=13)	Pre-test	Post-test	t-value
Description	6.54 (.59)	7.73 (.90)	-3.48**
Analysis	6.35 (.98)	7.19 (.75)	-2.38**
Synthesis	6.62 (.77)	7.96 (.83)	-4.82**
Evaluation	6.12 (1.26)	7.50 (.86)	-3.26**
Inference	6.35 (1.07)	8.08 (1.08)	-4.14**
Content knowledge	4.96 (3.11)	7.73 (.90)	-2.98**
All	6.15 (1.59)	7.70 (.91)	-7.30**

**p<.05

Table 1 shows that significant improvement was found in all the targeted critical thinking skills: description ($t=-3.48$, $p<.05$), analysis ($t=-2.83$, $p<.05$), synthesis ($t=-4.82$, $p<.05$), evaluation ($t=-3.26$, $p<.05$), and inference ($t=-4.14$, $p<.05$). Students also made significant improvement in the learning of content knowledge ($t=-2.98$, $p<.05$).

Table 2. *Line Graph on the Group's Pre- and Post Tests Performance of Content Knowledge and Critical Thinking Skills*

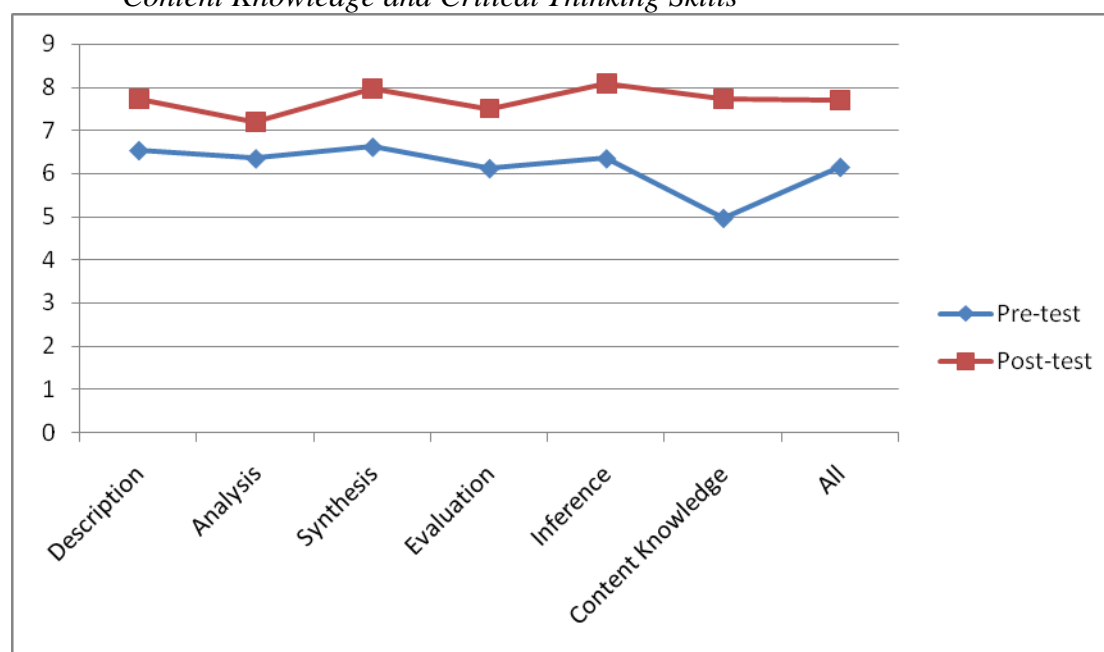


Table 2 shows that students made greater improvement in acquiring content knowledge than critical thinking skills as the result of the instruction. Content knowledge received the lowest score in the pre-test and exhibited the highest range of improvement in the post-test. The high jump in test result is probably due to students' being exposed to content knowledge in the training session versus not being trained on the subject ever before. The pre-test low score in content knowledge is due to students' lack of the vocabulary as well as the concept necessary for ad analysis. This study shows that content knowledge in ESP courses can be learned at an efficient rate as short as five weeks.

On the other hand, critical thinking skills are being applied in many other subject areas

(Tsui 1999) and students are already familiar with the skills. Therefore, the result of extra training in critical thinking yields less dramatic improvement compared to the improvement shown from content knowledge. However, the critical thinking skills that the students had been exposed to included both low level and high level skills (Bloom 1956). The test results showed greater improvement gained in the high level skills, including synthesis, evaluation and inference. This implies that the ad analysis syllabus for this study is more effective in the high level critical thinking categories. Table 2 clearly indicates that ad analysis instruction helped improve students' learning of critical thinking skills.

It is also suggested in Table 2 that a correlation existed between content knowledge and critical thinking skills in reference to writing. When students gain in content knowledge, they also improved on critical thinking. Similar finding was reported by Stapleton (2001) with his Japanese undergraduate subjects that the quality of critical thought depends on the topic content. With a familiar topic, Japanese students were able to generate better critical thinking in their L2 production. This result suggests that ESP practitioners need to provide solid instructions on content knowledge in order to facilitate students' writing in contexts that require the application of various levels of critical thinking, such as Business Writing.

Table 3. *Comparison of Students' Overall Writing Performance between the Pre-test and the Post-test*

(n=13)	Pre-test	Post-test	t-value
Content	35.19 (3.50)	41.19 (3.00)	-7.23**
Organization	33.62 (3.52)	40.81 (3.15)	-8.30**
Total	68.81 (6.96)	82.00 (6.08)	-7.93**

**p<.01

Besides content knowledge and critical thinking skills, students' overall writing performance was also evaluated. To simplify the grading process, vocabulary, language use and grammar were graded under the category of content; whereas the category of organization includes features such as topic sentence, supporting ideas, and transitional phrases used in an essay. Table 3 showed that students made very significant improvement in both the content and organization categories, resulting in a very significant advancement made in their overall writing performance.

Improvement shown in the content category may be attributed to the students' applying knowledge they acquired on advertising to the post-test essay. The training in critical thinking skills; however, is probably the main reason why students improved significantly in the organization of their essays. This is in line with Bahous' (2001) and Willis' (1992) findings that critical thinking helps students think clearly about their claims and build logical arguments. Consequently, their essays are better organized.

Results of the Questionnaire

Table 4. *Result of the Survey Questions (N=13)*

The lecture, discussion and activities of the unit help me to:	Strongly Agree	Agree	Unsure	Disagree	Strongly Disagree
1. identify what kind of an ad it is when I see one (e.g. service, product, public service...)	8 (62%)	5 (38%)			
2. identify different parts of an ad	5 (38%)	8 (62%)			
3. identify the target consumers of an ad	9 (69%)	4(31%)			
4. use jargons to describe and talk about an ad (e.g. "captions", "sales pitch"...)		8 (62%)	5 (38%)		

5. analyze how an ad captures consumers' attention	7 (54%)	6 (46%)			
6. analyze how an ad build consumers' desire or confidence	4 (31%)	9 (69%)			
7. understand how different parts of an ad work together to convey messages	7 (54%)	4 (31%)	2 (15%)		
8. know how to use various materials/information to prove my point	2 (15%)	5 (38%)	6 (46%)		
9. make judgments on the effectiveness of an ad	3 (23%)	5 (38%)	5 (38%)		
10. understand how ads influence us	7 (54%)	6 (46%)			
11. understand the subtext	3 (23%)	7 (54%)	3 (23%)		
12. pay attention to relevant details when I write	4 (31%)	8 (62%)	1 (8%)		
13. organize my ideas in an essay	4 (31%)	6 (46%)	2 (15%)	1 (8%)	
14. think clearly when I write	3 (23%)	7 (54%)	3 (23%)		
15. improve my writing in general	3 (23%)	5 (38%)	4 (31%)	1 (8%)	
16. I think what I learn in this unit is interesting	8 (62%)	5 (38%)			

Table 4 displayed the results of the questionnaire on the perceived effect of the ad analysis instruction. The questionnaire was concerned about four major issues: the acquisition of content knowledge (Questions 1-4); the learning of critical thinking skills (Questions 5-11); the improvement on writing skills (Questions 12-15); and students' interest in the course (Question 16).

The results of Questions 1-4 showed that students either strongly agreed or agreed that the course helped them learn content knowledge related to advertisement. However, Question 4 revealed that five students (38%) were not sure if they were competent enough to put the knowledge into application. This suggests that either the students needed more instruction on content knowledge or they needed more practice on how to apply content knowledge to classroom discussion and essay writing. Future programs should strengthen the instruction as well as the application of content knowledge to address this issue.

As for critical thinking skills, Questions 5 and 6 investigated analytical skills; Questions 7 and 8 covered synthesizing abilities; Question 9 dealt with evaluation skills; and Questions 10 and 11 examined the ability of making inferences. The survey result showed that students generally agreed that the training session was helpful to them in the learning of all four critical thinking skills, particularly in the areas of making analysis and inferences. However, about 1/3 of the students were not sure whether the training session helped them improve their abilities to synthesize and evaluate. This result does not completely correspond to the findings obtained from the quantitative data. As mentioned in the previous section, the test results showed that the students made the least improvement in their analytical skills whereas they improved the most in making inference and evaluation. The discrepancy between learners' perception and their actual performance on the analytical ability is probably due to the reason that they are more familiar with the skill; therefore, they are more confident in their performance in this category. As to the difference regarding other critical thinking skills, further research needs to be done to provide a clear explanation.

When asked whether the ad analysis classes were helpful to their thinking process when they write, most students gave positive answers in this respect. They believed that it helped them to organize their ideas better (77%); think more clearly (77%) when they write; and help them to improve writing in general (61%). However, Question 16 also showed that 39% of the students were not sure whether their writing had improved or not. This may be

due to the fact that the training session was too short to allow the students to detect any difference in their writing performance.

Finally, all students agreed either very strongly (62%) or strongly (38%) that the ad analysis training program was interesting.

Results of the Open-ended Question

In the open-ended question of the questionnaire students discussed the strengths and weaknesses of the ad analysis training session. Their comments on the advantages of the instruction were generally directed to four areas: content knowledge, critical thinking, writing skills, and the simulated business meeting activity.

Most students expressed that the instruction helped them acquire content knowledge and guided them to look at advertisements from a more “professional” angle. They learned to pay attention to different parts of an ad, understood how the parts work together to achieve a marketing purpose, and were able to identify the various sales pitch employed in an advertisement. Some of their comments in this respect were:

“By teaching how to analyze the details of an ad, the course shows us the details which we may not pay attention before. It helps us know more jargons in the field of ad so we can know more about ad.”

“I can read more messages than usual. I understand that every detail of the advertisement is important.”

“It helps me to notice some details, such as blanks and figures in the ad.”

“The part of ‘sales pitch’ makes me surprised. I started to think about what is my ‘button’.”

The second advantage mentioned by the students was regarding critical thinking skills. Some students thought that analyzing ads helped them think more critically and deeply about advertisements.

“It lets me thin deeper into the ad, not merely feel touched or interested. And I can understand the messages conveyed by ad by using the skills learned in class better than before.”

“It shows me that there are some advertisements which contain deeper meaning and it makes me think much critically.”

“I can think critically when I write.”

Most students reflected that the training session helped them learn to better structure their ideas while writing. They were able to develop their arguments more effectively and write with better organization.

“For my writing skills, I think it makes me organize many materials and supporting ideals well. And lets me know how to be coherent.”

“I know how to write coherently and relevantly, and my thoughts would

“It help me think and write in a more organized way.”

“I realize that it’s hard to write a coherent ad analysis, but I’ve known some writing strategies.”

The last advantage brought up by the students was the simulated business meeting. Students found that it was challenging and inspiring because it not only pushed them to think clearly about their own ad analysis but also provided them with the opportunity to observe and learn from other students’ analysis. In addition, the formal setting of a business meeting allowed students to practice using English in real-life situations. As one student put it: “It was not only interesting and motivation-encouraging but also useful. I learn a lot about what to say and how to do in business meetings.”

As to the drawbacks of the instruction, a couple of students responded that writing the ad analysis essay was difficult. They expressed frustration over not knowing how to apply the content knowledge they learned into writing. Another significant problem pointed out by the

students was the length of the training session. They thought that “the course is too short” and they needed more time to do more practice on writing ad analysis essays. Most students did not mention any drawbacks in their survey. However, when they did, it seemed that their critiques were more related to writing apprehension than the training course itself.

In general, the results of the open-ended question correspond closely to the findings established by the quantitative measures. Students reported that the training session was helpful to them in three aspects: content knowledge, critical thinking and essay organization. Furthermore, students found that the design of the course was stimulating and useful. Many students wrote “it is a special experience in writing” and they felt “glad and lucky to be able to participate in the activities.”

CONCLUSION AND IMPLICATION

The result of this study shows that incorporating ad analysis instruction in a college ESP writing course benefits learners in at least four ways: (1) it helps them to acquire content knowledge in Advertising; (2) it fosters critical thinking skills; (3) it improves the organization of their writing; and (4) it creates meaningful contexts for learning to take place. These findings confirm that effective learning takes place when class activities are manifested in forms of manageable tasks (Jacobs et al., 1996).

The result of this study does not provide substantiate evidence to support the claim that there are different levels of critical thinking skills (Bloom 1956), however, it confirms that there are different components of critical thinking. And the ad analysis instruction designed for the study is particularly helpful for the learning of synthesizing, evaluating and making inferences. In terms of application to ESP writing, learners are found to improve most in the organization of their essays when critical thinking skills are applied.

Content familiarity is found to play a significant role in facilitating both critical thinking and the writing of ESP essays. Content knowledge not only serves as a basis for discussion in ESP courses but also helps foster the learning of critical thinking and writing skills. The findings of this study suggest that Business English writing teachers need to provide solid instructions on business concepts in order to allow students to exercise effective critical thinking as well as writing skills.

The design of the ad analysis curriculum is found to be effective in that it creates meaningful contexts for students to acquire content knowledge, reinforce content knowledge acquired and stimulate critical thinking to achieve communicative goals. Students’ positive response to the designed curriculum suggests that ESP teachers should adopt a variety of pedagogical options that both involve learners in frequent, natural writing tasks and engage them in stimulating activities that arouse their interest to learn.

Although the current study provided positive evidence of incorporating ad analysis instruction in the teaching of ESP writing, there are some limitations. First, our use of simulated business meetings as a measure of instruction may not be applicable to all EFL college students, particularly to those whose English proficiency is at a lower level. Second, the sample size of the study is very small and is limited to students who enrolled in the same program of the same university. A larger sample size and a more diversified sample body are needed to test the generalization of the result of the study.

REFERENCES

- Alagozlu, Nuray (2007). Critical thinking and voice in EFL writing. *Asian EFL Journal*, 9(2).
- Bahous, J. (2001). Teach ESP through critical thinking and problem-solving skills. Paper presented at the Meeting of the International Association of Teachers of English as a Foreign Language, Bielefeld, Germany.
- Brennan, M. and Van Naerssen, M. (1989). Language and content in ESP. *ELT Journal* 43(3),

196-205.

- Bloom, B. S. (1956). *Taxonomy of educational objectives, handbook 1: Cognitive domain*. New York: Longmans Green.
- Cook, G. (1992). *The Discourse of Advertising*. London: Routledge.
- Dahmeroglu, D. & Vanci-Osam, U. (2005). Leading students to critical thinking in tertiary level reading classes: Is it biting off more than one can chew? *Proceedings of the Conference on Quality Teaching for Quality Learning: Education for Critical Thinking*. May 2005, Cluj-Napoca, Romania.
- Ellis, M., & Johnson, C. (2002). *Teaching Business English*. Shanghai: Shanghai Foreign Language Education Press.
- Gatehouse, K. (2001). Key issues in English for Specific Purposes (ESP) curriculum development. *The Internet TESL Journal* 7(10). <http://iteslj.org/>
- ICAT Critical Thinking Essay Examination (1996). The International Center for the Assessment of Thinking. Dillon Beach, CA.
- Jacobs, E. et al. (1996). Cooperative learning: context and opportunity for acquiring academic english. *TESOL Quarterly*, 30(2), 253-80
- Jacobs, H.L., Zingraf, S.A., Wormuth, D.R., Hartfiel, V.F., & Hughey, J. B. (1981). *Testing ESL composition: A practical approach*. Rowley, MA: Newbury House Publishers.
- Pateman, T. (1983). How is understanding an advertisement possible? In *Language, Image, Media*. Edited by Howard Davis and Paul Walton. Oxford: Basil Blackwell, 187-204.
- Rottenberg, A. T. (1991). *Elements of Argument: A Text and Reader*. Boston: Bedford Books.
- Stapleton, P. (2001). Assessing critical thinking in the writing of Japanese university students. *Written Communication* 18(4), 506-548.
- Tsui, L. (1999) Courses and instruction affecting critical thinking. *Research in Higher Education*, 40(2), 185-200.
- Willis, A. S. (1992). Integrating levels of critical thinking into writing assignments for introductory psychology students. Paper presented at the Annual Meeting of the American Psychological Association, Washington, DC.

Effective English Courses for University Students

Wen-li Tsou (鄒文莉)
National Cheng Kung University
wtsou@mail.ncku.edu.tw

With the expansion of the global village, English has become a medium for advanced studies and research. In the higher education institution in Taiwan, there is an increasing tendency for the educators to adopt English-written textbooks or articles as teaching materials. However, most students are overwhelmed by anxiety or frustration when encountered such reading materials. To bridge the gap between the learners' English proficiency and their mainstream courses, this proposal presents a holistic ESP (English for Specific Purposes) program for students of one top university in Taiwan.

In this university, where students from fields of engineering, computer science, theoretical science, and designing form about 65% of the total student body, students will need more than the traditional approach to English training to succeed in their future careers. They need not only English proficiency for general communication or socialization, but more importantly specific language tools for their respective professions in order to perform on-the-job tasks, conduct research projects, pursue advanced studies, and publish their findings.

Need analyses, course evaluation questionnaire, interviews and the official TOEIC test are applied to evaluate the learning effects of the ESP program.

INTRODUCTION

Given the general consensus that maintaining competitive in the age of globalization is the keystone to Taiwan's future economic success, the government has seen the need for developing leaders with global vision in various professional fields. The backbone of such government's endeavor is Ministry of Education's program aimed at strengthening institutions of higher education and key disciplinary areas as a national priority for the 21st century. Aligned with the enhancement effort, National Cheng Kung University (NCKU) has re-visioned and re-designed English programs in the undergraduate level.

The University's Department of Foreign Languages and Literature (FLLD) has initiated the "NCKU Eagle Project" in order to 1) promote the average second language proficiency of college students, and 2) address specific academic and occupational needs in students' English learning. In the past NCKU has required the university's undergraduate students to fulfill the English language requirements by successfully passing Freshman and Sophomore English courses. While it is important to strengthen students' general language proficiency, globalizing academic environment and market place have called for English programs that equip students with language skills to function in English-only classrooms while also preparing them for specific career-based language challenges. In contrast to Freshman English that in functionality belong to English for General Purposes (EGP), courses designed to develop academic and professional skills are often referred to as English for Specific Purposes (ESP).

According to Laurence Anthony, from the early 1960's English for Specific Purposes has grown to become one of the most prominent areas of EFL (English as Foreign Language) teaching today. Its development is reflected in the increasing number of universities offering graduate programs in ESP (e.g. The University of Birmingham, and Aston University in the UK) and in the number of ESP courses offered to foreign students in English speaking countries. There is now an established international journal dedicated to ESP discussion, "English for Specific Purposes: An International Journal," and the ESP SIG groups of the

IATEFL and TESOL are also active at their national conferences.

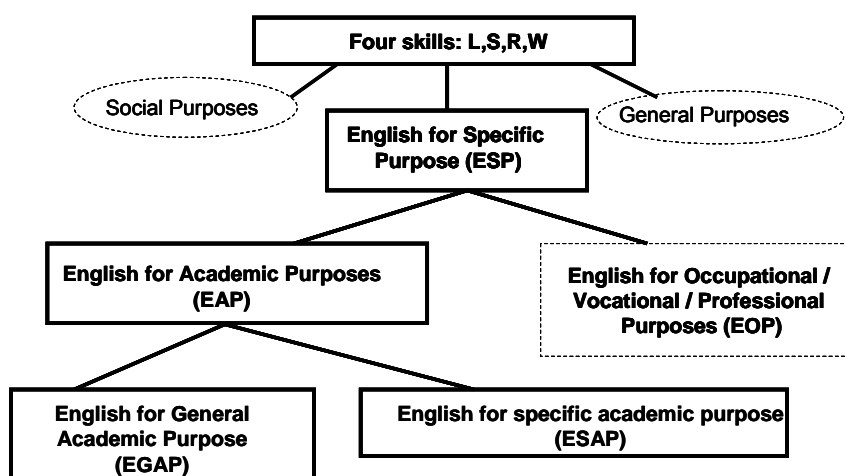
In NCKU, where students from fields of engineering, computer science, theoretical science, and designing form about 65% of the total student body, students will need more than the traditional approach to English training to succeed in their future careers. They need not only English proficiency for general communication or socialization, but more importantly specific language tools for their respective professions in order to perform on-the-job tasks, conduct research projects, pursue advanced studies, and publish their findings.

In our effort to create a learner-based and content-based approach that helps build students' language competence in their specific disciplines, we're very fortunate to receive support and grant from Office of Academic Affairs and The organization of the Office of University Advancement. In addition, we've aligned with field experts such as Bookman Books and Education Testing Service to provide material, funding, and training support. The program currently involves seven full time professors, two full time project teachers and one full time assistant. All these resources together weave a very promising future for the landmark project to build the country's first ESP Teaching Center.

NATURE AND RATIONALE BEHIND NCKU'S ESP PROGRAMS

According to Dudley-Evans, ESP is an approach to language teaching in which all decisions related to content and method are based on learners' reason for learning. ESP courses bridge the gap between learners' English proficiency and their mainstream courses while helping students develop language, study and research skills appropriate for study in a particular college or profession. The primary aim of NCKU's ESP program is thus to present a holistic English learning program for all students.

To better illustrate how ESP relates to the language studies, the following diagram briefly summarizes the various purposes in learning of English:



Diagrams 1: English for Specific Purpose (ESP)

On the top we see the four skills involved in language learning. They're followed by three categories of purposes: General Purposes, Social Purposes, and English for Specific Purpose (ESP). English for General Purposes, as the name suggests, help build learners' general proficiency in four skills. In Taiwan, English classes in middle schools and Freshmen English programs are for general purposes. English for Social Purposes, on the other hand, address language skills for communicational purposes to be used in situations such as greeting, telephoning, shopping, survival English, and etc.

ESP programs, as already mentioned, include EAP (English for Academic Purposes) and EOP (English for Occupational Purposes) courses. The former addresses communicational needs in classrooms where English is the primary (if not the only) language of instruction. EOP courses, on the other hands, are developed with specific careers in mind. Professionals such as engineers, doctors, or airline pilots, for example, need respective language trainings in English so they can acquire vocabulary and learn to express themselves in order to perform on the jobs.

NCKU ESP PROGRAM

The overall project is targeted at advanced language training in a chosen professional area and to fulfill the courses with authentic teaching materials and tasks. Eventually the roll-out program is scheduled to begin in the third year of students' undergraduate studies so that they may be prepared for on-the-job professional challenges. The process of teaching and learning will be monitored by a graduate student and the project teachers. Evaluation for curriculum, teaching effectiveness and student performance will also be conducted at the end of the semester. In addition to in-class achievement test, student' language performance will also be assessed by standardized proficiency test—Test of English for International Communication (TOEIC), the expense to be fully supported by ETS.

The University has introduced 15 sessions of five EOP courses: Career English (three classes); Economics & Trade English (two classes); Business & Management English (two classes); Travel English (two classes); Presentation English (six classes). These courses currently involve 1000 undergraduate students from College of Liberal Arts and College of Management. The pre-program survey and analysis has been administered for these students in March 2008.

Career English is designed to teach students how to employ their speaking ability in the business topics. The course will provide new vocabulary for students and teach them how to apply them in the discussion and topic-related contexts, such as meetings and customer service. Student will sharpen their note-taking skills and know how to response through the situational listening and video clips. They will be given various related reading passages to get the main ideas of the texts and express their opinions in the discussions.

Learning objectives for Career English include basic understanding of career English, acquisition of vocabulary with focus on the work specific terminology, developing strategies for comprehending various text, developing strategies for responding requests and enquiries in contexts, developing note-taking and comprehension skills in the situational listening, using critical thinking of all types of the related texts, designing business events, including preparation, approval and disapproval, and upgrade presentation and negotiation skills in business.

Presentation English, on the other hand, is designed to teach students how to employ their English ability in a number of business topics. Students will be given audio and video clips to understand how a presentation proceeds. By teaching professional strategies and demonstrations, students acquire and perform the presentation skills under authentic conditions, such as business plan, new product development, and project report.

Through personal and group presentations, students will learn to present the overview and organization of presentations, integrate visual aids to capture the audience's attentions, improve verbal and non-verbal expressions in the presentation, organize materials and present figures and charts systematically, and evaluate their own presentations.

Travel English, as the course title suggests, is designed to teach students how to deploy their language ability in the authentic situations during business travel. The course will provide new vocabulary of tourism for students and teach them how to apply them in the discussion and topic-related contexts. Student will sharpen their note-taking skills and know

how to respond through situational listening.

Through related reading and video clips, target language skills will be integrated in class to simulate possible conditions. Students will gain understanding in tourism and improve vocabulary related to business travel. They will learn to design package tour, practice booking information and develop budget, while acquiring strategies for comprehending text and responding to requests and enquiries.

REFERENCES

- Dudley-Evans, T. (1998). *Developments in English for Specific Purposes: A multi-disciplinary approach*. Cambridge University Press.
- Flowerdew, J., & Miller, L. (1995). On the notion of culture in second language lectures. *TESOL Quarterly* 29(2), 345-374.
- Johns, A. M. (1988). The discourse communities' dilemma: Identifying transferable skills for the academic milieu. *English for Specific Purposes*, 7, 55-60.
- Jordan, R. R. (1997). *English for Academic Purposes: A guide and resource book for teachers*. Cambridge: Cambridge University Press.
- Kaewphanngam, C., Broughton, M.M., & Soranasataporn, S. (2002). [Corpus-based analysis: Guidelines for getting practical language input in materials development](#). *SLLT: Studies in Language and Language Teaching*, 11, 16-32.
- Laurence, A. (1997). [English for Specific Purposes: What does it mean? Why is it different?](#) *On-CUE*. 5(3), 9-10.
- McDonough, J. (1984). *ESP in perspective: A practical guide*. London: Collins ELT.
- McDonough, J. (2005). Perspectives on EAP: An interview with Ken Hyland. *ELT Journal* 59, 57-64.
- Snow, M., & Brinton, D. M. (1988). *The adjunct model of language instruction: An ideal EAP framework* (S. Benesch, Ed.), pp. 33-52.

The Historical Review of ETJCEE/DRET Grammar Tests

Hsing-cheng Tung (董幸正)
Chia Nan University of Pharmacy and Science
tung2610@ms13.hinet.net

This paper primarily reviewed the historical developments of grammar tests in the English Tests of Joint College Entrance examinations (ETJCEEs) as well as the Departmental Required English Tests (DRETs), dating from the year of 1956 to the year of 2006. Three major issues were addressed as follows: (1) types of test formats employed to measure the grammar knowledge, (2) the variables to drive the developments of the ETJCEE/DRET grammar test formats, and (3) types of grammar constructs assessed in the ETJCEEs/DRETs. The developments of ETJCEE/DRET grammar test formats would be expounded on the grounds of the language proficiency hypotheses: (1) the divisible competence hypothesis, (2) the unitary competence hypothesis, and (3) the multi-dimensional competence hypothesis. Additionally, the features of pre-scientific testing period would be included for illustrating the first two issues. The grammar constructs assessed in the ETJCEEs/DRETs were defined by the Purpura (2004: 91), in which two major fields received the central focuses: (1) grammatical knowledge, and (2) pragmatic knowledge.

The overall results of this paper were sketched in the following. Indeed, types of ETJCEE grammar test formats were in tandem with the developments of the language proficiency hypotheses, transforming from the decontextualized grammar knowledge in the era of the divisible competence hypothesis to the cohesive devices in the heyday of the integrative approach. Particularly, the budding interest in the unitary competence hypothesis in the 1980s resulted in the drastic changes of the ETJCEE grammar test formats as well as their measured constructs. Despite the innovation of the grammar test formats was sluggish in the DRETs, the cohesive devices outweighed the traditional form-focused knowledge.

INTRODUCTION

This paper primarily reviews the historical developments of the grammar test types employed in the English Tests of Joint College Entrance Examination (ETJCEEs) and the Departmental Required English Tests (DRETs) from the year of 1956 to the year of 2006. For decades, the ETJCEE grammar tests have played a crucial role in the senior high school English classroom instruction, especially in the heyday of structuralism. This was further warranted by newspaper reports and the periodical comments in which the dull lexical memorization and the deductive analysis of grammar rules were still believed as the best bet for the English language study in the mid-1990s (Chang Yang, 1968; Chen Yung-fu, 1969; Chen Yung-sen, 1970; Yeh Kung-chao, 1965). Consequently, students zoomed in on the vocabulary and grammar study, letting their writing and listening abilities slide (Hsueh Wen-lang, 1974: 5). Similarly, through the analysis of ETJCEE test types by Professor Huang Tsan-suei (1997: 5), quite a number of discrete-point grammar tests sprang up in the 1970s, such as (1) the multiple-choice sentence meaning test in the year of 1960, (2) the matching sentential completion test in the year of 1961, (3) the gap-filling insertion test in the year of 1964, (4) the rearrangement test in the year of 1965, (5) the deletion test in the year of 1971, (6) the punctuation test in the year of 1972, (7) the multiple-choice error correction test in the year of 1974, and so forth. These creative grammar tests, virtually, received the impetus from the structural linguistics. Despite Huang (1997: 5) sweated his guts out to classify the ETJCEE grammar tasks in the past two decades, he did not give detailed accounts for the innovation of grammar tests in relation to the testing or teaching rationales. Further, the ETJCEE was promulgated in the year of 1956, but a paucity of information was available

regarding the attributes of ETJCEE test types from the mid-1950s to the early 1970s in Huang's study. Accordingly, the writer of this paper was intended to provide the panorama of the ETJCEE grammar tests from the year of 1956 to the year of 2006. In addition to classifying the ETJCEE grammar test types, the following three major hypotheses of language proficiency provided the references for keeping track of the variables leading to the innovations of ETJCEE grammar test formats, along with their pedagogical implications: (1) the divisible competence hypothesis, (2) the unitary competence hypothesis, and (3) the multi-dimensional competence hypothesis. Additionally, the Grammar-Translation paradigm, which shaped the language instruction and assessment in the early 1990s, were also included to detect the features of ETJCEEs in the 1950s. The pedagogic implications of the ETJCEE grammar test types, as stated earlier, had been widely discussed in the newspapers and related academic reports. However, a dearth of qualitative studies detected the relationships between the grammar testing trends and their pedagogical influences. Thus, the pedagogic impacts of the ETJCEE grammar tests were another intriguing motivation for this research paper.

Other than the classified ETJCEE grammar tasks and their possible pedagogic influences, this research paper laid its another target at the assessed grammar knowledge in this nationwide English entrance examination. Virtually, Chen Ming-shan (2004) carried out the taxonomies of the ETJCEE grammar rules into eighteen types as follows: (1) tenses, (2) aspects, (3) subjunctive mood, (4) infinitives and gerunds, (5) participles, (6) auxiliaries, (7) the agreement of subject and verb, (8) verbs with ambiguous meaning, (9) tag, (10) direct and indirect interrogatives, (11) nouns, (12) articles, (13) pronouns, (14) relative pronouns, (15) adjectives, (16) adverbs, (17) prepositions, and (18) conjunctions. Here, these eighteen categories could be concisely presented as three major elements: (1) the morphosyntactic form, (2) the interactional form, and (3) the information management form (Purpura, 2004: 91). Under the dichotomies of grammar ability by Purpura (2004: 91), the writer of this paper was able to discern the multiple layers of grammar abilities from the year of 1956 to the year of 2006.

On the grounds of the purposes and motivation stated earlier, three research questions were addressed as follows:

- Q1 How many grammar test types have been employed in the ETJCEEs/DRETs from the year of 1956 to the year of 2006?
- Q2 Does the innovation of ETJCEE/DRET grammar test types also bring about the changes in the senior high school classroom English instruction?
- Q3 Which grammar components have been included for the assessment in the ETJCEEs/DRETs in the past five decades?

HYPOTHESIS OF LANGUAGE PROFICIENCY AND THEIR RELATION TO WAYS OF ASSESSMENT

This section reviewed three major types of language proficiency hypotheses and their corresponding theories to language teaching and testing. This review purposed to provide references for addressing the features of ETJCEE/DRET grammar test formats in the follow-up section. The language proficiency, as being noted in the beginning of this section, was defined in terms of three major hypotheses as follows: (1) the divisible competence hypothesis, (2) the unitary competence hypothesis, and (3) the multi-dimensional competence hypothesis. These three hypotheses were elaborated in the discrete subsections that followed. Before illustrating these three language proficiency hypothesis, we sketched the features of pre-scientific period which did not receive the rationale support but still exerted profound influences in the language teaching and testing in the early 1990s.

The Pre-scientific Period & The Grammar-Translation Paradigm

The pre-scientific period, as its name implied, did not receive the rationale support. At that time, the fundamental target for the language teaching and learning was to read the literary passages. Besides, the deductive grammar analysis paved the foundation for the successful performances in the translation and writing. Due to a lack of unified measurement rubrics, foreign language learners' performances were subjectively evaluated by language teachers. The Grammar-Translation Teaching Method and the Essay-Translation Testing Approach were both rooted in the pre-scientific period (also intuitive period). The Grammar-Translation Method laid its focus on the deductive analysis of grammar rules and the lexical memorization which were believed to lay the foundations for the perfect translations of stilted or literary prose. Since the aural/oral communication was of little interest in the Classic Method, the Grammar-Translation Method did not require language teachers to conduct classes in the target language. Therefore, in the heyday of the Grammar-Translation paradigm, mother tongue was allowed as the primary medium of instruction. Finally, literary works were still the predominantly reading materials regardless of the readability levels.

The Divisible Competence Hypothesis & The Structuralism

Since the Grammar-Translation method had nothing to do with improving language learners' oral communication ability, it was gradually replaced by the audiolingualism which received the attention from the 1950s to the 1960s. The structuralism was derived from the divisible competence hypothesis which consisted of various kinds of competence, and each discrete linguistic component was distinguishable (Oller & Perkins, 1980: 14). The structural linguistics was synonymous with the divisible competence hypothesis to some extent. As the structuralists claimed, the language was viewed as a system consisting of discrete linguistic elements which were linearly and pyramidally structured in a rule-governed way (Richards, 2001: 55). In other words, the levels of the linguistic elements began with the phonemic systems which resulted in the morphemic systems. The morphemic systems, in turn, resulted in the higher level systems of phrases, clauses, and sentences (Richards, 2001: 55). For this reason, structuralists sought to classify the discrete elements of languages they studied. Another distinctive feature of the structural linguistics was oral: Speech was language. Brooks (1964) claimed that most people learned to speak before reading and writing. This became apparent to the people whose language did not have written forms were still able to communicate freely. The American linguist William Moulton (1961) also proclaimed that language teaching methodology ought to be based on speech instead of writing. Thus, the structural linguistics had profound influences on the fields of grammar as well as the field of anthropology. In other words, structural linguistics offered information concerning how speakers in different speech communities conveyed their ideas to listeners, especially the same ideas with different kinds of expressions. Furthermore, the structuralist approach was also based on the contrastive analysis hypothesis (hereafter CAH) in which the principal barrier to the target language acquisition/learning was the interference of the first language system (Brown, 2000: 208). Pedagogically implied, the contrastive analysis placed a strong emphasis on the surface structural differences between learners' native language (L1) and the target language they were learning (L2). Based on the structural differences between L1 and L2, language teachers were able to predict their pupils' learning difficulties. Then, teachers constructed a hierarchy of learning difficulties that pupils were supposed to confront with.

The Unitary Competence Hypothesis & The Integrative Approach

The structuralism guided the English test design, leading the linguistic elements to be assessed in distinct test sections (i.e. pronunciation test, vocabulary test). However, the flip side of the structuralism lied in its inability to integrate the language components for the

language use under a unified context. The *unitary competence hypothesis* (also *indivisibility hypothesis*) was rooted in opposition to the discrete-point philosophy which was limited to measure the surface structures of linguistic elements (Brière, 1971: 385). The unitary competence hypothesis was defined as the processes of comprehending and producing utterances which were governed by one-dimensional, unitary, and indivisible intellectual force. This global view of language proficiency was in harmony with *the integrative approach* or the *internalized expectancy grammar* which generated a sequence of linguistic elements by means of the *analysis-by-synthesis* process. In other words, language learners activated types of linguistic knowledge simultaneously to process or predict the incoming messages (Heaton, 1990: 16; Oller, 1979). Additionally, another distinctive feature of the integrative approach *redundancy* came into play to decode to textual meaning. The redundancy, being overlooked by the structuralists, offered enough contextual clues for interlocutors to grasp the essential meaning and obviate misunderstanding. In so doing, interlocutors would make valid guesses about the unknown elements when they had the capacity to perceive the redundancy in a particular message.

The Multi-dimensional Hypothesis & The Communicative Approach

The integrative approach focused primarily on the psychologically linguistic analysis, failing to cultivate the autonomous communication. Thus, based on the hypothesis of the multi-dimensional competence, the language proficiency was defined on different levels of competence such as the linguistic competence, socioeconomic competence, and so forth. The linguistic competence, therefore, was no longer regarded as the only one dimensional entity to measure learners' language proficiency. In addition, as some linguists claimed such as Fodor *et al.* (1974) and Rosenberg (1979), two levels of competence were broadly categorized: comprehending an input (receptive skills) and producing an output (productive skills). These two major types of competence were operated in two different processes. In the 1980, Canale and Swain postulated *the communicative competence* which consisted of the grammar competence, sociolinguistic competence, and strategic competence. Driven by the multi-dimensional competence hypothesis, the EFL language teaching in the 1980s highlighted the fundamental properties of language. Also, the language practice activities were aimed at simulating the daily authenticity. Here, the daily authenticity was not limited to the pragmatic features. The social and cultural aspects had to taken into consideration in the foreign language learning and instruction. Therefore, the successful language learning was suggested to leave the womb of the classroom so that learners were able to carry out authentic and meaningful tasks. In addition to the meaningful values of the language learning, learners were treated as cooperative partners rather than individuals to grapple with the language tasks. While dealing with the communication tasks, language learners lay their immediate priority to the meaningful interaction rather than the accuracy of pronunciation. Thus, non-linguistic clues such as the facial expressions were accepted as the alternative communication strategies. If possible, language learners had to be adept at different styles of language under specific situations so that they were able to freely communicate with native speakers at multiple layers of socio-economic status. Finally, in order to stimulate students' motivation in the oral interaction, language teachers played a role as a facilitator rather than a dominator. In so doing, learners were highly motivated to interact with the "facilitators" whenever confronting with the communication blackout.

The aforementioned statement sketched the distinctive features of the communicative language teaching. However, several interfering factors deterred language teachers from pursuing the communicative goals in the classroom. To begin with, EFL teachers had to be highly proficient in various styles of language so that they were able to manipulate certain communication strategies in different contexts. Second, it was extremely to precisely define the notion *authenticity*. For instance, the issue *Mass Rapid Transit* was indeed authentic for

students dwelling in the urban areas. However, such a conversational topic might be arcane to the ones in the rural areas. Furthermore, it was far fetched to communicate fluently without certain conversational formulae. Therefore, students had to equip with certain conversational rules for the proper communication.

Upon the influence of the multi-dimensional competence hypothesis and the communicative/language competence, there was an apparent improvement in the content of language tests between the 1970s and the 1980s. As compared to the traditional discrete test formats, the language tests in the 1980s covered a much broader range of language abilities such as the knowledge of cohesion, functions, and sociolinguistic appropriateness. Besides, the test content was assumed to be authentic, which had strong bearings on what people did with language in real life rather than dull mechanical drills (Heaton, 1990:5; Skehan, 1988). For instance, John Upshur (1971) advanced a *productive communication test* in which testees were required to pick out the differing points between two pictures. In this task, only the success in communication was what counted. John Upshur offered an innovative idea for the follow-up communication tests such as the simulated communication tests devised by Bartz and Schulz (1974). As Bartz *et al.* (1974) described, the simulated communication tests purported to offer the maximum contextualization for test takers to demonstrate their oral performances, and no particular language skills were sought to elicit in this kind of test. If necessary, the realia or visuals were considered for adoption to make the tests more authentic. What's more, four types of language skills (i.e. listening, speaking, reading, and writing) had to be in an active interplay in a communication test.

According to the multi-dimensional hypothesis and the communicative competence, the features of communicative tests were summarized from the following studies: Carroll (1961), Clark (1972), Murphy (1978; 1980), Swain (1984), Weir (1983; 1990: 10-14), and Widdowson (1983).

1. The communicative testing was sought to offer a realistic setting.
2. The contents in a language test had to be integrative.
3. Sampled tasks and texts had to be authentic and genuine in a communicative language test. In addition, several variables had to be accounted for in the selection of the texts and tasks: where, when, how, with whom, and why the language was to be used.
4. The communicative performance was affected by testees' prior knowledge/experience/abilities.
5. Unsimplified language (i.e. non doctored, genuine) texts were best regarded as ideal inputs for test takers.
6. Tasks related to the evaluation of oral interaction ability had to be conducted under the normal time constraints.
7. A communicative language test was ought to be as direct as possible. In other words, candidates engaged in the communication under realistic linguistic, situational, cultural, and affective constraints. Besides, in the communicative language testing, the focus was placed on the functional meaning rather than the linguistic forms.
8. A communicative language test primarily measured test takers' productive skills. Therefore, the multiple choice format might be particularly suspect in this respect.
9. As for ways of test scoring, the statistical analysis did not satisfactorily reflect test takers' communicative language ability. Instead, testers had to briefly address the test results to test takers so that test takers' performance processes would be elicited by this introspection technique.

On the grounds of these nine features of a communicative language test cited above, test

takers' overall language proficiency was evaluated in diversified language tests which encompassed different areas of communication (Morrow, 1979: 155). For example, mixed categories such as conversation/discussion (listening + speaking), correspondence/note-taking (reading + writing) were touted as the authentic communicative tests because they were able to evaluate test takers' particular language and communicative skills.

METHODOLOGY

At the outset, this section briefly described two major English assessment tools for senior high school graduates: (1) the English Tests of Joint College Entrance Examination (ETJCEE) from 1956 to 2001, and (2) Departmental Required English Tests (DRET) from 2002 to 2006. Then, a detailed account was provided for the framework of grammar ability by Pupura (2004:91), which served as the reference for the analysis of the grammar components assessed in the ETJCEEs/DRETs in the past fifty years.

English Test of Joint College Entrance Examination (1956-2001)

The ETJCEE, being promulgated in July in the year of 1956, had long been the single English proficiency test for the senior high school graduates who were desired to receive the university education. This nationally administered English test primarily measured the general English proficiency of senior high school graduates. Thus, the content of ETJCEE was not limited to particular senior high school English textbooks even though it was well proven to have strong bearings on the six volumes of senior high school English textbooks which were compiled by the National Institute for Compilation and Translation (國立編譯館 NICT) (Chang *et al.*, 1993; Tang *et al.*, 1993). In this study, all of the ETJCEEs in the past fifty years were collected from the *China Daily News* and the book *The English Tests of Joint College Entrance Examination from 1956 to 2001* by Learning Publishing Company (Liu, 2001).

In the early 2000s, under the impact of the newly-implemented system *Diversified University Admissions* (大學多元入學方案) by Ministry of Education (MOE), the Departmental Required English Test was groomed to take over the position when the ETJCEE was suspended in the year of 2002. The features of the DRET would be outlined in the subsection that followed.

The Departmental Required English Test (2002-2006)

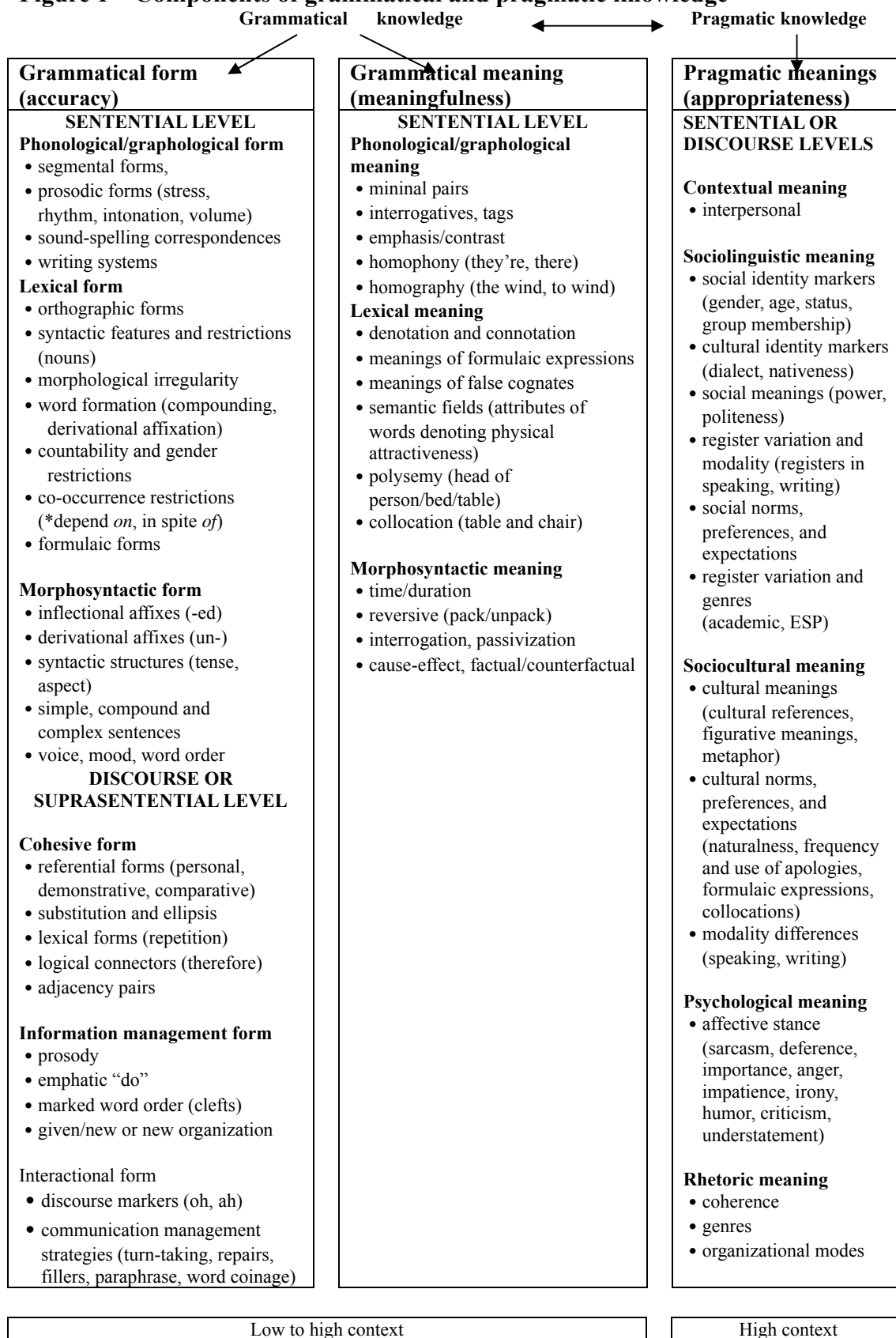
The multiple enrollment scheme *Admission via Examination* (考試分發入學) was reputedly akin to the traditional JCEE enrollment system, requiring senior high school graduates to participate in the *Departmental Required Test* (指定考試科目DRT) if they sought to advance into the university education. In addition, the DRET (指定考試科目英文科) was also defined as a general English proficiency test since its test content was not limited to particular senior high school English textbooks. In this study, the DRETs in the past two years (i.e. from 2002 to 2003) were collected from the book *The English Tests of Joint College Entrance Examination, Joint Evening Division College Entrance Examination, and the Scholastic Aptitude Tests from 1994 to 2003* (Chen, 2003). The DRETs from the year of 2004 to the year of 2006 were downloaded from the website of CEEC (大學入學考試中心) in the following: <http://www.ceec.edu.tw/AppointExam/AppointExamPaper.htm>.

Instruments Used for the ETJCEE/DRET Grammar Test Analysis

Before embarking on our analysis of assessed language proficiency, we had to be explicit in types of grammatical ability defined by Purpura (2004: 91). As Purpura (2004: 91) denoted, the grammar constructs encompassed two major fields: (1) grammatical knowledge, and (2) pragmatic knowledge. The former was operated at multiple layers (i.e. sentential level, discourse level, and suprasentential level). The latter was realized at the

sentential or discoursal levels. Figure 1 offered the types of grammatical ability as follows:

Figure 1 Components of grammatical and pragmatic knowledge



(Purpura, 2004: 91)

As Figure 1 sketched, the grammatical knowledge at the sentential level consisted of three major elements: (1) phonology/graphology, (2) lexis, and (3) morphosyntactics. The phonological/graphological knowledge pertained to the micro-skill speaking, ranging from the recognition of segmental forms, minimal pairs, suprasegmental elements, to synonyms. The knowledge of lexicon encompassed: (1) orthographic forms, (2) lexical meanings, (3) co-occurrence restrictions (i.e. idioms/phrases), and (4) collocation. The knowledge of morphosyntactics referred particularly to: (1) affixes, (2) syntactic structures (i.e. sentence patterns), and (3) tenses and aspects. These three gross categories lent themselves admirably to the taxonomy of assessed linguistic elements in our study. Additionally, at the suprasentential level, there remained three major grammatical abilities as follows: (1) cohesive form/meaning, (2) information management form/meaning, and (3) interactional form/meaning. These three types of grammatical knowledge were more aptly realized in the cloze tests or the dialogue tests, but they were possibly measured in the discrete-point tests (i.e. sentential coherence, logical connectors within a sentence). Thus, while analyzing the linguistic elements in the traditional discrete-point tests or in the passage-based tests, we had to take these six elements into account (i.e. from graphological form/meaning to the interactional form/meaning).

The other dimension of grammatical ability *the pragmatic knowledge* was relevant to the sociocultural norms and the psychological process. The former would be considered occasionally in the process of analyzing the dialogue tests. The latter, however, would be set aside because it laid beyond the scope in our study. Therefore, Purpura's framework of grammatical ability was given credit for us to discuss which grammatical constructs were invoked by an array of assessment tasks in the ETJCEEs/DRETs.

THE DEVELOPMENTS OF ETJCEE/DRET GRAMMAR TESTS

This section, on the grounds of the language proficiency hypotheses and their corresponding testing and instructional rationales, reviewed the periodical developments of ETJCEE/DRET grammar tests in the division of the following three major phases: (1) the impacts of the Grammar-Translation Method from 1956 to 1959, (2) the controlling powers of structuralism from the 1960s to the mid-1970s, and (3) the increasing popularity of integrative approach from the mid-1970s to the early 2000s. Other variables would be added for the description if the ETJCEE/DRET grammar test developments went awry of then-current testing or instructional approaches, such as the official policies regulated by Ministry of Education in Taiwan (MOE) and the Senior High School English Curriculum Standards (SHSECS).

To begin with, Table 1 provided the general ETJCEE testing trends in the late 1950s as follows:

Table 1. *The general testing trends in the ETJCEEs from 1956 to 1959*

Language abilities	Test sections	Years			
		1956	1957	1958	1959
vocabulary, idiomatic usage, syntactic patterns	Tests of explanation	20	30		
	Gap-filling correction	10	20	20	
	Gap-filling translation	20	10	20	
Idioms/phrases (preposition)	Choice of Words	20		20	
Vocabulary	Definition	6		10	20
	Spelling			10	20
	Parts of Speech				20
Sentence Patterns	Substitution	10			
Grammar rules	Syntax				20
Micro-speaking skills	Accentuation		10		
Reading	Comprehension	12	30	20	20

As Table 1 presented, the grammar test did not receive the central attention since a battery of measured linguistic elements was amalgamated into a test section, such as (1) Test of Explanation, (2) Test of Gap-filling Correction, and (3) Test of Gap-filling Translation. Few discrete test sections were spared to the grammar rules to be assessed. Virtually, through the glance at Table 1, the ETJCEEs from the year of 1956 to the year of 1959 attached heavy weights to the vocabulary knowledge, of which assessment was divorced from the context. Typical examples included: (1) Test of Definition, (2) Test of Spelling, and (3) Test measuring lexical declensions. The test *Choice of Words* ostensibly measured the idioms/phrases out of context. Remarkably, the ETJCEEs in the mid-1990s were rooted in the traditional Grammar-Translation Method, in which the lexical memorization out of context was crucial to the English study. Impacted by the Grammar-Translation paradigm, the rote memorization of vocabulary was commonly regarded as the building blocks to the successful English language ('The best ways to learn English', 1954: 2). Both English teachers and students zoomed in on the amounts of lexicons to be mastered, setting aside the language use (Chin Ling-chieh, 1958: 7).

Despite the fact that the ETJCEEs in the late 1950s were primarily grounded on the Grammar-Translation paradigm, the Audiolingualism/Structuralism (i.e. the divisible competence hypothesis) gradually wielded its power toward the end of the 1950s. This was best evidenced by the abolishment of the gap-filling translation task in the year of 1959 (See Table 1). For the audiolingualists, learners' native language had to be avoided so that the perfect foreign language behaviors would be achieved. In the meantime, a discrete grammar section *The Test of Syntax* won the nascent attention in the ETJCEE because the audiolingualists laid the structural accuracy as the foundation for the foreign language study. The example was presented as follows:

(1) Syntax 造句法

下面有二十題，每題包括四個答案，其中只有一個是正確的。請把正確答案的號碼寫在試卷內。

(The following are 20 test items. Each test item is followed by four options. Only one option is the correct answer. Please write the numbers of correct answers in your examination sheet).

1. He asked ① to the people for money ② for money to the people
 ③ the people for money ④ money to the people.
2. How much ① do you cost the books ? ② do the books cost ?
 ③ the books cost you ? ④ do cost the books ?

(ETJCEE 1960)

- (3) 下面左邊有二十個未完成的句子，請在右邊選擇適當的項目以完成之，並將代表適當項目的字母(A, B, C...等)填寫在試卷內。(20%)

(The following are 20 incomplete sentences in the left. You have to restore the chunks in the right column to make the whole sentence meaningfully complete. Please fill the letters A, B, C, in your examination sheet. 20%)

注意：①在選擇項目時，句意與文法並重。

②右邊有二十三個項目，其中有三個是沒有用的。

(Notice: ① The sentential meanings and grammar rules are of paramount importance.

② In the right are 23 candidates to combine with their corresponding incomplete sentences in the left. In particular, 3 chunks in the right function as the distractors.)

左邊未完成句子 (The incomplete sentences in the left)

1.The stranger told the little boy

2.None of the men in the room had

右邊的項目 (The chunks in the right)

(A)begin to work on her wedding clothes.

(B)of returning home after midnight....

Key: 1. (O)that he had a collection of old stamps.

2. (L)eaten anything for several days.

(ETJCEE 1961)

- (4) 下面每一句中間遺漏兩個字，這兩個字列在句子的後面。請找出這兩個字在句中原來的位置。答案的寫法是將所遺漏的字連同位置的前後兩字，共三字，依次填寫在試卷內。

例：The examination was long for us finish in an hour.(A) to (B)too

答案：(A) us to finish (B)was too long

注意：每一答案之三字中，如有任何一字拼寫錯誤，或次序顛倒，不予計分。

(Directions: Two words are missing in each of the following sentences. These two missing words are given after the sentences. Please find out the syntactic positions of these two words provided in each of the test items. The ways of your written Your answers should be presented with a missing word and its nearby two words such as the following example:

(Example) The examination was long for us finish in an hour.(A) to (B)too

Key: (A) us to finish (B)was too long)

(Note): You will receive no points if any of your written responses are misspelled or disordered.

1.The mayor was a nice person I enjoyed the visit. (A)such (B)that

2.The large drawer in the papers were kept locked up. (A)was (B)which

3.The fellow recommended the committee know my name. (A)does (B)by

Key: 1. (A) was such a (B) person that I; 2. (A) in was the (B) kept which locked; 3. (A) committee does know (B) recommended by the

(ETJCEE 1964)

(5) 請將下面每題中的五個項目組成一句通順的句子，並將各項目的標號按組成的順序填寫在試卷內。請特別注意大小寫及標點符號。

(Please reorder five elements in each of the following sentences into a well-structured sentence. Please mark the numbers in order in your examination sheet. Please be careful of the capitalization and the punctuation.)

1. do you know where is John, Mary going?

1 2 3 4 5

2. Tom Should be found guilty, I would have hanged. him

1 2 3 4 5

Key: 1. 31425 2. 21354

(ETJCEE 1965)

These four creative ETJCEE grammar tests, as compared with the multiple-choice syntax test in (1), coincided with a distinctive feature of structuralism which processed the grammar rules inductively rather than deductively. Traditionally, *Test of Syntax* in the year of 1959 explicitly guided test takers to retrieve specific grammar knowledge to achieve the intended test responses. These four innovative grammar tests laid the overemphasis on test takers' underlying linguistic knowledge to achieve the sentential paraphrase (See the example 2), the cohesive meanings (See the example 3), and the rearrangement of mutilated chunks (See the examples 4 and 5). The structural accuracy was still the foundation for the test achievement, but it was generated reversely in comparison with the syntax test. Besides, according to Richards (2001: 61-62), the tests of insertion and rearrangement were typical activities in the heyday of the audiolingualism. The former was also known as the activity of expansion in which language learners supplied a cued lexicon in a certain place. The latter was also named as the activity of restoration which required language learners to rearrange the discrete chunks to its original forms with a minimum of changes and additions.

It was inane to conclude the influences of structuralism were on the wane in the late 1960s, with the overall test scores decreasing sharply from 50 percent in the year of 1967 to 6 percent only in the following year (See Table 2). As a matter of fact, it was the translation task that seized 32 percent test scores in the year of 1968. The translation task was swept under the rug in the year of 1959 (See Table 1) on the consideration of its corrosive impacts upon test takers' renderings (i.e. the negative influences of test takers' native languages). Why was such a limited writing task brought into vogue in the year of 1966? In fact, it was not uncommon to see the revival of the translation tasks since the Senior High School Curriculum Standards (SHSECS) in the year of 1962 clearly regulated mastering the translation abilities was one of the instructional objectives. Besides, the senior high school English education was supposedly amenable to the ETJCEE testing trends. Without the translation tasks, senior high school English teachers would not commit themselves in the English writing instruction. Did the ETJCEE translation tasks realistically measure test takers' writing performances? Prior to the explanations, the following presented two translation tasks in the year of 1966:

- (6) 對下面 8,9,10,11,12,13,14,15 各題，在每題四個答案之中，選擇其中一個你認為對於這句中文是最恰當的英文譯句，把它的標號填寫在試卷內。

(Directions: Each of the following test items (i.e. from item 8 to item 15) is followed by four sentences. Among these four English sentences, please pick out the one which best fits the Chinese equivalent in the stem. Then write down the letter of a correct answer in your examination sheet.)

8.我伯父的兩個孩子都是男孩子。

- (A)My uncle's both children are boys.
(B)My uncle's children both boys.
(C)Both my uncle's children are boy.
(D)Both my uncle's children are boys.

9.在這屋裏有一些桌子。

- (A)This room have several desks.
(B)There are several desks in this room.
(C)There have several desks in this room.
(D)In this room have several desks.

10.昨天下很大的雨。

- (A)Yesterday rained very hard.
(B)Yesterday have very hard rain.
(C)It fell very large rain yesterday.
(D)It rained very hard yesterday.

Key: 8.(D) 9. (B) 10. (D)

(ETJCEE 1966)

- (7) 對下面問題，請照中文的意思，從每個括弧中選擇一個最恰當的字，並且加上必要的標點符號，都把它按照標號填寫在試卷內。

(Please select one word in each test item that best matches its Chinese counterpart. If needed, the punctuation should be added to your responses. Please fill these answers in the examination sheet.)

A.他向來是個懶惰的人，所以他現在沒有錢，也沒有食物。

- 1.(He/She/They/We) 2.(do/did/has/have) always been
3.(a such/such a) lazy man 4.(so/that) now he hasn't
5.(any/some)money or 6.(food/foods)
7.(too/either)

B.你有什麼傢俱可以借給我們嗎？

- 8.(Are/Have/Has/Is) there any
9.(furniture/furnitures) of 10.(your/yours) you can
11.(borrow/lend) 12.(I/us/you)

C.如果昨天醫生不在那裏，那些可憐的嬰孩全都死亡。

- 13.(did not be/had not been/was not/would not be) there yesterday,
14.(the all poor, little/all the poor, little/all the little; poor/the all little poor) babies
15.(was dead/would be died/would dead/would have died)

Key:

- A. 1. He 2. has 3. such a 4. that 5. any 6. food 7. either.
B. 8. Is 9. furniture 10. yours 11. lend 12. us?
C. 13. had not been 14. all the poor, little 15. would have died

(ETJCEE 1966)

As seen in the examples (6) and (7), such translation tasks were just backdrops for traditional grammar tests which bypassed test takers' real written performances. For instance, the Items 8 and 10 in (6) were akin to the previous test of rearrangement in (5), reordering the discrete language components. Another translation task in (7) primarily elicited grammar structures by means of the multiple-choice scoring method. These two translation tasks were advantageous to economically evaluate test takers' achievement due to their fixed responses, but they failed to reflect the real writing abilities. Such a defect was improved when the translation task in the year of 1968 realistically measured test takers' limited writing performances. Typical example was provided as follows:

(8) 說明：把下列四句譯成英文。

(Directions: Please translate the following four sentences into English)

1. 沒有知識，你不能當教員。
2. 人生和夢一樣，沒有多少歡樂。
3. 你應該好好利用你的寶貴光陰。

Key:

1. Without knowledge, you cannot be a teacher.
2. Life is just like a dream, without much pleasure in it.
3. You should make good use of your valuable time.

(ETJCEE 1968)

Obviously, the controlling writing task in (8) still rested in the mastery of linguistic elements rather than the real written communication. Nevertheless, this sentence-oriented translation task was expected to awaken senior high school students the importance of productive writing skills. However, this pioneering work led to several controversies. Firstly, the sentences posed for translation in (8) were meaningfully irrelevant which resulted in the communication blackout (i.e. not assessing test takers' ability to organize the events logically). Besides, without any universally accepted scoring rubrics, the ETJCEE examiners' test evaluation would become a free-for-all (Lu, 1969: 3). Professor Liu Hsi-ping (Lu, 1969: 3), for example, commented that the unsettling grading problem was prone to the great disparity on the test score results, tarnishing the impartiality of the ETJCEEs. Given the selective function of the ETJCEE (i.e. passing the examination and then getting admitted to the universities), the gap-filling translation test format was reinstated in the year of 1970, acting as a guided writing device to elicit test takers' grammatical competence.

To sum up, the structuralism wielded its overwhelming powers in the 1960s since various innovative grammar tasks sprang up in the ETJCEEs. This held true for the translation tasks which were believed as the alternative assessment task to measure test takers' grammar knowledge. Pedagogically implied, the structural accuracy was indeed the instructional objective in the local senior high school English classroom education. However, with the consideration of the limited instructional hours per week, the deductive grammar teaching was still regarded as the cost-effective strategy. Thus, in spite of the popularity of the structuralism in the ETJCEE test design, senior high school English teachers were reportedly mesmerized by the traditional Grammar-Translation Method (Hsu, 1970: 3;

Lin, 1967: 10; Yeh, 1965: 2; Yuan, 1974: 2). As a result, it was not uncommon to see senior high school English teachers dwelt much on the grammar rule analysis in class. Was the structuralism playing a leading role in the ETJCEEs as well as the local senior high school English teaching in the 1970s? The results were presented in Table 3:

Table 3. *The general testing trends in the ETJCEEs in the 1970s*

Assessed Language Abilities	Test Sections	Years									
		70	71	72	73	74	75	76	77	78	79
Vocabulary Test		12	17	0	20	10	10	0	10	20	5
Grammar	Syntax	20		20	10		10			5	
	Substitution	10	20		10						
	Affixes	13									
	Rearrangement							10			
	Deletion		10								
	Punctuation & Capitalization			7	6						
	Rearrangement (Coherence)			3							
	Error Correction					20					
	Sentential completion						20				
	Analysis of grammar rules								5	5	
Sentence Combination									10		
Vocabulary & Grammar	Translation (Gap-filling)	10									
	Translation (Chinese-English)		10	15	10	20	5	10	10	10	10
	Translation (English-Chinese)						5	10	10	10	10
	Discrete-point test			20		20		30	5		15
Speaking Test		20	10	20	20	10	30	20	12	12	11
Reading Comprehension Test		10	33	15	24	20	20	30	70		45

As Table 3 demonstrated, the structuralism still extended its impacts in the ETJCEEs in the 1970s, producing a number of innovative grammar tasks as follows: (1) *Test of Deletion* in 1971, (2) *Test of Punctuation & Capitalization* in 1972, (3) A variation on *Test of Rearrangement* in 1972, (4) *Test of Error Correction* in 1974, (5) *Test of Sentential Completion* in 1975, (6) *Test of Analyzing Grammar Rules* in 1977, and (7) *Test of Sentential Completion* in 1978. Particularly, the *Test of Error Correction* in 1974 and *Test of Sentential Completion* in 1975 marked major breakthroughs in the ETJCEEs. The former was driven by the error analysis, in which the target language linguistic systems were also impeding factors to the English study. The latter implied the forthcoming of the integrative testing approach because the goal of assessment focused primarily on the sentential coherence. The following presented these two grammar tasks:

(9) Writing Ability (寫作能力) 20%

下面十個句子，每句有五處劃底橫線部份，其中有一處在寫作造句上是錯誤的。請將該錯誤的項目選出。按照規定作答。本題為單一選擇。每題2分。

(Directions: In the following are ten sentences, along with five underlined parts in each test item. In particular, one part is grammatically incorrect. Please find out the incorrect part in each item, and then supply your answers in the examination sheet. Only one grammatical error is found in each test item.)

31. The money (A) being stolen and we had no chance (B) of getting it back, so I (C) explained to her that there was no use (D) crying (E) over spilt milk.
32. I'd (A) like to get all the (B) informations you have (C) on a man (D) named Evans (E) who used to work in your office.

Key: 31. (A) 32. (B)

(ETJCEE 1974)

The multiple-choice editing test was still of the structural attribute in its test design,

seeking the erroneous linguistic elements as the primary target. Despite the major concern was laid to the lexical and the morphosyntactic forms, the prompted sentences seemed not highly supportive to discern grammatical errors. For instance, the Item 32 in (9) assessed the countability of the lexicon *information* which did not necessarily called for the support of the sentential meaning. Thus, such an indirect writing task had an unusual concentration on the grammar elements, remaining ignorant about other ingredients necessary for the real English writing task (Brown, 2004: 198). Due to the quibble of construct validity, the task of editing in (9) was abolished shortly after its administration in the ETJCEE.

Another grammar task *Test of Sentential Completion*, distinct from the previous structure-centered grammar tasks, cast specific attention to the sentential coherence through lexical and grammatical cohesion devices. Here was the following example:

(10) (The Sentence Combining Test)

說明：下列 31 至 35 題，每題含有若干短句。今按各短句之綜合原意重寫成一長的句子。新句中有一個空白，試在四個答案中找出一個最適合填入空白的答案。每題 2 分。

(Directions: The test items from 31 to 35 consist of several short sentences, respectively. Please combine these short sentences into a long one. There leaves a blank in each combined sentence. Please find out the correct answer to make the combined sentences meaningfully complete. Each test item accounts for 2 points.)

31. The rain is over. You must not stay any longer.
You must not stay any longer _____ the rain is over.
(A) when (B) that (C) now that (D) as for
32. One of the women left. A little later, she came back, carrying a chair. She put it down before me.
One of the women left and then returned with a chair _____ she put down before me.
(A) that which. (B) which (C) with which (D) of that
33. I don't like the books you sent me. I dislike all of them. The one I most dislike is The Utter Failure.

I like none of the books you sent me, _____ The Utter Failure.
(A) among them all (B) best off all (C) worst of all (D) least of all

Key: 31.(C) 32.(B) 33.(D)

(ETJCEE 1978)

Evidently, the integrative testing approach took its root in the sentence combination task in (10), inasmuch as test takers availed themselves of the proper connectors in concert with specific background knowledge. For instance, due to the discursal functions of the cued sentences in the item 31 (i.e. the cause-effect function), the option (C) *now that* was the best candidate.

With the encouragement of the cohesion-oriented grammar task in (10), the integrative testing approach would have been thought to guide the ETJCEE grammar test design in the late 1970s. Contradictorily, two innovative grammar tasks ran counter to then-current ETJCEE testing trends in the mid-1970s: (1) the sentence-based English-Chinese translation test in the year of 1975, and (2) the test of syntactic rule analysis in the year of 1977. These two types of grammar tests seemed passé in the era of the traditional Grammar-Translation paradigm. Which factors triggered the developments of these two grammar tests in the 1970s? Here were some examples.

(11) 翻譯 (Translation)

A. 第 41~45 題中，每題有一句英文，底下有四句中文，請把與英文句子最接近的中文句子選出來。每題 1 分。(四選一)

(Directions: The following are unrelated English sentences and their possible Chinese equivalents. Please select one Chinese equivalent for their counterpart in the stem.)

例 (Example) : Stop talking so loudly, please!

- (A) 請大聲講話。 (B) 請停止講話。
(C) 請別大聲講話。 (D) 請別停止講話。

41. That incident made his hair stand on end.

- (A) 那件事使他栽了跟斗。 (B) 那件事使他怒髮衝冠。
(C) 那件事使他毛骨悚然。 (D) 那件事使他費盡腦汁。

42. Oh, boy! Can't you just mind your own business?

- (A) 噢，孩子！你能不管你自己的生意？
(B) 孩子！你不能自己做自己的事嗎？
(C) 咳！你不能只管你自己的事嗎？
(D) 啊呀，孩子！你管得了自己的事嗎？

Key: 41. (C) 42. (C)

(ETJCEE 1975)

(12) 說明：每題有一個句子，其中畫有底線的字在文法上作什麼用？四個答案中有一個或多個是正確的。找出正確的答案。

(Directions: Each test item is constructed in the form of a single sentence which contains a word in underline. Please circle out the correct grammar explanation(s) of each given word in underline.)

6. I would never think of doing such a thing.

- (A) 是形容詞。 (B) 是動名詞。 (C) 作“of”的受詞。 (D) 作“I”的主要動詞。

7. She asked me who I thought would be invited.

- (A) 是代名詞，指“she”。 (B) 是代名詞，指“me”。 (C) 是“thought”的主詞。 (D) 是“would be invited”的主詞。

Key: 6. (B)(C) 7. (D)

(ETJCEE 1977)

The *English-Chinese translation task* in (11) remained the traditional Grammar-Translation paradigm which required the language learners to translate sentences into and out of the target language. Such a translation task received a lethal blow in its test design because test takers' real ability of renderings was plastered over with the multiple-choice scoring method. In other words, the task in (11) was the reading task as such. In a similar vein, the test of grammar rule analysis in (12) was the product of the Grammar-Translation Method, inasmuch as the deductive grammar rule analysis was the primary target. It was not uncommon to implement these two ETJCEE grammar tasks because the deductive grammar rule instruction was still believed as the cost-effective way in high school English education (Yeh Kung-chao, 1965: 2). Professor Tung Tsung-shuan (1972) concurred with Yeh's statement, saying that senior high school students would go berserk if English teachers did not dwell much on the deductive analysis of grammar rules. Accordingly, it is possible to see the revival of the deductive grammar tasks if the testers were the followers of the Grammar-Translation paradigm. Nevertheless, the ETJCEE

grammar test design still kept abreast of then-current testing trends, sweeping the Grammar-Translation tasks under the rug in the following decades (i.e. 1980s). This was further warranted by the general testing trends in the following table.

Table 4. *The general testing trends in the ETJCEEs in the 1980s*

Assessed Language Abilities	Test Sections	Years									
		80	81	82	83	84	85	86	87	88	89
Vocabulary	A single lexical spelling	5		5	5						
Grammar	Sentence combination	10									
Vocabulary, grammar, idioms/phrases	Translation (Chinese-English)	10									
	Translation (Chinese-English) (Gap-filling)		10	10							
	Translation (English-Chinese)	10									
	Translation (English-Chinese) (Gap-filling)		10	10							
	Discrete-point tests	20	20		10		5				
Speaking Test		5	10	5	5	10	15	5	20	20	
Reading Comprehension Test		40	30	50	30	55	60	55	45	40	50
Writing	English Composition (Lexicons and idioms/phrases)		20								
	English Composition (Free composition)			20	20	20					
	Situational composition						20	20	20	20	20
	Direct English-Chinese translation (single sentences)				10	10	15				
	Direct Chinese-English translation (single sentences)				10	10					
	Chinese-English translation (Passage)							20	20	20	20

As seen in Table 4, the integrative testing approach was gradually palatable to the ETJCEE test administrators. This was further envisioned in the declining test score percentage in the traditional discrete-point tests. In the beginning of the 1980s, only the sentential completion task remained in the grammar test section. With the wielding powers of the integrative testing approach, the traditional discrete-point grammar tests were eventually supplanted by the highly integrative assessment tasks, such as: (1) the multiple-choice cloze task, (2) the direct writing task. Virtually, in addition to the integrative testing approach, the official test policies proclaimed by the Ministry of Education (MOE) acted as the driving forces to speed the innovation of the ETJCEE test design. In response to the waves of comments on senior high school students' less proficient writing proficiency (Chang, 1977; Li, 1976; Yuan, 1974), the Ministry of Education (MOE) was in more active terms about the direct English composition task, proclaiming to carry out the English writing tests in the year of 1981. Meanwhile, another highly integrative task *Multiple-choice Cloze Test* was resuscitated to indirectly measure test takers' manipulation of linguistic elements within a unified context. Due to the spree on the integrative testing approach, the traditional form-focused grammar tasks were rendered obsolete in the 1980s. The following presented the multiple-choice cloze test:

(13) 綜合測驗：(含讀音、拼字、用字、文法、文意等)

(Cloze Test: (including speech sound, spellings, word usage, grammar rules, and textual meaning))

說明：下面三段文章，有三十個空白，編號 1-30，每個空白含有四個可能的答案。請存細閱讀各段文章後，選出一個最適當的答案，並將其標示在答案卡上。

(Directions: The following are three short reading passages with 30 blanks which are numbered from 1 to 30. Each item is followed by four options. Please browse through each reading passage and then circle out the letter standing for the correct

answers.)

A good teacher is many things to many people. I suppose everyone has his definite ideas about (1) a good teacher is. As I (2) back on my own experience, I find the teachers that I respect and think about the most are those who demanded the most (3) from their students.

I think of one teacher in (4) that I had in high school. I think he was a good teacher because he was a very strict person. He just (5) no kind of nonsense (6) all in his classroom. I remember very vividly a sign over his classroom door....

- | | | | |
|----------------|---------------|---------------|----------------|
| 1.(A)how | (B)that | (C)who | (D)what |
| 2.(A)figure | (B)figured | (C)look | (D)was looking |
| 3.(A)dicipline | (B)disciprine | (C)dissepline | (D)discipline |
| 4.(A)peculiar | (B)particular | (C)curio | (D)curiosity |
| 5.(A)endure | (B)intoliet | (C)tolerated | (D)toils |
| 6.(A)at | (B)after | (C)in | (D)to |

Key: 1.(D) 2. (C) 3. (D) 4. (B) 5. (C) 6. (A) 7. (C) 8. (D) 9. (B) 10. (D)

(ETJCEE 1984)

The multiple-choice cloze test, as seen in the example (13), was credited to accommodate the multiple language components at once, such as the phonological forms, orthographic forms (e.g. spelling), morphosyntactic forms, interaction forms, and the like. Being highly praised for economically and efficiently measuring test takers' linguistic knowledge as well as the reading comprehension ability, the multiple-choice cloze task made its impact so complex in the ETJCEEs. One year later, a gap-filling cloze test was given credit to detach the ubiquitous link between the reading and writing ability. As Professor Cheng Heng-hsiung *et al.* (1991: 4) illustrated, the gap-filling cloze task was also defended as the alternative assessment tool to measure students' limited writing ability, tapping the proper lexicons into certain contexts (Cheng Heng-hsiung *et al.*, 1991: 4). If students were not expert in this limited writing task, they were supposedly less proficient in the writing performances. That was why the gap-filling cloze task was reputed as the warm-up writing activity. The example was presented in the following:

(14) 填充題：(A Gap-filling cloze test)

說明：下面這段文章共有十個空白，編號(a)至(j)。每個空白內均應填入一個字，才能使語法正確、文意清楚。請依照上下文的意思，按順序把各空白內應填的字，連同編號，寫在「非選擇題試卷」上。

(Directions: The following is a text with ten blanks which are numbered from (a) to (j). Please insert a word in each blank so that the whole text is grammatically and meaningfully correct.)

Margaret Bourke-White was the first photographer for *Fortune* magazine. Later she became (a) of the first photographers for *Life*. (b) a result of her photo essays in *Life*, she became famous. Her name was (c) known than any other photographer's. Surprisingly, Margaret chose her career almost (d) accident. She started college with the idea (e) becoming a scientist. But she found (f) necessary to support herself. She turned to taking pictures. From (g) moment she fell under the spell of photography, she was obsessed by (h) she and the camera could do together. "After I

found the camera,” she once said, “ (i) never really felt like a whole person unless I (j) planning pictures or taking them.”

Key: (a) one (b) As (c) better (d) by (e) of (f) it (g) the (h) what
(i) I (j) was

(ETJCEE 1985)

According to the example (14), the gap-filling cloze task did not receive considerable attention in the ETJCEEs. On the basis of the exact scoring method, only the closed class lexis, such as the connectors, pronouns, determiners, interrogative forms were enumerated in such a slot-filler reading and writing task. If ETJCEE examiners failed to rein in the expected responses from test takers, they have to exhaust themselves in checking the answers. Under the consideration of the economy of scoring, the gap-filling cloze task sparsely appeared in the ETJCEEs.

A direct writing task was reputed to assess test takers' creative ideas and their manipulation of various linguistic elements. Impacted by the integrative testing approach, the direction translation task was assessed at the discursal level, with a strong emphasis on context and meaning (Brown, 2004:220). Below was the typical example:

(15) 翻譯：中譯英 (Translation: a Chinese-English translation test)

下面一段短文共有五個中文句子，請譯成正確、通順而達意的英文。每題4分，而各題都分前後兩個半句，各佔2分。答案請寫在「非選擇題試卷」上。

(Directions: The following is a short passage with five Chinese sentences posed for translation. Please translate these five sentences meaningfully and accurately in English. Each sentence accounted for four points. Besides, in speaking of scoring, one sentence is divided into two parts, each part accounting for two points. Please write your responses in the given answer sheet.)

(1) 每天早晨有一位穿綠色制服的人(2%)，把信件送到我們的社區來(2%)。

(2) 每論是晴天或下雨(2%)，天天都可以看到他(2%)。

(3) 如果有掛號信要送給我們(2%)，他就會在門口按一下鈴(2%)。

(4) 他認識我們附近的每一個人(2%)，而我們每一個人也都認識他(2%)。

(5) 有時候我們在上學的途中遇到他(2%)，他就以微笑跟我打招呼(2%)。

Key:

(1) Every morning a man in a green uniform delivers letters to our community.

(2) Rain or shine, he can be seen every day.

(3) If there are any registered letters for us, he will ring the doorbell.

(4) He knows everyone in our vicinity and vice versa.

(5) Sometimes I meet him on my way to school, and he greets me with a smile.

(ETJCEE 1986)

In summary, contrary to the structuralism, the integrative testing approach accommodated the grammar knowledge in the cloze test or the performance-based tasks. The traditional discrete-point test was considered passé in the 1980s. Pedagogically implied, with the advent of two highly integrative tests (cloze test and direct writing test), there indeed arouse wide discussions on the strategic pathways in bolstering test takers' writing proficiency as well as the intensive reading skills (Hsiao, 1981:3; Huang, 1997: 93). With

the washback effects of the ETJCEE writing tasks, test takers begged for the short cuts to the efficient writing skills, such as memorizing model texts (Huang, 1997: 93), drilling sentence patterns for the renderings (Hsiao, 1981:3). Nevertheless, senior high school English education did not fly out of the womb of the structuralism, inasmuch as the cloze tasks were blamed for seeding with the linguistic elements as well as laying the overemphasis on the intensive reading skills.

Despite the local senior high school English education lagged a little bit behind the testing trends, the ETJCEE test developers still endeavored to assign equal weights to both linguistic and discorsal knowledge in the cloze tasks. However, as Table 5 revealed, the ETJCEE assessment tasks from the 1990s to the beginning of the 2000s remained the same as the past decade:

Table 5. *The general ETJCEE testing trends from the 1990s to the year of 2001*

Test Sections	Years											
	90	91	92	93	94	95	96	97	98	99	00	01
Vocabulary & Idioms/phrases			10	10	10	10	10	20	15	15	15	15
Dialogue	10	10	10	10	10	10	10	10	5	5	5	5
Reading (Cloze)	30	20	30	20	20	20	20	10	20	20	20	20
Reading (Text/Passage)	20	30	20	20	20	20	20	20	30	30	30	30
Writing (Translation)	20	20	10	20	20	20	20	20	10	10	10	10
Writing (Composition)	20	20	20	20	20	20	20	20	20	20	20	20

As Table 5 demonstrated, the ETJCEE was sluggish in its test reform processes. Heavy weights were still attached to the reading- and writing-oriented tasks. It was not until the year of 2002 that newly-developed grammar-oriented task appeared in the newly-implemented Departmental Required English Test (DRET): (1) the matching textual meaning test in the year of 2002, and (2) the gap-filling textual structure test in the year of 2002. The former was akin to the gap-filling cloze test, seeding with linguistic elements. The latter placed its focus on the coherence between textual segments.

(16) 文意選填 (A Gap-filling Textual Meaning Test)

說明：第 31.~40.題，每題一個空格，請依文意在文章後所提供的(A)到(J)選項中分別選出最適當者，並將其字母代號標示在答案卡之「選擇題答案區」。

(Directions: The following items (from No. 31 to No. 40) contain one blank, respectively. Please select the options in the following to make the whole text meaningfully complete.)

Are you someone who practically lives in front of the computer—a *mouse potato*? Or are you nervous about new technology—a *technophobe*? _____ 31. _____, if you want to master the English language, you will need to be _____ 32. _____ those new computer words that seem to be popping up everywhere.

Luckily, most computer words are easy to learn. _____ 33. _____, many of these words probably already have similar forms in your own language. _____ 34. _____, the German word for computer is *Computer*, in South American Spanish it is *computador*, and in Japanese we _____ 35. _____ *konpyuta*.

Another reason why computer words are easy to learn is that many of them are so _____ 36. _____. They are words that often make us smile when we first hear them such as _____ 37. _____ (traditional mail rather than Internet-based mail) or *wysiwyg* (what-you-see-is-what-you-get).

To get a feeling for computer words, it helps to _____ 38. _____ the world that created them—*cyberculture*, as it is often called. The computer industry is _____ 39. _____ young people who think of themselves as very different from _____ 40. _____

business people in suits. It is a word that avoids heavy scientific-sounding language in favor of words that are simple, fresh and playful.

- (A) For one thing (B) full of (C) understand (D) familiar with
 (E) find (F) In either case (G) For example (H) traditional
 (I) snail mail (J) colorful

Key: 31 (F) 32 (D) 33 (A) 34 (G) 35 (E) 36 (J) 37 (I) 38 (C) 39 (B) 40 (H)

(DRET 2002)

In the extract (16) cited above, the guessing rate was hardly achieved because each given written prompt was indexed with specific linguistic attributes. Despite the economic and efficient means to evaluate a considerable number of linguistic components in (16), the cloze assessment task took a serious beating for not completely realizing the cohesive functions of discrete clauses. Such a discursal viewpoint came into play in another innovative cloze task which was named as the matching test of textual structure. The following is a typical example:

(17) 篇章結構 (Textual Structure) :

說明：第 41. 至 45. 題，每題一個空格，請依文意在文章後所提供的(A)到(J)選項中分別選出最適當者，填入空格中，使篇章結構清晰有條理，並將其英文字母代號標示在答案卡之「選擇題答案區」。

(Directions: The following reading passage leaves 5 blanks which are numbered from 41 to 45. Please select the proper options in the following to make the whole text meaningfully complete.)

Windshield wipers of a car were invented by Mary Anderson on a trip in New York City in 1903. While touring the city on a streetcar, Mary was not interested in the views on the streets. 46 Repeatedly, the motorman had to get out of the streetcar to wipe off the snow and ice collected on the windshield. New York streetcar motormen at that time had tried various ways to solve this problem. 47 Mary, sitting on her seat, quickly drew her device in her sketchbook. 48 Mary's device allowed the motorman to use a lever inside the streetcar to activate a swinging arm on the windshield to wipe off the snow and ice. Because the device was first designed for cold weather, it could be easily removed when warmer weather arrived. 49 Even though her friends teased her about her awkward invention attached to a streetcar, Mary didn't give in to peer pressure. 50 By 1913, her invention became standard equipment on American cars. Windshield wipers save lives and make it easier to drive through storms.

- (A) Her solution was simple.
 (B) A year later, she received a patent for it.
 (C) Its function of wiping rain was later considered and added.
 (D) Instead, she paid much attention to the streetcar motorman.
 (E) Wiping off the snow and ice by hands, however, seemed to be the only solution.

Key: 46. (D) 47. (E) 48. (A) 49. (C) 50. (B)

(DRET 2002)

Unlike the traditional cloze formats, the extract in (17) laid its emphasis on the functions of discrete clauses, rather than pieces of discrete linguistic elements. As a matter of fact, each assessed clause carried cohesive relationships with its nearby segments.

To sum up, the ETJCEEs/DRETs were both of reading-focused nature. This ran counter to the communication-based EFL instruction depicted in the 1990s. Besides,

another feature of the DRET was the abolishment of the conversation test due to its inability to measure students' real communication ability. This might be detrimental to the classroom English speaking instruction because senior high school English teachers possibly shifted their focus back to the reading and writing skills. Such a deleterious washback effect was further verified by the speech of Professor Douglas Scholar in Taiwan. As Professor Scholar pointed out, only the first ability (i.e. the ability to tell the whole story) was the major concern in English teaching in Taiwan. Besides, reading was the primacy of the EFL classroom instruction. Little attention was devoted to the skills of listening, speaking, and writing. As a result, students who obtained high scores in the paper-based English exams failed to communicate with native speakers of English (Chang, 2005). In Scholar's words, both the ETJCEEs and the DRETs from 1999 to 2005 failed to keep up with the current language teaching and testing trends. Thus, senior high school graduates were unable to skillfully display their aural and oral abilities. So far, we dedicated our attention to the general developments of grammar test item types in the ETJCEEs/DRETs from the year of 1956 to the year of 2006. However, we had not delved into types of grammar knowledge which were realized in ETJCEE/DRET assessment tasks. In the following section, we spared our efforts to analyze types of grammar knowledge assessed in the ETJCEEs/DRETs.

TYPES OF ETJCEE/DRET ASSESSED GRAMMAR KNOWLEDGE

The preceding section outlined the historical developments of the ETJCEEs/DRETs in the past fifty years. This section addressed types of measured grammar knowledge in: (1) the traditional discrete-point grammar tests and cloze tasks, and (2) the structural complexity in the translation tasks.

Types of Grammar Knowledge in the ETJCEE Grammar Tests

To begin with, we resorted to Purpura's (2004: 91) framework of grammar ability to delve into types of grammar knowledge in the traditional ETJCEE grammar tests or the cloze tasks. The results were revealed as follows:

Table 6. *Types of assessed grammatical knowledge in the ETJCEEs/DRETs*

Grammar	Morphosyntactics	Coherence	Information management	Total
Number	480 (66.57%)	220 (30.51%)	21 (2.91%)	721 (100%)

Among types of grammar knowledge in Table 6, heavy weights were attached to the morphosyntactics which accounted for 66.57%. In addition, 30 % around pertained to the knowledge of coherence which linked separate sentences. It was not surprising that scant attention was devoted to the knowledge of information management since this category was limited to the structural parallelism, foregrounding, and emphatic meaning. Here, Table 6 clearly illustrated the predominant position of the morphosyntactics in the grammar tests. However, such a research finding failed to describe the historical developments of grammatical knowledge being assessed in the ETJCEEs/DRETs during these five decades. For this reason, the following table outlined the weights of assessed grammatical ability in a chronological order.

Table 7. *The weights of assessed grammatical knowledge in each decade*

Types of grammatical ability	Years						Total (A)
	1950s	1960s	1970s	1980s	1990-2001	2002-2006	
Morphosyntactics	69 (75.82%)	209 (78.87%)	84 (61.76%)	52 (57.78%)	48 (47.06%)	18 (48.65%)	480

Coherence	21 (23.08%)	45 (16.98%)	45 (33.09%)	37 (41.11%)	53 (51.96%)	19 (51.35%)	220
Information management	1 (1.12%)	11 (4.07%)	7 (5.26%)	1 (1.11%)	1 (0.97%)	0 (0%)	21
Total (B)	91	265	136	90	102	37	721

Note1: The percentage is available in terms of the following formula: $[N/Total (B)] \times 100\%$

As Table 7 demonstrated, more than 75 percent of the grammatical knowledge in the 1950s pertained to the morphosyntactics, and the remaining grammatical components were relatively distributed across the knowledge of coherence (23.08%) and the knowledge of informational management (1.12%). This finding was reasonable since the grammar tests in the late 1950s placed greater focus on the affixes (The Test of Parts of Speech: 1959), types of clauses (The Multiple choice of Syntax: 1959), and the like. In addition, this result was congruent with the traditional Grammar-Translation paradigm which laid its focus on the declensions. In the 1960s, the form-focused grammar testing trend remained popular, but the overemphasis was laid on the syntactic structures rather than the affixes. This could be further evidenced by various types of ETJCEE grammar tests in the 1960s. Examples include: (1) The Gap-filling Test of Insertion, (2) The Multiple-choice Test of Rearrangement, and (3) The Matching Test of Syntax. The surface structures were still the major concern in these three grammar rules, with a modicum of weight in the sentential coherence and the affixes. Impacted by the form-focused grammar tests in the ETJCEEs, senior high school students firmly believed that the deductive explanation of grammar rules provided the ground for the successful English learning (Tung, 1972). For instance, Professor Tung Tsung-shuan (1972) shared his experiences of high school English instruction in the early 1970s. As Tung (1972) stated, senior high school students expected their English teachers to dwell much on the grammar rules in class. They believed that mastering a finite set of grammar rules offered the efficient means to succeed in the ETJCEEs. Tung's report provided the convincing evidence of the deductive grammar instruction in senior high schools in the mid-1990s. The inductive grammar teaching met few responses in senior high school English teaching.

Nevertheless, as shown in Table 7, the knowledge of morphosyntactics gradually lost its preeminent position when the unitary competence hypothesis received the increasing attention in the ETJCEEs in the 1970s. As opposed to the testing trend in the past two decades, the rate of morphosyntactic forms gradually declined in the 1970s. The knowledge of coherence, by contrast, emerged as a crucial consideration in the assessment of grammar ability. This was further verified by its increasing percentage values from 16.98% to 33.09%. Meanwhile, there was a slight increase in the knowledge of information management from 4.07% to 5.26%. This result indicated that the knowledge of surface grammatical rules was no longer the major concern. Senior high school students had to be wary of the cohesive forms so that they could present their written accounts cohesively and logically.

The cohesion-centered grammar testing trend continued in the 1980s, especially in the heyday of the cloze test formats. As Table 7 demonstrated, approximately 42% of cohesive knowledge was capitalized to process the reading comprehension in the cloze tasks. Even though a host of morphosyntactic rules were still utilized in the grammar or cloze tests in the 1980s, their importance was continually declining (i.e. a gradual decrease from 61.76% in the 1970s to 57.78% in the 1980s).

In the 1990s, with the frequent use of the cloze tests in the ETJCEEs, the cohesive knowledge became the central focus. From then on, memorizing surface syntactic patterns was no longer the only short cut to obtaining high scores in the ETJCEEs. This was the similar case in the grammar testing trend in the newly promulgated DRETs in the early 2000s.

Both knowledge of morphosyntactics and cohesion were of paramount importance in the grammar assessment tasks. Virtually, the cloze tasks were treated as an alternative to the reading comprehension task, and the cohesive forms were essentials to conjugate the discrete sentences into a logical and meaningful reading passage. Without the cohesive forms, senior high school students failed to get access to the comprehension of a reading extract. That was the major reason why the knowledge of cohesion was outweighed in the ETJCEEs and DRETs from the 1980s.

The Syntactic Complexity in the Direct Translation Tests

Here, we shifted our attention to the structural features in the direct translation tasks. Virtually, this issue had been tackled by Chen Kun-tien (1992) who discovered that the sentences posed for translation from the year of 1983 to the year of 1991 were not structurally complicated. Chen's findings were tabulated in the following:

Table 8. *Types of sentence patterns in the translation tests from 1983 to 1991*

Sentence patterns	Numbers	Percentage
1. Simple sentence	20	42.6
2. Complex sentence	21	44.7
3. Compound sentence	4	8.5
4. Complex-compound sentence	2	4.2
Total	47	100%

(Chen Kun-tien, 1992: 12)

According to Table 8, the total amounts of simple sentences reached 42.6%, while the number of complex sentences totaled 44.7 % on average. Fortunately, little weight was allotted to the complex-compound sentence pattern (only 4.2%). This result denoted that the syntactic complexity was not the primary consideration in the direct translation tasks. What mattered rested in test takers' ability to convey their written messages in simple ways.

Here, we expanded Chen's research findings to offer the panorama of syntactic features in the direct translation tasks ranging from the year of 1968 to the year of 2006. The result was provided as follows:

Table 9. *Types of sentence patterns in translation tests from 1968 to 2006*

Years	Simple	Complex	Compound	Complex-compound	Total
68	2	2			4
69	2	1			3
92		2			2
93	4	1			5
94	1	3	1		5
95	1	1	2		4
96	2		3		5
97	5				5
98	1	2	2		5
99	2	1	1	1	5
00	1	3	1		5
01	2	3			5
04		1	1		2
04(M)		2			2
05		1	1		2
06		2			2
Total	23 (37.70%)	25 (40.98%)	12 (19.67%)	1 (1.64%)	61

Still, the results in Table 9 paralleled to Chen's previous research finding, in which both simple and complex sentences received the central focus in the ETJCEE/DRET direct translation tasks. The complex sentence in this sampled corpora was particularly referred to an independent clause in conjugation with a dependent clause through a connector (e.g. when, although) rather than a relative form (e.g. that, who, where, when). Therefore, the ETJCEE/DRET direct translation tasks in the past five decades were indeed not structurally and lexically difficult.

CONCLUSIONS

After reviewing the historical developments of the ETJCEE/DRET test formats and the analysis of assessed grammar elements, it is concluded that changes of ETJCEE grammar tests were in tandem with the global testing trends: (1) the overemphasis on the morphosyntactic forms from the late 1950s to the 1970s due to the impacts of the Grammar-Translation paradigm as well as the structuralism, (2) the increasing interest in the knowledge of coherence with the popularity of the integrative testing approach from the 1980s. Despite the DRET assigned equal weights to both morpho-syntactic forms and the coherence, senior high school English teachers still focused primarily on the former type of grammar knowledge. An educational inspector Wang Pu-yu was gloomy about the local EFL instruction precisely because quite a number of teacher dwelt much on the discrete language components, relied heavily on Mandarin as the medium of instruction, and set aside the aural and oral components (Chien Li-chun, 1995). Lin Show-jen (1990) investigated the ETJCEE washback effects upon the senior high school monthly English examinations (hereafter HSE). As Lin revealed (1990), the ETJCEEs were most likely to exert great influences upon the third grade HSE tests only. The test formats used in HSE tests on the lower grades were not greatly affected by the ETJCEE. The reason was adduced that most of the sampled teachers still placed more emphasis on the grammatical and lexical instruction to the lower grade senior high school students (Lin Shiow-jen, 1990: iii). Therefore, the senior high school EFL instruction remained the structural tradition despite that the ETJCEEs/DRET were moving toward the integrative testing approach.

Now that the grammar rules, especially the morphosyntactic forms, were regarded as the building blocks to the EFL grammar study, it was not uncommon to see the disparities between the structure-centered classroom EFL instruction and the current testing trends. Nevertheless, due to the predominant roles of the coherence knowledge, it will be weird for senior high school students to exhaust themselves in memorizing sentence patterns by rote. Instead, building up the concept of coherence/cohesion by reading short passages will be an alternative grammar instructional strategy for senior high school English teachers.

REFERENCES

- Bartz, W. H. & Schulz, R. A. (1974). *Approaches to the testing of communicative competence*. Paper presented at Central States Conference, April, in Milwaukee. (ERIC Document Reproduction Service No. ED 107108).
- Brière, E. (1971). Are we really measuring proficiency with our foreign language tests? *Foreign Language Annuals*, 4(3), 385-391.
- Brooks, N. (1964). *Language and language learning: theory and practice*. NY: Harcourt Brace.
- Brown, H. D. (2000). *Principles of language learning and teaching*. NY: Pearson Education Limited.
- _____ (2004). *Language assessment: principles and classroom practices*. NY: Pearson Education Limited.

- Canale, M. & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics* 1, 1-47.
- Carroll, J. B. (1961). Fundamental considerations in testing for English language proficiency of foreign students. In *Testing the English proficiency of foreign students* (pp.30-40). WA, DC: Center for Applied Linguistics.
- Clark, J. L. D. (1972). *Foreign language testing: theory and practice*. Philadelphia, PA: Center for Curriculum Development.
- Fodor, J. A., Bever, T., & Garret, M.(1974). *The psychology of language*. NY: McGraw Hill
- Heaton, J. B. (1990). *Writing English language tests*. NY: Longman.
- Huang, T. S. (1997). *A qualitative analysis of The JCEE English tests*. Taipei: The Crane Publishing Co., Ltd.
- Moulton, W. G. (1961). Linguistics and language teaching in the United States: 1940-1960. In C. Mohrman, A. Sommerfelt, & J. Whatmough (Eds.), *Trends in European and American linguistics, 1930-1960* (pp. 82-109). Utrecht: Spectrum.
- Lin, S. J. (1990). *A study of the influence of the English test of the Joint College Entrance Examination upon the senior high school monthly and final English exams*. Unpublished M.A. Thesis at National Taiwan Normal University.
- Morrow, K. E. (1979). Communicative language testing: revolution or evolution. In C. J. Brumfit and K. Johnson (Eds.) (pp.143-158).
- Murphy, R.J. L. (1978) Sex differences in objective testing performance, mimeo. Associated Examining Board Research Report RAC/56.
- _____ (1980). Sex differences in GCE examination entry statistics and success rates. *Educational Studies*, 6/2, 169-178.
- Oller, J. W. (1979) *Language tests at school*. UK: Longman.
- Oller, J. W. & Perkins K. (1980). *Research in language testing*. Rowley, Mass: Newbury House.
- Purpura, J. E. (2004). *Assessing grammar*. UK: Cambridge University Press.
- Richards, J. C. (2001). *Approaches and methods in language teaching*. UK: Cambridge University Press.
- Rosenberg, M. (1979). *Conceiving the self*. NY: Basic Books.
- Skehan, P. (1988). Language testing: Part I. *Language Teaching*, 21, 211-221.
- Swain, M. (1984). Large-scale communicative language testing. In S. J. Savignon & M. S. Berns (Eds.) *Initiatives in communicative language teaching: a book of readings*. Reading, MA: Addison-Wesley.
- Upshur, J. (1971). Productive communication testing: Progress report. In G. E. Perren & J. L. Trim (Eds.), *Applications of linguistics* (pp.435-441). Cambridge: Cambridge University Press.
- Weir, C. J. (1983). The associated examining board's test in English for academic purposes: An exercise in content validation. In A. Hughes & D. Porter (Eds.) *Current developments in language testing* (pp.147-153). London: Academic Press.
- _____ (1990). *Communicative language testing*. UK: Prentice Hall International Ltd.
- Widdowson, H. G. (1983). *Learning purpose and language use*. Oxford: Oxford University Press.
- 張騫 (Chang, C.)。(1977, 9月27日)。考試領導教學。台灣新生報, 7版。
- 張錦弘 (Chang, C. H.)。(2005, 8月3日)。很多高中生 a, an 不分。聯合報, C7版。
- 張淑瑛 (Chang, S. Y.)、丁雍嫻 (Ding Y. H.)。(1993)。大學聯招英文科和英文教材的相關性。英語教學, 17(4), 55-61。

- 張焉 (Chang, Y.)。 (1968, 6 月 10 日)。關於大專聯考的英文入學考試。中華日報, 5 版。
- 陳坤田 (Chen, K. T.)、黃燦遂、林素娥、林啟一、蘇順發、康淑娥 (1992)。大學入學考試英文科寫作能力命題及評分之探討研究報告。台北市: 大學入學考試中心。
- 陳鳴弦 (Chen, M. S.)。 (2003)。大學英文試題詳解。台北市: 建宏出版社。
- 鄭恆雄 (Cheng, H. H.)、廖樹輝 (Liao H. H.)、斯定國 (Si T. K.)。 (1991)。大學入學考試英文試題編制及預試計畫 (一) 研究報告。台北市: 大學入學考試中心。
- 簡麗春 (Chien, L. C.)。 (1995, 5 月 20 日)。多聽多說英語琅琅上口駐區督學指教學待改進會話應重於文法。中國時報, 2006 年 4 月 2 日, 取自 http://vip.tol.com.tw/CT_NS/CTSearchList.aspx
- 金靈節 (Chin, L. C.)。 (1958, 10 月 19 日)。你能運用多少英文字。聯合報, 7 版。
- 陳鳴弦、蘇全富 (Su, C. F.) (2004)。英文文法分類題庫。台北市: 建宏出版社。
- 陳永富 (Chen, Y. F.)。 (1969, 7 月 8 日)。談英語學習的方法。青年戰士報, 8 版。
- 陳永森 (Chen, Y. S.)。 (1970, 5 月 27 日)。漫談英語學習。青年戰士報, 8 版。
- 徐文閱 (Hsu, W. Y.)。 (1970, 8 月 2 日)。請停止英語教學或從國小一年級開始教起 (上)。台灣日報, 3 版。
- 薛文郎 (Hsueh, W. L.)。 (1974, 3 月 18 日)。英語四月通嗎? 中華日報, 5 版。
- 李宜涯 (Lee, Y. Y.)。 (1976, 8 月 13 日)。英語奠基差。青年戰士報, 5 版。
- 林語堂 (Lin, Y. T.)。 (1967, 7 月 17 日)。論台灣的英語教學。中央日報, 10 版。
- 劉毅 (Liu, Y.)。 (2001)。歷屆大學聯考英文試題全集: 大學聯考第一屆到最後一屆。台北市: 學習出版社。
- 湯廷池 (Tang, T. C.)、丁雍嫻、張淑瑛。 (1993)。最近五年來大學聯考英文科試題與國編本英文課本之相關性研究報告。載於黃自來、湯廷池、鄭恆雄、吳耀敦、施玉惠、黃燦遂、丁雍嫻、陳坤田 (合著), 英文試題與中學英文教材之建立及命題人才之培訓研究報告 (頁 3-19)。台北市: 大學入學考試中心。
- 董崇選 (Tung, T. S.)。 (1972, 3 月 16 日)。有感於當前: 台灣英語教學。聯合報, 9 版。
- 葉公超 (Yeh, K. C.)。 (1965, 6 月 27 日)。英語教學問題。中央日報, 2、3 版。
- 袁蘭芳 (Yuan, L. F.)。 (1974, 4 月 25 日)。當前中學英語教學, 豈非一大浪費? 自立晚報, 2 版。

Senior High School Students' Perceptions of L2 Classroom Climate

Chang-O, Wang (王嫦娥)
Ta-Hwa Institute of Technology
moonwan@ms53.hinet.net

The study aims to investigate the construct of L2 classroom climate (hereafter climate) and develop a questionnaire on senior high students' attitudes towards climate, which is defined by teacher support, peer support, and students' preparedness. For the purpose of constructing a 5-point Likert-type questionnaire on climate, 204 first and second-year senior high school students in central Taiwan responded to 24 statements reflecting their positive or negative attitudes towards climate. After the administration of the Climate questionnaire in April, 2007, we performed reliability tests and item analyses, and principal component analyses for item selections. The results indicated that Cronbach's alpha value of the finalized 15 items was .8003 indicating good reliability and those of the scales on teacher support, peer support, and students' preparedness were .8594, .6227 and .7500.

In this study, descriptive statistics, t-test, and Pearson Product-moment correlation were utilized for data analyses. The main results showed that students' perceptions of climate varied significantly in terms of grade and sex at a significance level of .05, and that teacher support, peer support, and students' preparedness were positively and significantly correlated with climate ($r = .8020, .4896$ and $.5296, P < .001$). It is concluded that social support from significant others (e.g., teachers and peers) as well as students' preparedness may influence climate. Future studies utilizing the Climate scale in this study can take different cultural and socio-economic settings as well as diverse populations into consideration. It is suggested that data collection triangulation (e.g., observation and interview) can also be used to remedy the weakness of the questionnaire.

INTRODUCTION

Researchers suggested that senior high teachers might find a gap between their teaching beliefs and practice in implementing communicative language teaching (CLT) approaches in the EFL context of Taiwan (Liao, 2002). In a CLT classroom, it is essential for senior high school teachers to enhance students' autonomous learning (Benson, 2001; Reeve, 1998), give students academic and social support and create a cooperative and supportive classroom climate (Chang & Chang, 2003; Chang & Lin, 2001). However, traditionally, an English teacher in Taiwan is usually regarded as a knowledge provider and the only one authority in the classrooms. Frequently, he or she might face challenges in training students as active and autonomous learners. Besides, there is insufficient information about senior high school students' perceptions of CLT-oriented classroom climate in Taiwan. Thus, the study aims to examine the role of L2 climate and explore significant factors influencing the L2 classroom climate from senior high school students' perspectives.

Recently, a few researchers (Hsu, 2008) had examined students' perceptions of climate and found climate was associated with teachers' instructional behaviors, and motivational factors such as self-determination. Others contended that instructors have a basic understanding of students' perceptions of L2 classroom climate, which may not only help L2 teachers increase teaching effectiveness (Liu, 2006). It may also enhance the quality of peers' social/ academic support and interpersonal developments among students (Fassinger, 1995;

Hsu, 2001). However, little research attention is paid to researching the construct of L2 classroom climate and clarifying the components and factors of climate in EFL classrooms. This motivated the author of the study to re-examine the above-mentioned issues.

The research questions in this study were addressed as follows:

1. What is the level of L2 classroom climate reported by secondary school students, in terms of teacher support, student support and student preparedness?
2. Are there any relationships between L2 classroom climate and such factors as teacher support, student support and student preparedness?
3. Are there any gender and cross-grade differences among senior high school students in terms of their perceptions of L2 classroom climate?

Based on the aforementioned research questions, the author of the study made three research hypotheses. The first one was that senior high school students had a moderate level of L2 classroom climate. The second was that there was a positive correlation between climate and each of the three components of climate (teacher support, student support and student preparedness). It is assumed that climate can be predicted by the previously mentioned variables. The third hypothesis was that there was no significant difference between first and second year students and between males and females in terms of their perceptions of L2 classroom climate.

LITERATURE REVIEW

Defining, measuring and researching classroom climate

Definition of classroom climate The construct of classroom climate (i.e., “Ban-fong” or “ke-shi-chi-fen” in Mandarin, see Tsai, 2008: 55) is sometimes referred to the concepts such as learning environment, psychological/ psychosocial classroom environment, classroom atmosphere, class climate, and learning climate (Adelman and Taylor, 2005; Chang & Chang, 2003; Chang & Lin, 2001; Hsu, 2008; Liu, 2006; Tsai, 1995; Wu, 1978; Zheng, 1998; Zhong, 1983; 蘇佳純, 2008). Despite the different definitions of classroom climate, it is not denied that the impact of classroom climate on students and staff can be beneficial for or a barrier to learning (Wang, 1999).

Classroom climate can be seen as a perceived quality of the learning setting or class atmospheres if we consider environmental factors such as physical, material, organizational, and social variables (Chiou, 2001; Liu, 2006; Wu, 1978). Both school and classroom climate may reflect the cultures and ideologies of schools and classrooms. And, of course, classroom climate and culture are shaped by school's political and socio-economic contexts.

Classroom climate can be investigated from alternative perspectives. Chang and Lin (1990) identified two types of classroom climate: global and local. The first type of climate (global climate) refers to the psychological and emotional atmosphere in student-teacher and student-student interactions. The second (local) climate is defined as a type of emotional atmosphere, which is closely associated with teachers' leadership or instructional approaches (Liu, 2006; Sheng, 2007).

Moos (1979) identified three dimensions of the factors influencing climate from the perspective of human environment. These dimensions were (1) relationship dimension (e.g., the extent to which people get involved in the environment, and give support), (2) personal growth or goal orientation dimension (e.g., directions along which personal growth and self enhancement might occur), (3) system maintenance and change dimension (e.g., the extent to which the environment is orderly and expected and responsive to changes).

Researchers (Hsu, 2008; Kuo, 2001; 蘇佳純, 2008) pointed out that students' high motivation may result in good classroom climates. However, there is limited information about whether classroom climate might influence EFL students' learning motivation in L2 classrooms. Research should clarify the relationship between climate and motivation in a

more comprehensive way.

Based on the discussion above, we find that climate researchers do not have any agreed definitions of classroom climates. It is thus warranted to explore the role of classroom climates and investigate crucial factors affecting L2 classroom climate in Taiwan.

Approaches to studying classroom climate Fassinger (1995) and Hsu (2001) argued that the research on classroom climate should consider the factors of teacher support and peer support. However, the former author contented that student preparedness be related to L1 climate while the latter did not examine the role of student preparedness in influencing climate. These two authors did not observe high school students' and teachers' perceptions of the L2 classroom climate and predictors of climate.

Among the climate researchers in Taiwan, Chiou (2001) is the first author that identified three types of variables influencing classroom climate: teachers, students and situational factors. Furthermore, she pointed out three major approaches to the study on classroom climate. The first approach to the study of climate is concerned with the impact of teachers' leadership (e.g., democratic, authoritarian and laissez-faire, see Sheng, 2007 and Wu, 1978) on classroom climate. It is argued that the more democracy teachers showed in the teacher-student interactions, the more positive climate students had (Wang, 1999). The second approach to exploring climate is proposed based on the theories of group dynamics. From this perspective, classroom climate is seen as a psychological atmosphere emerging mainly from teacher-student interactions. The factor of teacher influence (Chen, 1997; Flanders, 1960) was seen as a predictor of classroom climate. The third approach to researching climate is related to the rationale that other elements of classroom environments (e.g., teacher control and students' involvement, and interpersonal relations) may influence the changes of classroom climate.

In addition to the approaches to the study of climate as previously discussed, Chiou (2001) emphasized the importance of researching climate from the fourth perspective. She insisted that future studies take sociometry into consideration and go beyond the current conclusions about the role of climates. Specifically, an exploration of climate from the perspective of sociometry may provide us with information such as about the levels of students' acceptance and refusal in a democratic learning environment.

It is noted that although there seems no systematic review of the approaches to L2 classroom climate, climate has been researched from the fifth (L2 motivation) perspectives. For example, Wang (2008) found that climate was a predictor of situational willingness to communicate and motivation based on L2 motivational theories.

Measuring classroom climate Based on diverse conceptualizations of climate, researchers developed a number of the classroom environment/ climate instruments. For example, Wu (1999) constructed a scale measuring the authentic and ideal climate in that it is assumed that types of climate can be directly or indirectly measured. Wang (1999) examined senior high school students' perceptions of school and classroom climate by means of the use of climate scales. Nevertheless, only a few of the researchers (Fraser, 1998; Lin, 1982; Wong, 2006) critically analyzed the reliability and validation of the developed climate scales.

Moreover, the past L2 researchers in Taiwan investigated the nature of L2 climate at primary and secondary schools (Chang & Lin, 2001; Chang & Chang, 2003; Liu, 2006; Wang, 2008) and in universities (Hsu, 2008). However, none of the previous authors investigated the issues of cross-grade differences in the climate studies.

According to the discussion above, we found that there are inconsistent definitions of climate and there is a lack of development and validation of the L2 climate scales. In this study, the author relied on recent research findings and theories on L2 classroom climate to develop a climate scale (Chang & Chang, 2003; Chang & Lin, 2001; Fassinger, 1995; Hsu, 2001). She identified three essential components of L2 classroom climate. They are teacher

support, student support and student preparedness. It is assumed that L2 classroom climate is a multi-dimensional construct.

Importance of L2 classroom climate

L2 Classroom climate (hereafter climate) is seen as a crucial factor influencing instructional practices and behaviors. First, as Fraser (1998) pointed out, an initial understanding of students' perceptions of climate may help instructors explore or evaluate the effect of teaching practice on the educational innovation (Chen et al., 2003, pp. 80; Fassinger, 1995; Liu, 2006). Besides, Hsu (2001) emphasized that a study on L2 learners' perceptions of climate contributed to a basic understanding of the differences in students' questioning practices. Recently, Hsu (2008) had found a positive correlation between instructional behaviors, motivation, and learning climate. Based on the discussion above, it is concluded that climate may be an indicator of instructional behaviors, motivation, and teaching effectiveness in L2 classrooms. In this study, the author only paid attention to the role of L2 climate and its relations to teacher support, peer support and students' preparedness.

METHODOLOGY

Participants

In this study, 204 first and second-year senior high school students in central Taiwan (e.g., in Nantou County and in Taichung County) participated in this study. Among them, 101 students were males and 103 students were females. It is noted that they were chosen from three senior high schools and they were taught by EFL instructors with at least 10-year teaching experiences.

Instruments

The study attempted to explore the construct of L2 classroom climate and developed a questionnaire measuring senior high students' attitudes towards climate. This author-tailored Foreign Language Classroom Climate Scale (FLCC, hereafter the climate scale) was constructed based on the recent theories and research on the construct of classroom climate (Chang & Lin, 2001; Chang & Chang, 2003; Fassinger, 1995; Hsu, 2001).

The climate scale utilized in this study included 24 items. The first eighteen items were adopted from Hsu's (2001) questionnaire on climate. Items 2, 5, 8, 11, 12, 13, 15, and 17 measured the variable of teacher support while the rest items of 1, 3, 4, 6, 7, 9, 10, 14, 16, and 18 examined the variables of peer support. The 19th to 24th items measuring students' traits of "preparedness", which were not originally included in Hsu's (2001) climate scale, were adopted from Fassinger's (1995) study.

The climate scale is a five-point Likert-type scale (from one to five). Senior high school students were requested to circle the number representing the levels of their agreement on the items (i.e., 1= strongly disagree; 2= disagree; 3= no opinions; 4 = agree; 5= strongly agree).

Before piloting the climate scale, the researcher translated the scale items into Mandarin Chinese and required two TESOL experts and two senior high school teachers to modify ambiguous scale items and enhance their content validity.

After collecting two hundred and four senior high students' responses to the items on the climate scale in the late April of the year 2007, item analyses and PCAs were performed to ensure the construct validity of the instrument (Wu, 2003). Before the item analyses, the scores of items 3, 4, 9, 14, 17, 19, and 20, which had negatively worded statements, were reversed in advance. Besides, the following were the criteria for item selections.

a. **Item-total (item-scale) correlation:** In the analysis of item-total correlation (ITC), items with statistical significance (e.g., $p < .05$) were retained for further analyses. Items that did not meet the above criterion were discarded.

b. **Comparison between extreme groups:** T-value (Critical ratio or CR) was calculated in the comparisons between extreme groups. To test the discriminant validity of scale items,

the mean score of each item between the two extreme groups of students (the top and the lowest 27 % of the sample) was compared. Statistical non-significance of CR (i.e., above .16, $p > .05$) is a criterion for removing the item.

c. **Reliability test:** The calculation of coefficient of internal consistency (Cronbach's alpha) is used to test the internal homogeneity of the scales.

Alpha was computed for every possible version of the scales with one single item removed at one time. Items were dropped when their deletion would substantially increase the alpha of the scale.

d. **Principal component analysis:** After the author decided to remove certain items based on the results of item analyses, principal component analysis was applied to analyze the factor structures of the instruments in the pilot tests.

Two profound criteria were utilized in the evaluation of the adequacy for analyses of factor structures. They were Kaiser-Meyer-Olkin (KMO, varying between 0 and 1) measure of sampling adequacy, and Bartlett's test of Sphericity (with a null hypothesis that all of the diagonal elements are 1 and all off diagonal elements are 0 in a matrix). The minimum criteria were the KMO over .6, and Bartlett's test of Sphericity with statistical significance ($p < .05$).

Oblique rotations (e.g., direct oblimin rotations, see McCroskey and Richmond, 1982) were used as a means of seeking a simple structure of the factor solution and extracting desirable and interpretable factors embedded in each questionnaire. After determining the number of factors, the author examined the pattern matrix and retained the item that had a major loading above .40. Items that loaded on two factors (loadings $> .40$) with a difference smaller than 0.2, were removed. Each factor included at least three items.

Two criteria were used to determine the number of factors for each scale. A scree analysis or Cattell scree test was first used as a criterion for determining the number of factors. Second, the number of factors that best reflected the structure of a particular construct was chosen. For example, three factors were extracted for climate because it supported Hsu's (2001) and Fassinger's (1995) theories on climate, which can be defined by three indicators--- teacher support, student support and student preparedness.

Data collection and analysis

In order to construct and validate a 5-point Likert-type questionnaire on climate, the author collected the data after 204 first and second-year senior high school students responded to 24 statements on the climate scale. After the administration of the climate scale in April, 2007, reliability tests and item analyses, and principal component analyses (PCAs) for item selections were performed (Wu, 2003). For data analyses, descriptive statistics, t-test, and Pearson Product-moment correlation were utilized in the study to explore students' perceptions of climate, gender and grade-level effects on climate, and the correlations between climate and teacher support, peer support and student preparedness.

RESULTS AND DISCUSSION

Summary of item analyses for Climate Scale

After the administration of the Climate scale in April, 2007, we started to perform reliability tests and item analyses, and principal component analyses. The main findings were briefly introduced in the following lines. First, based on the criteria of item selection listed above, items 3, 4, 9, 17, the author deleted 19, and 20 at the stage of the item analyses (See Table 1). As shown in Table 1, since item 4 had a negative (-.0021) and non-significant ITC ($p > .05$) and those of items 9 and 20 were not significant ($p > .05$), they were deleted. Besides, all of the 24 items were able to discriminate extreme groups except items 4 and 20 (i.e., CR values were not statistically significant, $p > .05$). It was also found that if items 4, 9, 17, 19, and 20

were deleted, the internal consistency of the scale would be improved. Cronbach's alpha values of the scale became .7541, .7454, .7584, .7432, and .7536, which were larger than that of the initial scale (= .7426). Based on discussions above, items 4, 9, 17, 19, and 20 were deleted.

Table 1. *Summary of Item analyses for Climate Scale*

Item	Item-total correlation	CR values	alpha if item deleted	Item deletion
1	.3714**	5.30**	.7286	
2	.6144**	7.82**	.7094	
3	.2124*	3.39**	.7393	
4	(-.0021)	(-.15)	(.7541)	√
5	.5212**	6.64**	.7169	
6	.2962*	3.93**	.7343	
7	.4716**	5.98*	.7214	
8	.5605**	.7.75**	.7138	
9	(.0462)	.91**	(.7454)	√
10	(.2959**	.361**	.7334	
11	.6206**	3.61**	.7103	
12	.6310**	8.74**	.7086	
13	.6469**	.607**	.7051	
14	.3734**	4.63**	.7283	
15	.5476**	6.54**	.7161	
16	.5306**	7.40**	.7174	
17	.2697**	5.64**	(.7584)	√
18	.4164**	5.16**	.7254	
19	.1816*	.287**	(.74324)	√
20	(.0282)	(.64)	(.7536)	√
21	.3514**	4.98**	.7299	
22	.3442**	4.20**	.7298	
23	.4456*	5.51**	.7231	
24	.4609**	4.77**	.7220	

Note: 1. Items with the symbol of “√” represented those deleted in item analyses. 2. ** p<.001* p<.05.

After the deletion of the previous five items, two PCAs were performed. The results of the first PCA suggested the deletion of items 6, 14, 16, and 18 (See Appendix 1). Since Items 6 and 14 loaded on the fourth factor and had less than three items, they violated the criteria for item selection listed above, and were deleted. Besides, it was found that all of the items loaded on the accurate components of teacher support, student traits and peer support except item 16 and 18 loading on the wrong factor. Thus, items 6, 14, 16, and 18 were deleted before the second PCA. The result of the second PCA (see Appendix 2) conformed to theoretical expectations of factor loadings. To be specific, items 2, 5, 8, 11, 12, 13, and 15 were loadings on the first factor of teacher support. Items 21 to 24 loaded on the second factor of student traits of preparedness, and items 1, 3, 7, and 10 on the third factor of peer support. Moreover, the Cronbach's alpha value of the finalized 15-item climate scale was .8003, and those of the three subscales were .6227, .7504, and .8594 (see Appendix 3).

Results on students' perception of climate

The main findings of descriptive statistic analyses (see Table 2) confirmed the hypothesis that senior high school students had a moderate level of climate (M=3.25; SD=9.29). The findings suggested students had a positive attitude toward their classroom climate.

Besides, the results of data analyses indicated that students had a moderate level of student preparedness, peer support, and teacher support (see Table 2). The results showed that students had a positive attitude to the above mentioned variables. Moreover, the students in

this study had higher scores on the teacher support scale than those on peer support or student preparedness scales. It seems that students obviously perceived teacher support rather than peer support and student preparedness in L2 classes.

As shown in items 23 and 24 of Table 2, the students had a positive attitude toward their comprehension of course contents and teachers' questions. However, we found the score on the item 22 eliciting their preparedness for course contents was the lowest ($M=2.51$; see in Table 2) in comparison with the scores on other preparedness scales. One possible reason was that the students in this study probably had to learn other subjects and it was difficult for them to find sufficient time to learn English when they were free.

Another two significant findings were noted. First, students had a moderate level of peer support, which was probably because the classmates had mutual respects and were familiar with and encourage each other (see items 1, 3, and 10 in Table 2). The second finding was that as shown in item 7 in Table 2, students scored low on the item of "My classmates like to complete L2 tasks in groups." The results indicated that students liked to work on L2 tasks by themselves, and they had a negative attitude toward peer assistance in L2 activities..

Table 2. A Summary of Means and Standard Deviation for Climate

Variable (label/ item)	Mean	Std Dev	N
Student preparedness			
G22 I am well-prepared for course content.	2.51	.90	204
G21 I am well-prepared for assignments.	3.10	.99	204
G23 I have good comprehension of course contents.	3.26	.97	204
G24 I understand English teachers' questions.	3.29	.92	204
Peer support			
G7 My classmates like to complete L2 tasks in groups.	3.17	.89	204
G1 My classmates encourage one another.	3.44	.01	204
G3 My classmates do not have mutual respects.	3.64	.97	204
G10 My classmates are closely acquainted with one another.	3.91	.94	204
Teacher support			
G8 Our speech is seldom interrupted by English teacher.	3.21	1.04	204
G13 Our English teacher is humorous.	3.46	1.13	204
G5 Our English teacher is an inspiring person.	3.62	1.06	204
G15 Our English teachers help us express opinions and ask questions.	3.63	.90	204
G12 Our English teacher answers to our questions clearly.	3.64	.99	204
G11 Our English teachers respect me and allows us to look at note books.	3.65	.93	204
G2 Our English teacher encourages us to speak English.	3.70	1.02	204

Results on the relationship between gender and climate

The second significant finding of the study was that students' perceptions of climate varied significantly with gender at a significance level of .05. First, Table 3 showed that females scored higher on the climate scale. It is also found that the gender difference reached a significant difference level. Regardless of students' educational level (junior and senior high school), the findings partially lend support to Chen et al.'s (2003) finding, which indicated that students' perceptions of climate varied with the gender factor. It seems that females might be more concerned about climate than males and had expectations of climate.

Table 3. *Results of descriptive statistical and T-tests on Gender and Climate*

(Part 1) Results of descriptive statistical analyses on gender and climate					
Variable	Number of Cases	Mean	SD	SE of Mean	
Male	101	3.12	8.704	.866	
Female	103	3.37	9.017	.888	

Mean Difference = -5.7122; Levene's Test for Equality of Variances: F= .029; P= .866

(Part 2) Results of T-tests on gender and climate					
Variances	t-value	df	2-Tail Sig	SE of Diff	CI for Diff
Equal	-4.60	202	.000	1.241	(-8.160, -3.264)
Unequal	-4.60	201.95	.000	1.241	(-8.159, -3.265)

Results on the relationship between grade and climate

The third significant research finding was that the first and second-year students had different perceptions of climate at a significance level of .05 (see Table 4). Specifically, the overall mean scores on the climate scale for the first-year and the second-year students were separately 3.16 (SD= 9.804) and 3.33 (SD= 8.519). Despite the factors such as students' grade (e.g., at first and second grade), the students in this study showed a positive attitude to classroom climate. However, when the cross-grade difference was taken into consideration, a significant difference was found among first and second-year students in terms of their perceptions of climate. One possible reason for the cross-grade differences was that second-year students might get more familiar with significant others (e.g., teachers and peers) or they had more learning experiences than the first year students. Another explanation was that the first year students might be less concerned with the L2 climate in comparison with the second-year students.

Table 4. *Descriptive statistical and T-tests on Grade and Climate*

(Part 1) Means and Standard Deviation for Climates					
Variable	N of Cases	Mean	SD	SE of Mean	
First grade	84	3.16	9.804	1.070	
Second grade	120	3.33	8.519	.778	

Mean Difference = -4.3167; Levene's Test for Equality of Variances: F= 1.728; P= .190

(Part 2) Results of T-tests on grade and climate					
Variances	t-value	df	2-Tail Sig	SE of Diff	CI for Diff
Equal	-3.35	202	.001	1.290	(-6.861, -1.772)
Unequal	-3.26	162.30	.001	1.322	(-6.929, -1.705)

Results on the relationship between the three components of climate

The fourth main finding of the study showed that teacher support, peer support, and students' preparedness were positively and significantly correlated with climate ($r = .8020, .4896$ and $.5296$, $P < .001$, see Table 5). The results suggested that climate could be predicted by the previously mentioned factors. The main reason is that there was a systematic

relationship (e.g., positive correlations) between climate and support from teachers and peers and between climate and student preparedness. Besides, an increase of social support from teachers and peers and the higher levels of preparedness probably resulted in better climate.

The findings of the study also showed that among the components of climate, teacher support played a more significant role than peer support and preparedness levels in influencing climate. One explanation was that the students got a positive experience in winning teacher support. It is also noted that peer support was the least correlated with climate in comparison with teacher support and student preparedness. One possible explanation was that the students got insufficient social or academic support from their peers or they learned English in a competitive learning environment.

One minor finding was that there was no significant correlation between peer support and student preparedness (see Table 5). The results indicated that if students got more peer support, they might not be encouraged to increase the levels of preparedness for the course contents or assignments. Similarly, if students were prepared for the classes, their peers might not give them more social or academic helps in L2 classrooms. In other words, there was no systematic relationship between peer support and student preparedness.

Table 5. Results on correlation among climates and its components

	PEER	TEACHER	PREPARE	Total
PEER	1.0000	-----	-----	----
TEACHER	.1952*	1.0000	-----	----
PREPARE	.1080	.2569*	1.0000	----
TOTA	.4896*	.8020*	.5296*	1.0000

Notes: (1) PEER= peer; TEACHER= teacher support; PREPARE= student preparedness; TOTA= total score; (2) * P< .05

CONCLUSIONS

The present study examined senior high school students' perceptions of L2 classroom climate and the relationship between climate and its components including teacher support, peer support and student preparedness. The first finding was that climate had three components based on the results of the PCA. The result confirmed Hsu's (2001) and Fassinger's (1995) theories on climate. It is proved that L2 climate is a multidimensional construct. Besides, it was found that not only teacher behaviors (e.g., teacher support) and peer support but also student factors (e.g., students' preparedness) were positively correlated with climate. The finding implied that social support from significant others (e.g., teachers and peers) as well as students' preparedness may influence climate.

The results of the study indicated that senior high school students had a moderate level of climate. Moreover, there were significant differences between first year and second year students and between males and females in terms of their perceptions of L2 classroom climate. Minor findings showed there was no significant correlation between peer support and student preparedness but a positive correlation between peer support and teacher support.

Based on the results of the study discussed above, some pedagogical suggestions were provided for senior high teachers in CLT classrooms. First, if climate can be utilized as a predictor of teaching effectiveness (Hsu, 2001; Liu, 2006; Zheng, 1998) and motivation (Hsu, 2008; Wang, 2008; 蘇佳純, 2008), EFL teachers had better pay attention to their students' perceptions of climate and explore the crucial factors that influence climate. Students' perceived climate may determine their perceptions of instructors' teaching effectiveness. Furthermore, based on the results of the study, EFL teachers can help first-year students with

the review/ preview of course contents and assignments, and during L2 tasks more frequently. The increase of teacher support might enhance their preparedness and climate.

Several suggestions are made for the future climate studies. Future research utilizing the Climate scale in this study can take different cultural and socio-economic settings as well as diverse populations into consideration. For example, the present study did not investigate climate from the perspectives of the students in college/ universities, vocational or aboriginal senior high schools. It is also worthwhile to examine climate in urban and rural schools (Su, 2007). Besides, data collection triangulation (e.g., observation and interview) can also be used to remedy the weakness of the questionnaire. Since the present study did not investigate climate from EFL teachers' perspectives, future studies can compare the teachers' and students' perceptions of climate both qualitatively (e.g., making classroom observations and interviews) and quantitatively (e.g., distributing climate scales). In this way, we may get a more holistic understanding of L2 climate. Finally, future studies can examine the impact of climate on willingness to communicate, motivation, and teacher immediacy (Wang, 2008).

REFERENCES

- Adelman, H. S. and Taylor, L. (2005). Classroom Climate. In Lee, S. W. et al. (Eds), *Encyclopedia of School Psychology*. Thousand Oaks, CA: Sage.
- Andersen, G. J. and H. J. Walberg (1968). Classroom Climate and Group Learning. *International Journal of Educational Sciences*, 2, 175-180.
- Benson, P. (2001). *Teaching and Researching autonomy in Language Learning*. London: Longman.
- Creemers, B. P. M. and Reezigt, G. J. (1999) The Role of School and Classroom Climate in Elementary School Learning Environments. In H. J. Freiberg (Ed.). *School Climate. Measuring, Improving, Improving and Sustaining Healthy Learning Environments*. London: Falmer Press.
- Fassinger, P. A. (1995). Understanding Classroom Interaction: Students' and Professor's Contribution to Students' Silence. *Journal of Higher Education*, 66 (1), 82-96.
- Flanders (1960). *Teacher Influence, Pupil, Attitude, and Achievement*. Minneapolis, Minn.:University of Minnesota Press.
- Fraser, B. J. (1999). Research on Classroom and School Climate, In D. Gabel (ed.), *Handbook of Research on Science Teaching and Learning*. New York: MacMillan.
- Fraser, B. J. (1998). Classroom Environment Instruments: Development, Validation and Applications. *Learning Environments Research*, 1 (1), 7-33.
- Hsu, Y. T. (2008). Perceived Teachers' Instructional Behaviors, Students' Self-determinations, and Learning Climate in English Class. Papers Selected in *The Proceedings of 2008 International Conference and Workshop on TEFL and Applied Linguistics*, pp. 205-218. Taipei: The Crane Publishing Co., LTD.
- Hsu. W. H. (2001). *How Classroom Questioning Influences Second Language Acquisition*. Unpublished Dissertation, University of Illinois at Urbana-Champaign.
- Liao, W. C. (2005). *Senior High School English Teachers' Beliefs Towards Communicative Language Teaching and Their Classroom Practice*. Unpublished Master's thesis, National Taiwan Normal University.
- Moos, R. H. (1979). *Evaluating educational environments*. San Francisco: Jossey-Bass.
- Reeve, J. (1998). Autonomy support as an interpersonal motivating style: Is it teachable? *Contemporary Educational Psychology*, 23, 312-330.
- Trickett, E. J. and Moos, R. H. (1971). *Assessment of the Psychological Environment of the High School Classroom*. Stanford, CA: Stanford Uni, School of Medicine.
- Wang, Chang-o. (2008). A Preliminary Study on the Predictors of Situational Willingness to Communicate (SWTC) in an EFL Context. Unpublished Dissertation, National Taiwan

Normal University.

- 王雅觀 (1999)。《高級中學班級組織型態及其班級氣氛與學習狀況之研究》。台灣師範大學輔導研究所，碩士論文，未出版。
- 吳怡瑄 (2002)。《主題統整教學、教室氣氛、年級及父母社經地位與國小學童科技創造力之關係》。國立中山大學教育研究所碩士論文，未發表。
- 吳明隆 (2003)。《SPSS 統計應用學習實務—問卷分析與應用統計》。台北：知城數位科技。
- 吳武典 (1978)。國民中學班級氣氛問題。《台灣教育》，3 (9)：46-51。
- 吳福源 (1999)。《國民小學教師班級經營策略與班級氣氛之研究-優良教師與一般教師之比較》。國立政治大學教育學系博士論文，未出版，台北。
- 吳福源 (2000)。國民小學優良教師與一般教師之班級氣氛比較研究。《花蓮師院學報》，10: 171-195。
- 沈奕成 (2007)。《[國中導師領導風格、學生學習動機與班級氣氛知覺之研究](#)》。嘉義大學國民教育研究所，碩士論文，未出版。
- 汪榮才 (2006)。國小班級創造氣氛量表之發展與效度研究。《致遠管理學院學報》，1: 1-22。
- 林寶山 (1982)。《班級氣氛量表簡介》。輔導月刊，18 (11, 12)，40-41。
- 邱惠君 (2001)。從班級氣氛的研究取向談班級氣氛之營造。《教育社會學通訊》，34，23-31
- 張玉茹、林世華 (2001)。全語言教學在國中英語課之實驗研究。《師大學報：教育類》，46 (2)：233-253。
- 張玉茹、張景媛 (2003)。多元智慧教學與歷程檔案評量對國中生英語學業表現、學習動機、學習策略與班級氣氛的影響。《教育心理學報》，34 (2), 199-220.
- 張春興、林清山 (1990)。《教育心理學》。台北：東華。
- 郭秀緞 (2001)。《[高雄市國小六年級班級氣氛、學習動機與學業成就關係之研究](#)》。高雄師範大學教育研究所，碩士論文，未出版。
- 陳奎熹 (1997)。《教育社會化研究》。台北：師大書苑。
- 劉依玲 (2006)。《教育改革下國小英語教師專業認知與班級氣氛、教學效能之關係研究》。高雄師範大學教育研究所，碩士論文，未出版。
- 蔡璧煌 (1995)。《班級氣氛與學生政治社會化》。台北：師大書苑。
- 蔡麗娟 (2006)。《國中數學學習動機低落學生的學習與其心智型態之個案研究》。東華大學教育研究所，碩士論文，未出版。
- 鄭詩釗 (1998)。《國民小學班級經營氣氛、教室衝突管理與教師效能關係之研究》。台灣師範大學教育研究所，碩士論文，未出版。
- 鍾紅柱 (1983)。《高中班級氣氛之研究》。台灣師範大學教育研究所，碩士論文，未出版。
- 蘇佳純 (2008)。《學習動機、班級氣氛與學習滿意關係之研究—以花蓮地區國中小英語課為例》。國立東華大學教育研究所，碩士論文，未出版。

APPENDIX 1

Results of the First PCA for the Climate Scale

Analysis number 1 Listwise deletion of cases with missing values
 Kaiser-Meyer-Olkin Measure of Sampling Adequacy = .80811
 Bartlett Test of Sphericity = 1174.0105, Significance = .00000
 Initial Statistics:

Variable	Communality	*	Factor	Eigenvalue	Pct of Var	Cum Pct
G1	1.00000	*	1	4.90752	25.8	25.8
G10	1.00000	*	2	2.04750	10.8	36.6
G11	1.00000	*	3	2.01653	10.6	47.2
G12	1.00000	*	4	1.27996	6.7	54.0
G13	1.00000	*	5	.99013	5.2	59.2
G14	1.00000	*	6	.91542	4.8	64.0
G15	1.00000	*	7	.82678	4.4	68.3
G16	1.00000	*	8	.75910	4.0	72.3

APPENDIX 1: Continued

G18	1.00000	*	9	.73567	3.9	76.2
G2	1.00000	*	10	.70436	3.7	79.9
G21	1.00000	*	11	.56982	3.0	82.9
G22	1.00000	*	12	.55955	2.9	85.9
G23	1.00000	*	13	.52764	2.8	88.6
G24	1.00000	*	14	.46404	2.4	91.1
G3	1.00000	*	15	.44606	2.3	93.4
G5	1.00000	*	16	.35818	1.9	95.3
G6	1.00000	*	17	.32194	1.7	97.0
G7	1.00000	*	18	.28868	1.5	98.5
G8	1.00000	*	19	.281130	1.5	100.0

Hi-Res Chart # 4:Factor scree plot (PC extracted 4 factors).

(Factor Matrix)	Factor 1	Factor 2	Factor 3	Factor 4
G12	.73638	-.16813	-.15610	-.05076
G11	.72089	-.21552	-.18440	-.05558
G13	.71955	-.13928	-.31126	.03035
G2	.68092	-.08702	-.24083	.15670
G8	.65777	-.14113	-.31589	-.02860
G15	.62789	-.20895	-.13865	-.04833
G16	.60935	-.13474	.14477	.02926
G5	.60047	-.05562	-.32370	.03329
G18	.43890	-.11387	.40205	-.30224
G22	.30092	.70712	.13324	.10295
G23	.42492	.66940	-.06496	.00311
G24	.45431	.65095	.10525	-.11658
G21	.30603	.57255	.13202	.10325

G10	.28264	-.07746	.60382	-.23683
G1	.32981	-.24821	.56229	.02988
G3	.10478	-.27410	.55089	.36985
G7	.50030	.00924	.51262	-.30370
G6	.14964	.03454	.04934	.73419
G14	.26326	-.06587	.31135	.54079

APPENDIX 2**Results of the Second PCA for the Climate Scale**

Analysis number 1 Listwise deletion of cases with missing values

Kaiser-Meyer-Olkin Measure of Sampling Adequacy = .80186

Bartlett Test of Sphericity = 972.24226, Significance = .00000

Extraction 1 for analysis 1, Principal Components Analysis (PC)

Initial Statistics: (deleting items, 4, 9, 20, 17, 19, 6, 14, 16 and 18)

Variable	Communality	* Factor	Eigenvalue	Pct of Var	Cum Pct
G11	1.00000	* 1	4.37413	29.2	29.2
G13	1.00000	* 2	2.03800	13.6	42.7
G8	1.00000	* 3	1.78534	11.9	54.6

APPENDIX 2: Continued

G12	1.00000	* 4	.95212	6.3	61.0
G5	1.00000	* 5	.93655	6.2	67.2
G2	1.00000	* 6	.75342	5.0	72.3
G15	1.00000	* 7	.64535	4.3	76.6
G22	1.00000	* 8	.61155	4.1	80.6
G23	1.00000	* 9	.57716	3.8	84.5
G24	1.00000	* 10	.51217	3.4	87.9
G21	1.00000	* 11	.47188	3.1	91.1
G1	1.00000	* 12	.39154	2.6	93.7
G10	1.00000	* 13	.36308	2.4	96.1
G7	1.00000	* 14	.30313	2.0	98.1
G3	1.00000	* 15	.28459	1.9	100.0

OBLIMIN converged in 4 iterations.

Factor Matrix:

	Factor 1	Factor 2	Factor 3
G13	.80886	.00538	-.07520
G11	.77483	-.04716	.08651
G12	.75563	.00679	.11369
G8	.74683	-.01405	-.07909
G2	.71607	.05331	.01856
G5	.68027	.05230	-.11652
G15	.66084	-.05709	.10207
G22	-.10977	.79982	-.00899
G24	.06341	.77854	.06095

G23	.12023	.75577	-.09160
G21	-.03707	.66562	.04184
G10	-.04631	.10866	.73521
G1	.09414	-.05975	.69550
G7	.14235	.24176	.65373
G3	-.08477	-.14775	.61071

APPENDIX 3**Reliability of the Revised 15-item Climate Scale**

(Part 1) The Internal Consistency of the Scale Items on FLCC

	Scale Mean if Item deleted	Scale Variance if Item deleted	Total	Corrected Item Alpha if Item deleted
G1	47.7941	52.1150	.2582	.8004
G10	47.3235	52.7421	.2370	.8012
G11	47.5784	48.3534	.5915	.7756
G12	47.5931	47.6218	.6098	.7733
G13	47.7696	46.7201	.5763	.7744
G15	47.5980	49.9460	.4839	.7838
G2	47.5343	47.7771	.5690	.7761

APPENDIX 3—Part 1 Continued

G21	48.1324	51.9479	.2784	.7987
G22	48.7157	52.4508	.2800	.7978
G23	47.9657	50.5949	.3869	.7906
G24	47.9412	50.5285	.4213	.7881
G5	47.6127	48.7311	.4724	.7838
G7	48.0588	50.4891	.4429	.7867
G8	48.0196	48.6794	.4905	.7824
G3	47.5882	55.2385	.0492	.8146

Reliability Coefficients (N of Cases = 204.0; N of Items = 15; Alpha = .8003)

(Part 2) Internal Consistency of Subscale Items on Peer Support

	Scale Mean if Item deleted	Scale Variance if Item deleted	Total	Corrected Item Alpha if Item deleted
G7	10.9853	4.3594	.4517	.5198
G1	10.7206	3.9954	.4521	.5141
G10	10.2500	4.2574	.4314	.5315
G3	10.5147	4.6944	.2851	.6368

Reliability Coefficients (N of Cases =204.0; N of Items = 4; Alpha = .6227)

(Part 3) Internal Consistency of Subscale Items on Preparedness

	Scale Mean if Item deleted	Scale Variance if Item deleted	Total	Corrected Item Alpha if Item deleted
G22	9.6520	5.0162	.5831	.6726
G23	8.9020	4.7588	.5793	.6725
G21	9.0686	5.1775	.4463	.7485
G24	8.8775	4.9356	.5812	.6726

Reliability Coefficients (N of Cases = 204.0; N of Items = 4; Alpha = .7500)

(Part 4) Internal Consistency of Subscale items on Teacher Support

	Scale Mean if Item deleted	Scale Variance if Item deleted	Total	Corrected Item Alpha if Item deleted
G5	21.2892	20.7386	.5583	.8498
G2	21.2108	20.4036	.6301	.8391
G15	21.2745	21.8947	.5463	.8503
G11	21.2549	20.6933	.6753	.8336
G13	21.4461	19.0956	.7011	.8285
G12	21.2696	20.3260	.6740	.8331
G8	21.6961	20.4688	.6089	.8422

Reliability Coefficients (N of Cases =204.0; N of Items = 7; Alpha = .8594)

Passive Voice in Research Articles: Revisiting Rhetorical Functions

Hung-chun Wang (王宏均)

Hsin Sheng College of Medical Care and Management
rogerhcwang@hotmail.com

Passive voice is often taken as a less preferred choice of sentence construction than active voice in grammar handbooks and academic/technical writing manuals. However, recent interest in the passive has displayed that the function of passive voice has been oversimplified with respect to its rhetorical value in writing. In academic and technical writing, the passive is used to serve various rhetorical functions. Several important studies that argue for the rhetorical value of passive voice include Givón (1981) and Tarone *et al.* (1998), both of which oppose to the weakness and lifelessness of passive voice in writing. However, although they proposed several rhetorical functions of the passive in general and technical writing, their analyses were conservative in presenting the strength of passive voice. In this study, we analyze passive voice sentences in research articles of two academic disciplines, namely, applied linguistics and earth science. In each discipline, three research articles were collected from one internationally-prestigious journal, and so a total of six articles were analyzed to discern the rhetorical functions of passive voice following the Grounded Theory. Results show that the functions of passive voice are not limited to those already discussed in the existing research, while its functions vary to serve different functions in academic writing.

INTRODUCTION

Passive voice is often taken as a less preferred sentence construction than active voice. Compared with strength and vigor that an active verb conveys (Hacker, 2003), passive voice is weak and lifeless (e.g., Houpp *et al.*, 1995). Hacker (2003: 169) stated that “active verbs express meaning more emphatically and vigorously than their weaker counterparts... verbs in the passive voice.” Stunk and White (1979) also agreed that active voice sounds “more direct and vigorous” than passive voice. By contrast, passive forms are often worse choices because they lead to “lifeless and wordy writing” (Houpp *et al.*, 1995, p. 97). When the passive form is overused in writing, the writing will fail to express a clear picture of the writing style and responsibility (McCaskill, 1990). In this way, many experts or grammarians urge writers to make active voice a priority and minimize passive forms unless for some specific purposes.

Passive form is also viewed as a less preferred style than active form in technical writing. For instance, Lay *et al.* (1995) stated that active voice can make writing “much more direct and explicit” (p. 296). In the handbook of technical writing, Anderson (1995: 272) even made a suggestion that readers “use the active voice unless there is good reason to use the passive voice.” Thus a conspicuous contention in academic or technical writing is that writers should make active voice a priority, with passive voice appropriate only on specific occasions.

Despite the oft-claimed stigmatized status of passive voice in writing, it should be noted that debate over significance of passive voice has spurred many researchers to put forward a strong argument that passive voice construction is never inferior. On certain occasions, using passive voice is even a better choice than its active equivalent (Trammell, 1981). According to Beason and Lester (2000), passive voice construction must never be considered an error; so far no universal rules object to its use. Jenkins (2003) pointed out that passive voice is associated with responsibility for action or textual emphasis. In sentences like “The vase was

broken,” the speaker could skillfully avoid mentioning that he was the one that broke the vase. Also, in detective stories, “the manager was actually killed by the hairstylist” might convey a more surprising ending than its active equivalent, “the hairstylist actually killed the manager.”

Givón (1981) may be the first one that analyzed the functions of passive voice from a pragmatic perspective. Pragmatic functions of the passive can be broadly categorized as topic identification, impersonalization, and detransitivization. First, passive form is characterized as a syntactic construction that involves the non-agent moving to the subject position and the agent to the object position. The subject position is often considered focal, so when a receiver of an action is moved to the focal position, it stays in the spotlight of reader’s attention. Also, using the passive is thought of as an impersonalization strategy that helps the writer hide the identity of an agent, particularly when the agent is unknown, completely known, or irrelevant. For instance, the agent of (1a) can easily be obtained from the context.

(1a) English is spoken much in the European countries.

The third dimension, detransitivization, concerns demotion of the agent and promotion of the non-agent (Givón, 1993). Agent demotion refers to moving the agent from subject to object position when the agent is unknown, universal, or easily predicted from the context. Agent demotion can also help the agent eschew culpability, as in (1b).

(1b) He drove too fast last. A man was hit and died.

Still, non-agent promotion means moving a non-agent to subject position. With a non-agent promoted to subject focal position, it is topicalized and marked as the topic of discourse.

Focusing on scientific and technical writing, Tarone *et al.* (1998) proposed rhetorical functions of the passive in astrophysics journal papers. By analyzing two published articles in astrophysics, they discovered that the use of voice in sentence construction can be determined by the overall rhetorical structure of the writing, such as being a logical argument paper or an experimental paper. In astrophysics, journal papers feature logical arguments rather than research experimentation, and use of passive or active voice is influenced by this rhetorical style. Tarone *et al.* proposed that the passive is used in opposition to *we* + an active verb, particularly with respect to four functions in astrophysics journal papers:

(1) *we* indicates the author’s procedural choice, while the passive indicates an established or standard procedure; (2) *we* is used to describe the author’s own work and the passive to describe the work of others, unless that work is not mentioned in contrast to the author’s in which case the active is used; (3) the passive is used to describe the author’s proposed studies; and (4) the use of the active or the passive is determined by focus due to the length of an element or the need for emphasis. (p. 113)

Furthermore, Tarone *et al.* argue that the first three rhetorical functions of the passive and active may be applied to other fields in which research articles feature the form of a logical argument like astrophysics, while the fourth function of the passive should be considered a rhetorical trait of passive voice in general English writing.

RATIONALE BEHIND THE STUDY

In academic/technical English writing, a growing idea is that passive voice performs different rhetorical functions. However, a phenomenon worth discussion is that handbooks of academic writing often encourage writers to use active form as a preferred style. For instance, in the publication manual compiled by the American Psychological Association (2001), a clear statement is made to favor active voice construction in academic writing. Active voice

is preferred because it is direct enough to harmonize with verbs that act as “vigorous, direct communicators” (p. 41). Their explanation for use of passive voice follows the conventional argument that passive voice is used in expository writing to highlight the role of the recipient versus agent, or to elaborate on methodology of the study. Nevertheless, in contrast to their contention that passive voice is a less preferred choice, investigation of academic writing has reflected that academic writing abound in the prevalence of passive voice construction (Quirk *et al.*, 1985; Swales, 1990). Riley (1991) analyzed twelve scientific research texts and discovered that more than one-third of all main verbs used in the texts were in passive voice. This mismatch between prescribed usage of passive forms in publication manuals and actual use in academic papers reveals that passive should not be regarded as a bad or less preferred choice in writing. Further, it requires massive research work to explore functions of passive voice sentences in research papers within different academic disciplines.

Previous studies on cross-disciplinary differences often stand on the writing conventions prioritized under the style of each academic discipline. Writing style is considered to be “a social imperative” rather than “a pattern of individual choices” (Gross, 1991, pp. 934-935). This social condition empowers a set of rhetoric styles to work as the shared norm in each academic community. A common comparison that has resulted in numerous research in discourse analysis concerns the cross-disciplinary differences between two contrasting academic disciplines, hard and soft sciences. Becher and Trowler (2001: 36) stated that hard fields such as material sciences tend toward “impersonal, value-free”, soft fields like social sciences being more “personal, value-laden.” To fulfill the impersonalization favored in hard sciences, passive voice construction plays a particularly important role. For instance, Hacker (2003) highlighted that passive voice allows emphasis on design and process of experiments, rather than on the researcher of the study.

With hard science writing more impersonal and soft science writing more personal, we conduct a cross-disciplinary study to complete the picture of functions of passive forms in academic writing by investigating two academic disciplines as a whole. The central research question guiding this study is “What are the rhetorical functions of passive forms in academic writing?” By focusing on this issue, on one hand, we aim to discuss the rhetorical functions of the passive in academic writing in greater detail; on the other, we suppose more functions should be added to those already discussed in Givón (1981) and Tarone *et al.* (1998).

ACADEMIC CORPUS AND METHOD

Data are drawn from two academic corpora that consist of three research papers: the *Applied Linguistics Corpus* and *Earth Science Corpus*. All papers contained in these two academic corpora were retrieved from internationally prestigious journals (Appendix A). In *Applied Linguistic Corpus*, three papers collected from the 2007 issues of *TESOL Quarterly* were fed into this corpus. In the *Earth Science Corpus*, three papers were amassed from one international journal, *Engineering Geology*. Both these journals are listed in the Social Science Citation Index or Science Citation Index.

Generally speaking, writers in academia must possess knowledge of writing conventions required within their academic discipline. To get their papers published in international journals, all writers need to know the academic writing conventions that are taken to reflect a standard by researchers in their fields, whether or not English is their first language. Our study focuses on how passive construction is manipulated in those writing conventions across two academic disciplines, and whether a writer is a native speaker of English is not thought to influence results. The papers fed in our database were written by either native speaker(s) of English, non-native speaker(s) of English, or co-authored by both. In the six research papers, all the papers excluding abstracts, tables, figures, and direct quotes from previous research were entered into our corpora; the *Applied Linguistic Corpus* amounted to 23,822 words and

the *Earth Science Corpus* contained 12,510 words.

This study features investigation of passive voice in research writing across academic disciplines. Data analysis consists of the following steps. First, to identify passive forms used in each corpus, all articles were read carefully and each passive voice construction marked. Identification of passive voice in this study follows Riggle's (1998) analytic method, with which *be/get/feel/appear* + *past participle* construction was first identified as passive in the databases. Yet only construction with a direct grammatical active equivalent is considered in this study. For instance, (2a) could be changed into an active statement "The researcher *conducted* this study to investigate the influence of first language on L2 learning." By contrast, (2b) has no direct active counterpart, so it was not included in our data analysis.

- (2a) This study *was conducted* to investigate the influence of first language on L2 learning.
(2b) The design of the questionnaire *was based* on Johnson's framework.

An easy way to ascertain whether passive construction can be inverted into active is to see if adding *very* to the sentence destroys its grammaticality. If not, a passive verb must be used as an adjective. Also, if a supposed agent in the form of *by* + *agent* can be added to the sentence without influencing its grammaticality, it must be passive voice that we should include in our analysis. Stative passives that function as adjectives were excluded from our analysis. For instance, quoting Riggle's (1998: 92) examples, *unfunded* in (2c) is used as an adjective, and thus should not count as a passive verb. However, *funded* in (2d) is taken as a passive verb.

- (2c) It will be *unfunded*.
(2d) It will not be *funded*.

Once all passive constructions in databases were marked, a series of qualitative analyses were conducted. Following Grounded Theory (Glaser & Strauss, 1967), all the identified passive verbs were analyzed to generate a category of their rhetorical functions in academic writing.

RESULTS

Rhetorical Functions of the Passive

Central to this study is probing the rhetorical functions of passive voice constructions in two academic disciplines. Ten rhetorical functions, generated through a series of text analyses, will be addressed subsequently.

Modifying the noun phrase to supplement additional information. One common usage of passive forms is found in relative clauses that modify previous noun phrases and provide additional information value, which can be significant or unnecessary. Meanwhile, adding further information to the antecedent may convey non-literal meanings. In (4a), the passive form *not thoroughly exploited previously* shows the proposed advantage as seldom discussed in the past, and it further implies that this advantage of case studies may be a niche for research. An example in the *Earth Science Corpus* (4b) also shows that passive verb can supply additional information if embedded in relative clauses. In (4b), the writer discusses the function of "pre-consolidation pressure value"; passive form *was introduced by Casagrande* clarifies who first recognized this notion. Also, with "pre-consolidation pressure value" as the sentence-initial, it is the focus of the sentence. Using an active relative clause to modify it, "which Casagrande introduced" weakens the focus on "pre-consolidation pressure value."

- (4) a. (Applied linguistics research paper)
Another important advantage of case studies *not thoroughly exploited previously* is

- their potential for highlighting the student writers' own perspective (Li, 2007, p. 56).
- b. (Earth science research paper)
Pre-consolidation pressure value, P_c , which was introduced by Casagrande (1932, 1936), "is generally considered to indicate the maximum load on a natural soil during its geologic history, although the effect may also be caused by aging" (Crawford, 1986, p. 80). (Reznik, 2007, p. 27)

Drawing readers' attention to an important issue. Writers at times attempt to draw readers' attention to a specific essential issue that will be discussed below, and using some formulaic sequences that are in passive forms can successfully direct readers' attention to the focus. These formulaic sequences including *it has been shown*, *it should be noted*, and *it is believed* are always followed by a *that* clause. In (5a), Li (2007) discusses a graduate student's progress from a beginner to mature writer in his discipline. Li initially presented the trend of previous research in the first sentence. The use of *it has also been shown that* in the second sentence directs readers' attention to another focus in the existing research. Similarly, (5b) and (5c) indicate two examples of this function, i.e. *it should be noted* and *it was not considered necessary to*. These examples make readers understand the writer's opinion as reflected by words *noted* and *necessary* first, then draw their attention to the focus—the issue that should be noted or considered necessary.

- (5) a. (Applied linguistic research paper)
Previous studies that investigated graduate disciplinary writing have tended to focus on how to local curriculum setting impinges on the students' writing process to explicate the social nature of graduate writing. Yet *it has also been shown* that as a graduate student writer transforms from "newcomer" to "embryonic researcher" and then... (Li, 2007, p.72)
- b. (Applied linguistic research paper)
It should be noted that because the context of this research was not an L2 writing class, *it was not considered necessary to* ensure that the narrative task constituted an ecologically valid writing task. (Sheen, 2007, p. 263)
- c. (Earth science research paper)
The results of these analyses (Fig. 5) indicate that use of vertical accelerations in slope stability assessment the scatter in the data is not minimized. This inference is in qualitative agreement with that deduced from the data presented in Fig. 4c. *However, it should be noted* that site specific vertical accelerations were not available in most cases for these analyses. (Singh *et al.*, 2007, p. 175)

Topicalization. Givón (1981) stated that when a noun is moved to subject position in a passive voice sentence, it may be identified as the topic. Topicalization signals that the noun in subject position is the focus of writing and will be further elaborated in the following text. In both academic corpora, we also found examples that manipulate topicalization to highlight a topic. In (6a), Mackey *et al.* (2007) addressed some limitations of their study. In the first sentence, they topicalized the focus, *the limitations of the current study*, which signals that this focus will be further elaborated. One interesting example as shown in (6b) reveals that topicalization may remind readers of a topic when the first identification of topic is a distance away. In this example, the first sentence shows that the graduate student's *critical orientation* will be discussed in the text. The noun phrase *his orientation* in the second sentence confirms the graduate student's *orientation* as the topic and takes up several more lines for explanation. In the last sentence, because the previous topic indicator, *his orientation*, is a distance away, so readers may easily lose focus. The writers then again

topicalize the focus *his interaction* so as to keep this focus fresh in a reader's mind. Last but not least, (6c) shows that topicalization can increase the cohesion in writing. The subject in the second sentence, *the derived formulas*, directly link to the object in the previous sentence and establish strong unity and cohesion.

(6) a. (Applied linguistics research paper)

In considering recommendations, however, the limitations of the current study must *be noted*. With only 40 learners, the study was relatively scale. In addition, because this was fundamentally a classroom, the lack of randomization is a factor...In addition, because of the small number and constrained type of tasks used in this study, it remains to be seen how other task characteristics (e.g., task type and complexity) interact with the various types of familiarity. (Mackey *et al.*, 2007, p. 307)

b. (Applied linguistics research paper)

Yuan's engagement with the global specialist research community also sometimes showed a *critical orientation*. *This orientation* shows in his rather cynical view of the need to impress the referees (who were, in his words, probably "snobbish") by boosting his work, that is, highlighting its "bright points" and citing a previous publication of his home group ("Hehe, citing our own work once, so that others know our group has also produced something—we're not nuts"). It also shows in his framing of the need to meet the expectations of the target journal (e.g., supplying a statement for "Safety Consideration" because "Foreigners are quite humanistic".) *His orientation* is also seen in his interaction with journal articles, where he had a critical eye in reading the texts of "big hands" to figure out the secrets of rhetorical argumentation in the genre of RAs. (Li, 2007, p. 72)

c. (Earth science research paper)

This paper presents an analytical method describing the relationships between physical properties of collapsible soils and their mechanical properties. The derived formulas are *illustrated by* published experimental results. (Reznik, 2007, p. 28)

Impersonalization. The fourth function of passive voice in academic writing is to impersonalize the researcher by completely hiding or weakening his or her identity as the conductor of the study. In (7a), Mackey *et al.* (2007) stated that adults are like children: they may easily get bored if asked to participate in the same tasks again and again. Because the source of adults' boredom in this case arises from researchers' failure to provide them with a variety of tasks, by omitting the agent *by researchers*, the writer attempts to avoid discussing the failure of the researchers.

Impersonalization can also be found in *Earth Science Corpus*, wherein research papers often report natural phenomena that researchers observed, whose writing convention is that writers tend to impersonalize observations without researchers' involvement. In Lee *et al.* (2007), the researchers discussed the problem of reservoir leakage in Korea in 1997 and 1998. Though all natural phenomena were surely observed and noticed by researchers, the writers showed inclination toward passive forms to hide involvement of the researchers in the report.

(7) a. (Applied linguistics research paper)

Like all children, those in second language classrooms can also become bored if *they are asked to* repeat identical tasks, or even the same ask type too many times. (Mackey *et al.*, 2007)

b. (Earth science research paper)

No distinctively visible reservoir leakage *was observed* before December 1997, except for the settlement holes, even though the reservoir level rose to nearly the

normal water level between June and August 1997. In December 1997, a large amount of water *was noticed* at a downstream area when workers were excavating to install a seepage measuring weir. The first leakage data (about 2000 m³ /day) *were collected* in March 1998 and the reservoir level rose to the normal maximum level of 150 m (Lee *et al.*, 2007, p. 303) .

Concluding existing research. Passive forms are also used to provide readers with a general research trend or direction conducted by previous researchers. In (8a), Mackey *et al.* (2007) discussed task familiarity and recognized many research studies that have investigated different types of task familiarity. They further elaborated on it, specifying three trends in this direction, including familiarity with one's interlocutor. On the other hand, in (8c), Lee *et al.* (2007) stated that many previous researchers showed concerns towards the leakage problem. Passive form presents researchers' concerns as a general phenomenon noticed by numerous persons in this field. Active counterpart like "Johnson (1980) and Swales (1997) have raised concerns that S + V" simply implied that only some researchers raised the concerns.

(8) a. (Applied linguistics research paper)

Different types of task familiarity *have been explored* in previous research, including (a) familiarity with one's interlocutor, (b) experience with carrying out a particular task or type of task before (also known as *task repetition* or *rehearsal*), and (c) knowledge of the subject matter or content that is the focus of the task. (Mackey *et al.*, 2007, p. 289)

b. (Earth science research paper)

Concerns were raised that some portion of the leakage through the dam might bypass the current leakage measurement system, which consists of a cutoff wall (collecting barrier) with V-notch weir arrangement at the downstream toe of the dam (see Fig. 1b), probably because the cutoff wall does not extend sufficiently into the left abutment (Hayward Baker International, 2000). However, the base of the cutoff wall as constructed has been taken down to about 1m below the foundation rock, and the ground and rock level rises toward the left abutment. (Lee *et al.*, 2007, p. 303)

Objectivity: Stating self opinions. Academic writers' ultimate goal is to have their papers accepted by other people involved in the same profession. To this end, it sometimes is important to make one's writing sound objective to other experts. In research papers, we have found that writers often embed their opinions in passive sentences in order to establish high objectivity in their statements. In (9a), as the last sentence of the article, Li (2007) stated that his study shed light on teaching of TESOL and TESOL professionals. Active sentence "I hope that the study has been a useful contribution" focuses on the writer's subjective impression of his own study. Unlike the active equivalent, passive form *it is hoped that* makes the author's statement sound more objective.

Two interesting examples appear in (9b) and (9c). In (9b), Li discussed that he discovered the graduate student's borrowing is close to plagiarism; Li chose to mitigate his language by using passive form that reveals a *possibility* that it is plagiarism, instead of the active form "we may understand Yuan's 'borrowing' from the literature" directly indicating plagiarism. Also, in (9c) Lee *et al.* (2007) intended to provide a summary of the findings they noticed. Still, using an agentless passive form increased the objectivity of the research findings. Taken together, all three examples show academic writers sometimes do not state their opinions subjectively and so embed these opinions in passive sentences to show higher objectivity.

(9) a. (Applied linguistics research paper)

In this sense, *it is hoped that* the study reported in this article has been a useful

contribution. (Li, 2007, p. 74)

b. (Applied linguistics research paper)

In addition, although Yuan's "borrowing" from the literature, apart from the specialist terms, *may be understood as* plagiarism in academic community, to Yuan himself it was all normal and fine. (Li, 2007, p. 72)

c. (Earth science research paper)

Based on the findings from the sequential tracer tests, the following conclusions about the dam seepage or seepage pathways *were reached*. (Lee *et al.*, 2007, p. 314).

Objectivity: Focusing on the experiment. One writing style worth noting is that a great number of passive statements are used in both academic disciplines, particularly in the methodology section. In general, researchers attempt to show high objectivity when depicting experimental design or procedure, and using passive forms can effectively direct readers' attention to details about the experiment. In the following examples, most sentences describe step-by-step instruction of materials design or data collection. Writers in both disciplines used a great deal of passive voice when presenting design of the study.

(10) a. (Applied linguistics research paper)

Then the learners who had remained in the classroom *were taken out* and *familiarized with* the procedure of a different task. Practice tasks *were done* twice, one week before the experimental tasks *were carried out*. At the conclusion of the study, the learners *were asked* whether they had ever done an activity of that type... (Mackey *et al.*, 2007, p. 293)

b. (Earth science research paper)

An alternative correlation for estimating deformation of an embankment dam based on the observation from the case history database on the observations from the case history database *is proposed* relating the average...An attempt has also *be made* to identify the main causes of scatter in the data. (Singh *et al.*, 2007, p. 175)

Obscuring responsibility. Passive voice statement sometimes obscures responsibility, thus helping the writer avoid direct criticism on his or others' incompetence. In (11a), Sheen (2007) quotes Ferris's (1999) finding that most students prefer their writing errors corrected and that teachers should correct student errors. Passive form in this sentence attributes this result simply to student's preferences without any direct involvement of the teacher. However, if this sentence was changed to active, "most students do want their teachers to correct their writing errors," focus is on teacher role and to some extent implies a problem associated with teachers: that they show no attempt to correct student errors at times. The following sentence would then turn into a threatening advice. On the other hand, passive voice is also found in earth science research to obscure the responsibility. Reznik (2007) adopted his technique to analyze stress-strain curves, but later he recognized that his technique failed to interpret all data. Instead of using active voice sentence, Reznik chose to use a passive form which has the subject *all the test results* stay in the spotlight, and attributes failure of interpretation to test results. An active sentence like "The author's method could re-interpret not all test results" poses direct criticism of Reznik's technique for its inadequacy or immaturity.

(11) a. (Applied linguistics research paper)

Criticisms in this vein include practical concerns about the quality of the CF that teachers provide and also the danger of raising the students' *affective filter*. However, Ferris (1999, 2004) points out that *most students do want their writing errors to be corrected* and that it is the job of L2 writing teachers to attend to their needs. (Sheen,

2007, p. 258)

b. (Earth science research paper)

The stress-strain curves obtained for Hawkeye loess *were interpreted* using the technique offered by the Author (Reznik, 2005). Using Eq. (17a), the following empirical relationships between σ_{sz} and S were obtained (Figure. 6)..... *It has to be noted that* not all test results *obtained by Kane could be re-interpreted* using the Author's (Reznik, 2005) method for calculation of structural pressure values. (Reznik, 2007, p. 34)

Pinpointing the responsibility-taker. Unlike the function of obscuring responsibility, one interesting function found in the applied linguistics corpus is that passive voice can be used to point out who needs to be responsible for a specific issue, usually negative. Li (2007) quoted the participant's recall that his first English writing received very bad comments from his supervisor. The verb *criticize* connotes a rather negative meaning, and the passive form attributes this problem to Yuan's personal incompetent writing skills and the poor quality of his writing. Criticism came down to the fact that his writing failed to meet the standard of his supervisor. If the sentence were written in active form, "his supervisor had severely criticized his very first English language writing," this sentence emphasizes traits of the supervisor and weakens problems associated with the writing itself. It relates to the supervisor's problem, for instance, the supervisor was rather picky or hard to please, although Yuan's first writing was good enough. In this way, passive voice construction can sometimes help writers to clarify who needs to take responsibility for negative action. However, we did not have any example that shows this function in the earth science corpus.

(12) a. (Applied linguistics research paper)

Yuan recalled that his very first English language article *had been severely criticized by his supervisor* in the master's program. (Li, 2007, p. 65)

Indicating figures, tables, or other related information. Passive voice also tells the reader where to find related information, such as figures and tables. Passive verbs are often associated with figures and tables supplying further information concerning the experiment. For instance, (13a) is quoted from the methodology section, and this example described two tasks adopted in Mackey *et al.*'s (2007) study. The passive sentence directs readers' attention to a follow-up figure that clearly vignettes the experimental procedure in this study. In (13b) and (13c) passive forms also perform the same function by indicating whether readers can find information concerning summary of results and outline of procedures in both studies.

(13) a. (Applied linguistics research paper)

One of these tasks had familiar content, and the other had unfamiliar content. This design (and the counterbalanced order of presentation for the tasks) *is illustrated* in Figure 1. (Mackey *et al.*, 2007, p. 294)

b. (Applied linguistics research paper)

These results *are summarized and shown* in Table 5. (Mackey *et al.*, 2007, p. 299)

c. (Earth science research paper)

An outline of the procedures for parameter estimation and data analysis *is provided* in the following subsections. (Singh *et al.*, 2007, p. 175)

CONCLUSIONS AND DISCUSSION

This study operationalized a series of text analyses to discern the rhetorical functions of passive voice sentences in academic writing. Investigation of two academic corpora, each of

which comprised three papers published in one internationally prestigious journal, demonstrated that rhetorical functions of passive forms should go beyond those cited in Givón (1981) and Tarone *et al.* (1998). We discovered that their findings do not account for passive usage in academic writing, and at least ten rhetorical functions have been identified in this study. Note also that uses we identified are not detached from one another but overlap to some extent.

As we stated at the outset, many writers of grammar and writing handbooks suggested that active voice statements be regarded as preferred writing style, passive used cautiously owing to its stigmatized label of “illogical” and “unstylistic” (Pruitt, 1968) and/or resulting in “lifeless and wordy writing” (Houp *et al.*, 1995, p. 97). Our investigation of research papers in applied linguistics and earth science, however, reflected that passive voice is no minority; it pervades our research papers for discourse functions like showing objectivity and obscuring responsibility. These findings are inspiring because they expose inadequacy of grammar and writing handbooks concerning use of passive verbs in writing, especially in some publication manuals like the APA manual. Care should thus be taken to present to L2 learners and writers how passive forms are adopted, notably as an essential teaching component in EAP-related writing books and classes. This study explores rhetorical functions of passive forms, its chief to afford a clearer picture of how passive voice is used in academic writing, as a complement to earlier discussion in Givón’s (1981) and Tarone *et al.* (1998). However, in its present form, this study is limited in the sense that only three papers in each discipline, totaling six papers, were analyzed as data sources. Future research can take more studies or academic disciplines taken into consideration, bringing to light more rhetorical functions of passive voice.

REFERENCES

- Anderson, P. V. (1995). *Technical writing: A reader-centered approach* (3rd ed.) Fort Worth: Harcourt Brace.
- Beason, L., & Lester, M. (2000). *A commonsense guide to grammar and usage* (2nd ed.). Boston, MA: Bedford/St Martin’s.
- Becher, T., & Trowler, P. (2001). *Academic tribes and territories* (2nd ed.). Buckingham: The Society for Research into Higher Education and Open University Press.
- Jenkins, D. R. (2003). Active voice? Usually best. Passive voice? Sometimes better. *Writing*, 25(5), 26-27.
- Givón, T. (1981). Typology and functional domains. *SL*, 5, 163-193.
- Givón, T. (1993). *English grammar: A function-based introduction* (Vol. 2). Philadelphia: John Benjamins Publishing Company.
- Glaser, B. G., & Strauss, A. (1967). The discovery of grounded theory and applying grounded theory. In N. K. Denzin & Y. S. Lincoln (Eds.), *The American tradition in qualitative research* (Vol. 2, pp. 229-243). Sage publications.
- Gross, A. G. (1991). Does rhetoric of science matter? The case of the floppy-eared rabbits. *College English*, 53, 933-943.
- Hacker, D. (2003). *A writer’s reference* (5th ed.) Boston: Bedford/St Martin’s.
- Houp, K. W., Pearsall, T. E., & Tebeaux, E. (1995). *Reporting technical information* (8th ed.). Boston: Allyn and Bacon.
- McCaskill, M. K. (1990). *Punctuation, and capitalization: A handbook for technical writers and editors*. Washington, DC: National Aeronautics and Space Administration, Office of Management, Scientific and Technical Information Division.
- Pruitt, J. D. (1968). Passive voice should be avoided by research writers. *The Journal of High Education*, 39(8), 460-464.
- Publication manual of the American Psychological Association (5th ed.) (2001). Washington, DC: American Psychological Association.

- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. New York: Longman.
- Riggle, K. B. (1998). Using the active and passive voice appropriately in on-the-job writing. *Journal of Technical Writing and Communication*, 28(1), 85-117.
- Riley, K. (1991). Passive voice and rhetorical role in scientific writing. *Journal of Technical Writing and Communication*, 21(3), 239-257.
- Swales, J. (1990). *Genre analysis*. Cambridge: Cambridge University Press.
- Tarone, E., Dwyer, S., Gillette, S., & Icke, V. (1998). On the use of the passive and active voice in astrophysics journal papers: With extensions to other languages and other fields. *English for specific purposes*, 17(1), 113-132.
- Trammell, M. K. (1981). Enlightened use of the passive voice in technical writing. In J. C. Mathes & T. E. Pinelli (Eds.), *Technical communication: Perspective for the Eighties*. Virginia: National Technical Information Service.

APPENDIX: List of research papers under investigation

Applied Linguistics

- Li, Y. (2007). Apprentice scholarly writing in a community of practice: An intraview of an NNES graduate student writing a research article. *TESOL Quarterly*, 41(1), 55-79.
- Mackey, A., Kanganas, A. P., & Oliver, R. (2007). Task familiarity and interactional feedback in child ESL classrooms. *TESOL Quarterly*, 41(2), 285-312.
- Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, 41(2), 255-283.

Earth Science

- Lee, J-Y., Kim, H-S., Choi, Y-K., Kim, J-W., Cheon, J-Y., & Yi, M-J. (2007). Sequential tracer tests for determining water seepage paths in a large rockfill dam, Nakdong River basin, Korea. *Engineering Geology*, 89, 300-315.
- Reznik, Y. M. (2007). Influence of physical properties on deformation characteristics of collapsible soils. *Engineering Geology*, 92, 27-37.
- Singh, R., Roy, D., Das, D. (2007). A correlation for permanent earthquake-induced deformation of earth embankments. *Engineering Geology*, 90, 174-185.

Triggering Classroom Interaction with Teachers' Questions: A Taiwan Case Study

Liang-chun Wang (王亮鈞)
National Chengchi University
97551021@nccu.edu.tw

Questioning is a common situation in classroom interaction, and, particularly in the EFL context, many teachers prefer to use questions to arouse involvement of students in the class and make sure students' understanding. This case study attempted to observe two classes, in a junior high school of southern Taiwan, and investigate which type of the questions would be used more. On the other hand, the interactinal aspects and modes, triggered by certain characteristics of questioning, would be investigated. The result showed that display questions, especially with "Predictable IRF", have been used the most in both classes, which suggested that this type of questions can trigger more "whole class responses", and the tone of voice quality was also higher than others. However, referential questions, particularly with "Negotiation of Meaning" can trigger more variations of interactional aspects, so that more interesting and individual discernment of learning can be found and valued. The implications of this case study were that teachers should be encouraged to use more referential questions in order to make classroom interaction more dynamic, and the modes of interaction should also be taken into account to adopt "Negotiation" and "Navigation" and give learners more opportunities to be engaged in language teaching and learning.

INTRODUCTION

Just as Gall (1970) emphasized the importance of questions and teachers' frequent use of questions during an average school day, questions are well-known to be used in second or foreign language classrooms, and many teachers would like to exploit questioning skills to arouse students' responses or even some interactions among students.

On the other hand, classroom interaction would be a vital element in language instructions, since it can help students more effectively process language knowledge or feel that they are in the real-world conversations. Furthermore, Pica (1987) proposed that the interaction through classroom discourse among teachers and students is intended to "seek clarification or check comprehension of each other message meaning" (p.3). This perspective explains more deeply and detailedly the functions for classroom interaction, especially for the interaction triggered by teachers' questions.

However, in the Taiwanese EFL context, most of the time, teachers usually use a great deal of code-switching or code-mixing, such as English, Chinese, or Taiwanese for talk or questioning. There may be some differences from other ESL or EFL contexts in other countries, and we can not be sure whether or not this and other features of the EFL teachers' questioning can really help students be involved in classroom interaction or assist their foreign language learning. For this reason, this present study attempts to propose some questions of how non-native English teachers in the Taiwanese EFL context ask students questions for classroom interaction. Additionally, we will investigate the characteristics of teachers' questions and evaluate whether these characteristics would trigger certain classroom interaction from every aspect or mode.

LITERATURE REVIEW

In order to make clear the structure of rationale in this study, we need to overview the previous studies related to teachers' questions as well as classroom interaction. Many studies have noticed the importance of teacher talk in the classroom, and they also discuss various types or the most appropriate proportions of teacher talk (Cullen, 1998; Gall, 1970; Roth, 1996; Thornbury, 1996).

Some of the research put more emphasis on the characteristics of the holistic aspects from teacher talk. For instance, Cullen (1998) and Thornbury (1996) discussed the "communicative" characteristics of teacher talk. According to their studies, firstly, in terms of the types of teacher talk, the use of "**referential questions**", such as "What did you do at the weekend?", where the teachers do not know the answer, may have a more genuine communicative purpose than the exclusive or excessive use of "**display questions**", where the teacher already has the answer. These display questions are only exploited for the class to display their understanding or knowledge. Secondly, with regard to teachers' feedback, "**content feedback**" by the teacher, where the teachers' response to student contributions, focuses on the content of what the student says, will be more effective than "**form-focused feedback**". Thirdly, "**the use of speech modifications, hesitations, and rephrasing**" in the teacher's own talk, e.g. when explaining, asking questions, giving instructions, etc., can help students' learning more than only "**echoing**" of students' response when the teacher repeats what a student has just said for the benefit of the whole class. Lastly, "teachers' attempt to **negotiate meaning** with the students", e.g., requests for clarification and repetition or opportunities for students to interrupt the teacher are more helpful than merely sequences of "**predictable IRF (initiation-response-feedback)**" discourse chains in which the teacher initiates the chain (asking a question), a student responds, and the teacher then gives feedback.

Although these two studies only provided general features of teacher talk instead of specific observation of teachers' questioning, we can still find that teachers' questions should also obey the "communicative" disciplines to make learners have more opportunities to communicate with peers or teachers and arouse the interaction of the classroom. Another concern of these two studies is that we can not be sure whether or not these communicative characteristics of teacher talk can be totally applied to the Taiwanese EFL context or whether we should have other considerations or modifications for the specific context.

Other scholars paid their attention to the use of teachers' questioning along with the questioning behavior's influence on students' learning or response. Ellis's study (1993) presented the study of how teachers' questioning behavior influences students' learning. The author also discussed the types of questioning designed by Cunningham (1987)—Factual recall, conceptual, convergent, divergent, and evaluative—to see which type of questioning may really help students' learning. On the other hand, she also theorized a well-organized set of questioning strategies—open, focus, analysis, keystone, individual, application (OFAKA)—to evaluate the effectiveness of questioning. Although the author intended to focus on the effectiveness of questioning on students' learning, yet we can hardly see any empirical and concrete evidence for this aspect. Nevertheless, the viewpoint of "questioning strategies" can be regarded as the enhancement for the quality of teachers' questions as well as the help for classroom interaction.

The above studies centered on teachers' talk or questions, but, to some degree, ignored the importance of the concept of classroom interaction. Nonetheless, there are still some studies noticing this aspect in classroom discourse. Lynch (1991) discussed the role of questioning among teachers and learners in EFL contexts. The author criticized the aspect of questioning, involving educational settings to ask a question to which learners have already know the answer. He also compared three groups of learners with various language

proficiencies from post-elementary, intermediate, to advanced, to investigate the differences of the proportions of contribution between teachers and learners. The result showed that advanced learners can make more percentages of talk in class, while the lower leveled students required more time to familiarize the teaching materials in order to be independent of teachers' prompting. However, in this study, it may not be easy to discover whether the differences of students' talk were due to other variables, such as teachers' teaching styles or the difficulty of teaching materials. Furthermore, a higher percentage of students' talk may not definitely indicate more interaction between teachers and students or students and students.

Another study written by Gall (1970), based on the comparison of the previous studies, outlined a set of criteria for students' responses to teachers' questions: (a) complexity of the response; (b) use of data to justify or defend the response; (c) plausibility of response; (d) originality of the response; (e) clarity of phrasing; (f) the extent to which the response is directed at the asked question. With these criteria, teachers can verify whether their questions do have some assistance or influence on students' learning behavior. Furthermore, this study pointed that the educational objective should increase the frequency and quality of students' questions in the classroom context for interaction. However, we can not find out the relationship between students' responses and different types of questions so that the phenomenon of classroom interaction might not be easily categorized and evaluated.

Also, the study by Shomoossi (2004) conducted a qualitative-quantitative study as to the two types of questions, **referential and display questions**, to investigate which type may trigger more classroom interaction. The finding indicates that display questions were used by the teacher more than referential ones. Moreover, it shows that though referential questions may trigger more classroom interaction, not all referential questions can create enough interaction. The referential questions with a shorter answer and followed by a period of silence or topic-change can indeed trigger interaction. However, from this study, we found that it only focused on the types of questions, and the analysis of the interaction triggered by referential questions was too vague to prove the supporting position, claiming that more interaction can be triggered by referential questions. Nevertheless, the intention of the author to point out the importance of teacher's questions may have valuable implications for further studies. On the other hand, there still exists another great difficulty to analyze the classroom interaction brought by teachers' questioning. Roth's study (1996) provided a plausible explanation for this problem. The author described a case study of an expert teacher's questioning strategies during an open-inquiry engineering curriculum in a Grade 4/5 mixed French immersion classroom. The finding shows that the interaction of content, context, and students' responses to questions were characterized as complexity. Additionally, the teacher's competence in questioning was related to her discursive competence in the subject-matter domain, so that the practice of appropriate questioning may not be easy to achieve. From this study, the difficulty of analyzing classroom interaction triggered by teachers' questions would be influenced by the complexity of the interaction itself as well as teachers' language competence. However, another factor to affect the difficulty of analysis may be the lack of clear designed criteria for classroom interaction.

In terms of classroom interaction, Allwright's study (1984) elaborated more on this issue and proposed four arguments used in favor of getting learners "communicating." First, getting learners communicating represents a necessary and productive stage in order to transfer classroom learning to the outside world. Second, getting learners communicating is important for pedagogical thinking, since it involves person-to-person or face-to-face communication in order to solve problems. Third, advocating "communication" would be the best way with regard to "involvement" and "investment", which can help learning become more effective and make learners be involved in their learning. Fourth, supporting

“communication” in the classroom is focused on peer discussion, which can make learners share various viewpoints with each other.

However, the most significant contribution of this study may be the provision of the four modes of participation and five aspects of interaction management for the analysis of classroom interaction. As for the four modes of participation, there are “*Compliance*” (learners simply do whatever is asked or expected by co-operative learners), “*Direction*” (learners totally follow teachers’ direction), “*Negotiation*” (any attempt to reach the consensus of the decision among students and the teacher), and “*Navigation*” (attempts to steer a course over the obstacles which the lessons present for the students and the teacher). Indeed, compliance and direction may be the two extremes, while negotiation and navigation occur in between. With the four modes of participation, we can use to analyze the extent of involvement and investment of classroom interaction.

On the other hand, the five aspects of interaction management refer to turn, topic, task, tone, and code. Turn-management means the contribution that each participant made, topic-management implies the content of each contribution, task-management refers to the demand one of the participants will make on others, tone-management implicates the significant business of establishing the appropriate socio-emotional atmosphere, and code-management refers to the decision of language use by participants, including regional accent, register, or code-switching. With these clear aspects of interaction management, we can more specifically analyze and categorize the various features of classroom interaction as well as the internal structure of interaction triggered by teachers’ questions.

Based on the overview of teachers’ questions and classroom interaction, the previous studies, sometimes, focused only on a single angle of this important issue without observing the interrelation of teachers’ questions and classroom interaction. Even though some of them posited the vital importance of classroom interaction led by distinct questioning types, the analysis may not be sufficiently concise and specified. Due to the deficiency of the clear analysis in question-triggered interaction and the interrelation of the two important factors in classroom discourse, the present study will focus on the investigation of interrelation between teachers’ questions including different types, foci, functions and classroom interaction containing four modes and five aspects of interaction management in the Taiwanese EFL context. Also, we would like to find out whether there are salient differences of question-triggered interaction among different EFL and ESL contexts so as to provide some pedagogical implications on the current Taiwan English education.

RESEARCH QUESTIONS

1. What types of questions do Taiwanese non-native teachers use?
2. How often do Taiwanese non-native teachers use each type of questions?
3. What characteristics do teachers’ questions possess?
4. What kind of classroom interaction may be triggered by questions?

METHODOLOGY

Participants

The participants were randomly chosen for Ming-Syong junior high school in southern Taiwan, containing two non-native Taiwanese English teachers and two classes of students—one was 32 male students; the other was 40 female students. All of the students were in the first grade. There was no operation in the two classes and the two teachers were their original English teachers.

Design and Procedures

The classroom observation was conducted within one month for three sessions of each class, and each session retained 40 minutes. There was no special instruction or guidelines for

participants to take part in the observation procedure. During the session, we took notes for the important situations of the classroom interaction. The whole process of observation was all audio-taped and some of the interaction situations were video-recorded. After the observation, one questionnaire was answered by the two teachers in order to elicit some of the reasons for their styles and ways of the questions as well as their belief for questioning in classroom discourse.

DATA COLLECTION AND ANALYSIS

With regard to the section of data analysis, it is twofold. The first part, we would like to investigate the characteristics of the questions which were extracted from two classes and observe the frequency of each type of questions in the two classes.

In terms of the criteria of the questions, based on the features of teacher talk in Cullen (1998) and Thornbury (1996), we designed four categories of characteristics, each of which possesses two levels. Some of the definitions may be modified in order to fit the context of this study.

Firstly, as for “**Type**”, we divided the questions into “**Display questions**”(1-DQ), referring to the questions to which teachers know the answer(s) *very certainly*, the answers of which may be relatively fixed and less alternative, and the functions of which are roughly used for students’ displaying the knowledge in textbooks. The other level of the “Type” characteristic is the “**Referential questions**” (2-RQ), concerning the questions to which teachers do not know the answer *totally or very certainly*, the answer of which may be comparatively alternative, and the general function of which may be used for students to provide their opinions as well as different perspectives towards the same question.

Secondly, the characteristic of “Focus” can also be separated into “**Content-focused**” (C), implying the focus of the questions is on meaning, and “**Form-focused**” (F), referring to the focus of the questions containing grammatical rules and applying metalinguistic terminologies to explain the usage of English.

Thirdly, the way of the questions may be diverse, and, therefore, we provided “**Non-Echoing**” (NE), referring to the way of the questions, structured as explaining, asking questions, or giving instruction without repeating what students earlier said in their previous answers, and the other “**Echoing**” (E), referring to the way of the questions, structured from repeating parts of what students said in the previous answers.

Fourthly, as to the process of the questions, we can classify them into “**Predictable IRF**” (P), referring to the process of the questions, belonging to common predictable discourse chains. For example, the teacher asked a question, then students give a predictable form of answers, and later the teacher may or may not offer some feedback to students’ answers, and “**Negotiating of Meaning**”(NOM), referring to the process, comparatively unpredictable and intended to request the clarification or elicitation of relatively alternative answers to the question.

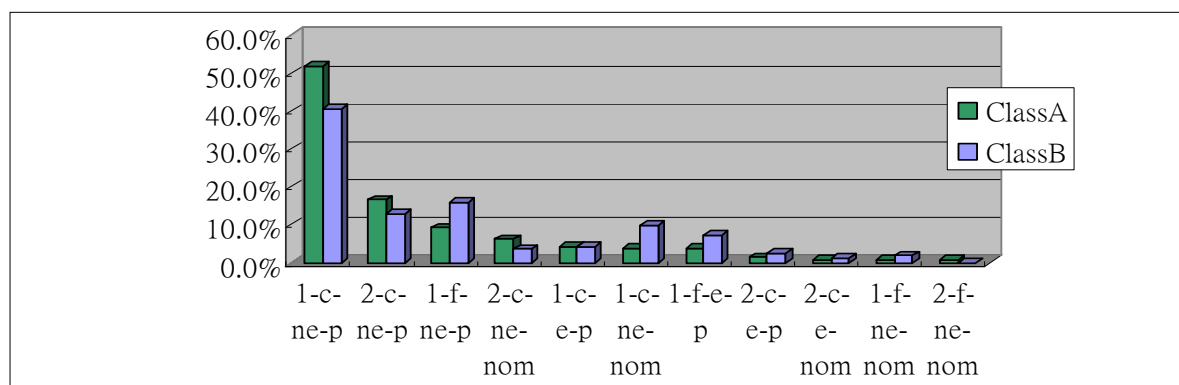
With eight levels in four variables—“1-DQ” (Display questions), “2-RQ” (Referential questions), “Content-focused” (C), “Form-focused” (F), “Non-Echoing” (NE), “Echoing” (E), “Predictable IRF” (P), “Negotiation of Meaning” (NOM), we can compose sixteen types of questions with different sequences of questioning characteristics, while, as a matter of fact, we only found eleven types of questions in the two classes.

RESULT AND DISCUSSION

Frequency of the Questions

FIGURE 1.

The Frequency of Every Type of the Questions in the Two Classes



423 questions were extracted from the two classes, 260 for Class A, and 163 for Class B. The standard for extracting questions in the two classes was that questions should be used for teaching, classroom management, and modification of teaching procedures.

From Figure 1., apparently, display questions were used more than referential questions. However, only with the dichotomy of classification may it be too vague to observe the characteristics of questions from other aspects. On the other hand, in order to investigate the aspects and modes of classroom interaction, the frequency of the feature sequences would provide more information for data analysis.

In terms of the sequences of question characteristics, we can find out Display questions-Content-focused-Non-Echoing-Predictable IRF (1-C-NE-P) were used most in both classes. However, the distributions of the second most used questions were different in two classes. The teacher in Class A used more Referential Questions-Content-focused-Non-Echoing-Predictable IRF (2-C-NE-P), while the teacher in Class B adopted more Display questions-Form-focused-Non-echoing-Predictable IRF (1-F-NE-P).

From the most used three types of questions, we understand that the both teachers in this study would prefer to use display questions to focus on the content without repeating students' answers to elicit the predictable IRF process, but the teacher in Class A also tended to use referential questions to focus on the content without repeating students' answers to elicit the predictable IRF process, while the teacher in Class B would like to use more display questions to focus on the form without repeating students' answers to elicit the predictable IRF process.

On the other hand, as to the process of the questions, "Predictable IRF" was often used more than "Negotiation of Meaning" in both display questions as well as referential questions. Nevertheless, we can still find that referential questions to focus on the content without repeating students' answers to elicit negotiation of meaning (2-C-NE-NOM) in Class A and display questions to focus on the content without repeating students' answers to elicit negotiation of meaning (1-C-NE-NOM) in Class B occupied a certain percentage in the use of questions. Another phenomenon is that both teachers seldom use the way of the questions to repeat what the students earlier said in their answers.

However, in this section, we can only observe the frequency of every type of questions,

but the classroom interaction triggered by the questions can not be scrutinized. Therefore, in the second section, we would like to investigate the classroom interaction triggered by every question.

Aspects and Modes of the Classroom Interaction triggered by Questions

In terms of this section, we would pay more attention to the analysis of classroom interaction. In order to make the analysis much clearer and more understandable, we design a set of criteria, adapted from Allwright (1984), including five aspects and four modes of classroom interaction.

With respect to the aspects of classroom interaction, there are “**Turn**”, “**Task**”, “**Topic**”, “**Tone**”, and “**Code**”. “**Turn**” refers to the contributions each participant, including students and teachers, would make. Here we would focus on the response triggered by every question, and, therefore, several subcategories can be introduced. There are “**Ss**” referring to the whole class response, “**Sp**” referring to the response produced by more than one student but less than half of the class, “**Sn**” referring to the response of a specific student, and “**T/S**” referring to the response of collaboration between the teacher and students, and “**Silence**” referring to no response and no turn-taking towards the question.

The second aspect is “**Task**”, which means that the participants demand others to make contributions, and here we focus on the degree of the control in the conversation among the teacher and students. “**T→S**” as the subcategory refers to the predominant control of the process of questions and answers, while “**S→T/S**” refers to the question produced by students towards teachers’ questions or students’ response to the question, requiring the teacher to make more contributions as well as consider the teaching process more. The third aspect is “**Topic**”, concerning the phenomenon of topic change or switch emerging in the conversation among the teacher and students.

As for the aspect “**Tone**”, it refers to the degree of voice quality for socio-emotional atmosphere and may be probably restricted to the sensible voice string. Thus, **High Tone (H.T.)**, implying louder volume and higher pitch, **Low Tone (L.T.)**, implying less loud volume and lower pitch, as well as **Laughter (Laugh.)**, showing the involvement of social interaction by students and teachers. The last aspect is “**Code**”, referring to the use of language by every participant in the questions and answers, including **Mandarin Chinese, Taiwanese, and English**.

In addition to the aspects of the classroom interaction, we would also like to use the modes of interaction to see the continuum of classroom interaction among the teacher and students. There are four modes of interaction, including “**Compliance**”—requirement for learners to simply do whatever is asked or expected by cooperative learners, “**Negotiation**”—any attempt to reach the consensus of the decision among students and the teacher, “**Navigation**”—the intention to overcome the obstacles presented in the lessons in the cooperation of the students and the teacher, and “**Direction**”—the mode for learners to totally follow teachers’ demands and to be directed by the teacher.

From Appendix 1., the result shows that the aspects of classroom interaction possessed by every type of questions extracted from the two classes. On the other hand, Table 1. presents the modes of classroom interaction. Generally speaking, most types of the questions possess Ss (whole class response) in “**Turn**” aspect, and “**silence**” does not reach higher percentages, and therefore we can infer from the data that both classes have a decent interaction in “**Turn**” aspect. Secondly, with respect to “**Task**”, most questions showed higher proportions of the teachers’ control, and the teacher in Class A controlled more task than the teacher in Class B did. Thirdly, “**Topic**” aspect to involve the change of the topics in Q&A procedure can be seen in the questions with “**Negotiation of Meaning**” feature. Fourthly, as to “**Tone**” aspect, “**High Tone**” phenomenon can be found in the questions with “**Non-Echoing**” and “**Predictable IRF**” characteristics, while “**Low Tone**” occurred more in the questions with

“Negotiation of Meaning” feature. Lastly, the “Code” aspect can be found that Mandarin Chinese and English achieved higher proportions, while Taiwanese has been exploited more in Class A., which can be explained in the questionnaires of the two teachers. The teacher in Class A mentioned the reason why she would use more Taiwanese was because of the male members in class, while the teacher in Class B thought that she would use Taiwanese only in particular situations to gain a socio-affective effect for classroom interaction. Here we can conclude that both teachers would prefer to choose Taiwanese as the teaching language for the purpose to pursue more class interaction or more involvement of students.

Table 1. *Modes of Interaction, and Functions in Questions in the Two Classes*

Modes	Compliance	Negotiation	Navigation	Direction	
Types of Questions		2-C-NE-P(B) 1-C-E-P(A) 2-C-E-P(B) 2-C-E-NOM(B)	1-C-NE-P(B)	1-C-NE-P(A)(B) 1-F-NE-P(A)(B) 1-C-E-P(A)(B) 1-F-E-P(A)(B) 2-C-E-NOM(A)(B) 2-F-NE-NOM(A)	2-C-NE-P(A)(B) 2-C-NE-NOM(A)(B) 1-C-NE-NOM(A)(B) 2-C-E-P(A)

On the other hand, Table1. shows the modes of classroom interaction, which are distributed unequally in four modes. In general, nearly all of the questions in the two classes belong to “Direction” mode, which means that the control of teacher was too predominant in the Taiwanese EFL context. However, “Negotiation” mode can be found in four types of questions, and most of them possess “Echoing” feature to make students come up with the clarification of the questions or more alternative answers to the questions.

The most used three types of questions in the two classes

In terms of the most used three types of questions (1-C-NE-P, 2-C-NE-P, and 1-F-NE-P), we find out some similar aspects among them. First of all, all of these three types are easy to arouse “Ss” (whole class response), and the frequency of this feature in the two classes reached more than 50%. However, still some of the “Sp” (less than half of the class response) and “Sn” (one specific student response) can be seen in these three types of questions. This perhaps can explain the reason why teachers love to exploit these three types of questions, since the whole class response would play a main role of the response, while, to some degree, individual responses or the responses from the minor groups of the class can also be elicited.

Additionally, we discovered that these three types induced more “High tone” of the voice quality, which may be related to the more proportions of the whole class response. Also, the “Laughter” phenomenon occurred particularly in 1-C-NE-P and 2-C-NE-P, while 1-F-NE-P possessed no “Laughter” phenomenon probably due to the form-focused factor and it can seldom interest students in socio-emotional aspects.

Table2. presents the examples of the most used question (1-C-NE-P) in both two classes. Although both of the classes use 1-C-NE-P to elicit students’ knowledge or provide chances to practice language use, Class A preferred to choose an indirect way to ask questions in English, while Class B tended to use a direct way in Mandarin Chinese. As was mentioned previously, the responses of 1-C-NE-P were normally whole class responses. However, in some situation as in the second example in Class A, students would use “Code-switching” to

involve the class by individual three students. This particular phenomenon may make the “predictable IRF” become less predictable by the teachers. Nevertheless, the mode of the interaction in 1-C-NE-P still belongs to “Direction” mode, which implies that the teacher controlled the class power most in this type of questions.

Table 2. *Examples, Mode of Interaction, and Functions of 1-C-NE-P in the Two Classes*

Example	1-C-NE-P(A)	1-C-NE-P(B)
Mode	Direction	Direction
Functions	Eliciting students' knowledge	I. Eliciting previous knowledge of words II. Providing chances to practice the use language.

Questions with the Most Aspects of Classroom Interaction

Another angle to analyze Appendix 1. is to find out which type of questions would possess the most aspects of the interaction, so that it can provide all the participants in class with a chance to produce more diversified phenomena and enhance learners' motivation of learning.

According to Appendix 1., we found that, in Class A, the referential questions to focus on the content without repeating what students said for the process of negotiation of meaning (2-C-NE-NOM) possess the most aspects of interaction, which can trigger whole class responses (Ss), the response of the class minority (Sp), and the most proportions of the individual response (Sn). The most important part is that the type of questions will trigger “Silence” response, which implies that diverse ways of interaction for discussion may be triggered by referential questions. Also, the “Task” aspect has T→S and S→T/S, which means the control of the task in class may be relatively various. Furthermore, the topic change or switching can also be triggered by this type of questions. “Tone” aspect can be

found in all of the three subcategories. Lastly, as for the code of language use, in Class A, we discovered that all of the languages—Mandarin Chinese, Taiwanese, and English—have been adopted for classroom interaction.

As can be seen in Table 3., the data show the similarities and differences among the use of 2-C-NE-NOM in the two classes. Both classes used 2-C-NE-NOM questions to elicit alternative or various answers from students' life experiences. However, the teacher in Class A seemed to put more emphasis on noticing the affective conditions of the students.

Further, in Class A, when the teacher asked "Do you wash your face today?", two individual students started to repeat the questions "Who", and the third student answered the question with a complete sentence structure "I wash my face." This phenomenon without whole class responses was similar in Class B when the teacher asked "Has any new album been released these days?" Originally, whole class remained silence, while, after one specific student mentioned a singer's name, the interaction become active and dynamic.

In short, we found the interesting part of interaction triggered by 2-C-NE-NOM. That is, whole class responses may be reduced, while immediately, owing to the individual response, classroom interaction becomes more involved by the teacher and students.

Another phenomenon of the similar types of questions is that the questions with "Referential questions" and "Negotiation of Meaning", as can be seen in Appendix 1., would trigger more individual student responses as well as the silence response in both classes, which represents the variation of interaction. Thus, it suggests that the participants need more time to consider the question, or teachers should improve the way to ask questions.

Table 3. *Examples, Mode of Interaction, and Functions of 2-C-NE-NOM in the Two Classes*

	2-C-NE-NOM(A)	2-C-NE-NOM(B)
Example	<p>I. T: Hey, 今天要錄影你那麼緊張喔? [speak to one student] S: 哪有。 T: 你今天心情不好齁? [speak to one student] S: 哪有。 T: 啊怎麼頭低低的?</p> <p>II. T: S13, <i>Lidualetsonshiang</i> (What did you do just now?) S16: What are you doing? [With great emotional expression] T: "What are you doing?" 誰說的? Ss: [Laughter.]</p> <p>III. T: Do you wash your face today? (1-C-NE-P) Ss: Yes. T: Who don't wash your face? S6: <u>Who?</u> S7: <u>Who?</u> S8: I wash my face. S9: <u>Who?</u></p>	<p>I. T: School 要改成? S11: Work. T: 也不一定是 school 跟 work, 還有什麼也用 for? Ss: ... (silence for two sec.) S7: 看電影。 T: <u>看電影. 看電影?</u> Ss: For T: <u>the movie. 有電影還有什麼? Show!</u> T: 你們有看過什麼表演嗎? 有沒有? 還有什麼地方可以用? 上? Ss: <u>課。</u> II. T: 來, 發行專輯? 這個單字, 我想你們應該不太會! Realise an album. T: <u>最近誰的專輯還不錯?</u> Ss: (Silence) S7: 鄧麗君。 T: <u>鄧麗君不是做古了嗎? (1-C-E-P)</u> S8: 她不會發專輯了! Ss: [Laughter] T: <u>還有沒有? Jay 的專輯? (1-C-NE-P)</u> S10: 阿妹的! S11: 周杰倫!</p>
Mode	Direction	Direction
Functions	<p>I. —Noticing students' affective anxiety.</p> <p>II. —Noticing students' attention to class.</p> <p>III. —Eliciting alternative answers.</p>	<p>I. —Eliciting students' knowledge of word use.</p> <p>II —Eliciting students' various answers to the open-ended questions.</p>

After analyzing the discourse data, then we would like to answer the research questions presented at the beginning of the study and discuss the reasons or comparison with the previous studies.

First of all, as for the question of which type of questions that Taiwanese non-native English teachers would use, the result showed that "Display questions" were used more than "Referential questions" in this study. This finding was rather similar to Shomoossi's (2004) study, in which display questions were used more by the teachers. One possible way to explain this phenomenon is that EFL teachers may focus more on the knowledge of the textbooks rather than the students' personal information of what they really want to know.

Besides, display questions to ask learners to display the already known knowledge would be much easier to arouse whole class responses. In the teachers' questionnaires, both teachers would like to draw students' attention to promote students' learning atmosphere. In this way of teaching philosophy, Display questions may be more effective to achieve this goal of teaching.

With respect to the research question of the frequency of every question and what characteristics would trigger which aspects and modes of interaction, we found that 1-C-NE-P, 2-C-NE-P, and 1-F-NE-P reach higher proportions than other types, while 1-F-NE-NOM and 2-F-NE-NOM achieved the lowest percentage among the distributions of every type. We can interpret the data in the way that the highest degree of questions with "Non-Echoing" and "Predictable IRF" can trigger the most whole class responses (Ss) as well as the predominant control of the task(T→S) by the teacher. Moreover, the "Tone" of the class interaction, such as "High Tone" and "Laughter", can be triggered through these features of interaction. As for the modes of interaction, even though "Direction" was occupied the first place among the questions, yet there were still some situations belonging more to "Negotiation" mode.

On the other hand, in regard to the type of questions possessing the most aspects of classroom interaction, we found that 2-C-NE-NOM, particularly in Class A, has the most various aspects of interaction. This finding provided the evidence that referential questions, especially with "Negotiation of Meaning" characteristic and "focus on the Content" without repeating what students said can trigger more various phenomena of classroom interaction. Tsui (2008) proposed that "variation is a necessary condition for discernment and therefore learning"(p.7). Her emphasis of variation was on students' awareness and noticing, and this dimension can be seen in our 2-C-NE-NOM. Although, referential questions with "Negotiation of Meaning" may also possibly trigger the "Silence response" and the Tone aspect was less significant than other types of display questions with "Predictable IRF" feature, yet this type of the questions does show more variations of classroom interaction.

CONCLUSION AND IMPLICATION

In conclusion, this study presented a detailed observation and analysis of the classroom interactional aspects and modes which can be triggered by the questioning characteristics. From the finding, we understand that display questions can be exploited more by the teacher because they would be easier to arouse more whole class responses, and, in some occasions, other types of responses as well as code-switching teaching strategies can draw students' attention more effectively.

However, referential questions, particularly with "NOM" feature triggered more variations of classroom interaction, which can give teachers chances to investigate different individual thinking as well as the peculiar opinion exchanges among students and the teacher.

On the other hand, the modes of interaction seemed to be the occupation of "Direction", and other three modes, "Compliance", "Navigation", and "Negotiation" can rarely be found in this study. Especially for "Compliance", this study focused on teachers' questioning, and "Compliance" is related more to student-centered teaching procedure, and therefore fewer question types observed in this study belong to "Compliance". However, "Negotiation" and "Navigation" should be encouraged to be adopted by the teacher, which can arouse more involvement of students to collaboratively change the teaching procedure and enhance more classroom interaction.

In the current Communicative Language Teaching (CLT), the questions to trigger classroom interaction have been encouraged. The significant implication of the current study aims to provide a detailed analysis of the natural current classroom discourse to discover the modification of the position in the continuum for the communicative purpose. With the finding, we can understand and verify the effectiveness of the present implementation of CLT

in the Taiwanese EFL context so as to offer some advice for its future development.

LIMITATIONS AND SUGGESTIONS FOR FUTURE STUDIES

Since this study was a small range of research, and the sessions extracted from the two classes may be insufficient, either. In the future studies, we suggested that more observations can be adopted in order to provide more evidence to support the hypothesis of this study or offer other insights towards the classroom interaction triggered by questioning.

Another limitation is that the design of the research in this study may be comparatively more complicated than the other previous studies. Also, the division of teaching settings in terms of questioning may be too vague and difficult to classify. The suggestion for further studies is to apply much clearer and more organized criteria to the setting and classification of the interactional aspects and modes triggered by questioning, so that this would make more contributions to the research on classroom interaction.

REFERENCES

- Allwright, R. L. (1984). The importance of interaction in classroom language learning. *Applied Linguistics*, 5(2), 156-171.
- Cullen, C. (1998). Teacher talk and the classroom context. *ELT Journal*, 52(3), 179-187.
- Cunningham, R. T. (1987). What kind of question is that? In W. Wilen (Ed.), *Questions, questioning techniques, and effective teaching* (pp. 67-94). Washington, D. C.: National Education Association.
- Ellis, K. (1993, February 13-16). *Teacher questioning behavior and student learning: what research says to teachers*. Paper presented at the 1993 Convention of The Western Dates Communication Association, Albuquerque, New Mexico. (ERIC Document Reproduction Service No. ED 359572)
- Gall, M. D. (1970). The use of questions in teaching. *Review of Educational Research*, 40(5), 707-721.
- Lynch, T. (1991). Questioning roles in the classroom. *ELT Journal*, 45(3), 201-210.
- Pica, T. (1987). Second-language acquisition, social interaction, and the classroom. *Applied Linguistics*, 8(1), 3-21.
- Roth, W-M. (1996). Teacher questioning in an open-inquiry learning environment: Interaction of context, content, and student response. *Journal of Research in Science Teaching*, 33(7), 709-736.
- Shomoossi, N. (2004). The effect of teachers' questioning behavior on EFL classroom interaction: A classroom research study. *The Reading Matrix*, 4(2), 96-104.
- Thornbury, S. (1996). Teachers research teacher talk. *ELT Journal*, 50(4), 279-289.
- Tsui, A. B-M. (2008, May 3-4). *Boundary-Crossing in Classroom Interaction Research*. Paper presented at the 2008 International Conference on English Instruction and Assessment, Chia-yi, Taiwan.

APPENDIX 1

Aspects of the Interaction in Questions in Two Classes

I.A Question	Turn					Task		Topic	Tone			Code		
	Ss	Sp	Sn	S/T	Sil.	T→S	S→T/S		H.T.	L.T.	Laugh	M.C	T	E
1-c-ne-p(A)	++	+	+			++			++		+	+	+	++
1-c-ne-p(B)	++	+	+	+		++	+		++	+	+	++		+
2-c-ne-p(A)	++					++			++		+	+		++

2-c-ne-p(B)	++	+	+			++	+		+	+		++		+
1-f-ne-p(A)	++					++			++			++		+
1-f-ne-p(B)	+	+	+	++		++						++	+	+
2-c-ne-nom(A)	+	+	++		+	++	+	+	++	+	+	+	+	++
2-c-ne-nom(B)	+	+	++	+	++	++		+		+	+	++		+
1-c-e-p(A)	+	+	++	+		++	+		+	+		+		+
1-c-e-p(B)	+	+	+		+	+	+			+		+		+
1-c-ne-nom(A)	++	++	++	+		++		+	+	++		++	+	+
1-c-ne-nom(B)	+	++	+		+	+	+	+	+	+		+		+
1-f-e-p(A)	++					++	+		++			++		+
1-f-e-p(B)	++	+		+		+	+		+			+		+
2-c-e-p(A)	++					++	+		+			++		+
2-c-e-p(B)		++	+			+	+			+	+	++		
2-c-e-nom(A)			+		+	++		+		+		++	+	+
2-c-e-nom(B)	+	+		+		+		+			+	++		+
1-f-ne-nom(A)			+			++						++		+
1-f-ne-nom(B)		++			+	+				+		++		
2-f-ne-nom(A)					+	++						++		
2-f-ne-nom(B)														

Note. ++ = Percentage of frequency more than 50% + = Percentage of frequency less than 50%
I.A.= Interaction Aspect

APPENDIX 2

Examples, Mode of Interaction, and Functions of Every Question in Two Classes

	Example	Mode	Functions
1-C-NE-P(A)	<p>I. T:那它說一條麵包多少錢? S: ㄟ 條 20 塊。 T:所以怎麼樣? S: [No particular answer] T:就怎樣? Ss:很便宜。</p> <p>II. T: OK, what is week? Ss:週。 T:S1, <i>Shiamingshi</i> “week.” (What is week?) S2:週。 S3:週。 S1:週。</p> <p>III. T: OK, this one? Ss:Baby..Baby. T:OK, are you a babies?</p>	<p>I.Direction</p> <p>II.Direction</p> <p>III.Direction</p> <p>IV.Direction</p>	<p>I. — Eliciting students’ knowledge in the textbooks.</p> <p>II. — Eliciting students’ knowledge from L2</p> <p>III. — Eliciting students’ knowledge of the vocabulary</p> <p>IV. — Eliciting students’</p>

	<p>Ss:No. T:Are you girls? Ss:No. T:Are you boys? Ss:No..Yes..No. T:Are you men? Ss:No. [Laughter] IV. T:OK, this one?(Showing a pic.) S2:餓死我。 S1:“Us.” S3: 餓死。</p>		<p>knowledge of what students saw.</p>
1-C-NE-P(B)	<p>I. T:Do you have a computer ... 這裡是自己的 computer 就要用? T/Ss: Of your own. II. T:In---the---class 在? Ss: 課堂/教室。 T:在課堂當中叫 in class,那 in the class? Ss: ㄟ教室。 III. T:“Don’t be late” 不要遲到。有遲到就會有? S6: ㄟ早退! Ss: [Laughter] T:不是拉哈。對啦早到。 IV. T:OK,我一天洗兩次澡? Ss:I take a both twice a day. T:OK,三次,我一年洗三次澡? Ss: [Laughter] T:那個 S3,一年洗三次? Ss: (Full of laughter for13 seconds) T: OK, Evan Ss: Haha,Evan takes a bath T: ㄟthree times a year. Ss: [Laughter again] V. T:這裡講的是魚還是魚肉阿? S3:魚。 VI.T:來第 5 題,how much are these spoons? S5:老師可以用 the 嗎? T:How much are the spoons?</p>	<p>I.Direction II.Direction III.Direction IV.Direction V.Direction VI.Navigation</p>	<p>I./II. — Inducing students to give answers. III.— Eliciting previous knowledge of words IV. — Providing chances to practice the language use. V.— Eliciting the correct answer. VI.— Dealing with the obstacles in the textbooks</p>

	可以！		
2-C-NE-P(A)	<p>I. T:就怎樣？ Ss:很便宜。 T:OK，這樣有沒有問題？ Ss:沒有。</p> <p>II.T:Do you brush your teeth today? Ss:Yes.</p> <p>III.T:Do you like social studies? Ss:Yes. ↳No.</p> <p>IV. T:Is your social studies teacher handsome? Ss:No. T:Is your English teacher beautiful? Ss:No. [Laughter] T:<i>Beaiga.</i>(I don't wanna teach) Ss: [Laughter]</p>	<p>I.Direction</p> <p>II.Direction</p> <p>III.Direction</p> <p>IV.Direction</p>	<p>I. — Checking students' understanding</p> <p>II.— Eliciting alternative answers</p> <p>III. — Eliciting alternative answers</p> <p>IV.— Eliciting alternative opinions</p>
2-C-NE-P(B)	<p>I. T:OK,都準備好了對不對？還有沒有沒寫完的？ S1:老師我沒帶！ T:沒有帶還是沒有寫？ S1:沒有帶 T:阿有沒有寫？ S1:有。 T:好明天拿來給我檢查。</p> <p>II. T:第一題，我題目一起唸還是唸答案就好？ Ss:唸答案！ T:我比較喜歡一起唸耶！ Ss:答案一直唸一直唸就是一直唸下來就好。 T:ㄟ,你們聽得懂比較重要。來從第一題喔</p> <p>III. T:They are not really cheap...ㄟ我覺得你們跟我一起唸比較好不好？你們告訴我答案比較快！ S6:老師可不可以用“so”？ T:他說“真的”不是“如此”所以我們用 really。</p>	<p>I.Direction</p> <p>II.Negotiation</p> <p>III.Direction</p>	<p>I. — Initiating the class</p> <p>II.— Involving students to change the teaching procedure</p> <p>III. — Involving students to change the teaching procedure</p>

1-F-NE-P(A)	<p>I. T:OK,this one? Ss:Them. T:OK, them. T:什麼格?主格、所有格、受格? Ss:受格。</p> <p>II. T: OK,this one? Ss : [pause]Mouth. T:OK,TH 怎樣? Ss: [No particular answer.] T:舌頭要放在? (Pause) 這有聲無聲我問你? Ss:無聲。 T: OK, very smart.</p>	<p>I.Direction</p> <p>II.Direction</p>	<p>I. — Eliciting grammatical knowledge</p> <p>II. — Eliciting the pronunciation</p>
1-F-NE-P(B)	<p>I. T:How many 老師講過嚕, how many 後面接單數還複數? S4:複數。</p> <p>II. T:來下一個單字叫什麼?這一課每個都要會喔,來第一叫? Ss:First. Second. Third. T: ↪第二? ↪第三? T:這些英文叫做? Ss:...數...基數...</p> <p>III. T:OK,還有一個擺的位置很討厭,是我們的“三太子”“<i>Samtatsu</i>” “Sometimes”對不對?它放在句中 Ss: ↪句尾 T: ↪句首。</p>	<p>I.Direction</p> <p>II.Direction</p> <p>III.Direction</p>	<p>I. — Eliciting grammatical knowledge</p> <p>II. — Eliciting the following grammatical knowledge</p> <p>III. — Making students aware of the position of words.</p>
2-C-NE-NOM(A)	<p>I. T:Hey,今天要錄影你那麼緊張喔? [speak to one student] S:哪有。 T:你今天心情不好齁? [speak to one student] S:哪有。 T:啊怎麼頭低低的?</p> <p>II. T:S13,<i>Lidualetsonshiang</i> (What did you do just now?) S16:What are you doing?[With great emotional expression] T: “What are you doing?”誰</p>	<p>I.Direction</p> <p>II.Direction</p> <p>III.Direction</p>	<p>I. — Noticing students' affective anxiety</p> <p>II. — Noticing students' attention to the class</p>

	<p>說的？Ss:[Laughter.]</p> <p>III. T: Do you wash your face today?(1-C-NE-P)</p> <p>Ss: Yes.</p> <p>T: Who don't wash your face?</p> <p>S6:Who?</p> <p>S7: Who?</p> <p>S8:I wash my face.</p> <p>S9: Who?</p> <p>IV. T:OK, what is science in Chinese?(1-C-NE-P)</p> <p>Ss:自然。</p> <p>T:Who likes science?</p> <p>S11:I like</p> <p>Ss: [Laughter]</p> <p>V. T:OK, turn to page 92,你看到了什麼？</p> <p>S1:看到了英文。</p> <p>Ss: [no particular answer but many students said some words]</p>		<p>III. — Eliciting alternative answers</p> <p>IV. — Eliciting various students' opinions</p> <p>V. — Encourage students' different consciousness of the textbook.</p>
<p>2-C-NE-NOM(B)</p>	<p>I. T:School 要改成?</p> <p>S11:Work.</p> <p>T:也不一定是 school 跟 work, 還有什麼也用 for?</p> <p>Ss:...(silence for two sec.)</p> <p>S7:看電影。</p> <p>T: 看電影. 看電影?</p> <p>Ss:For</p> <p>T: 看 the movie.有電影還有什麼?Show!</p> <p>T:你們有看過什麼表演嗎?有沒有?還有什麼地方可以用?上?</p> <p>Ss: 課。</p> <p>II.T:來,發行專輯?這個單字,我想你們應該不太會!</p> <p>Realise an album.</p> <p>T:最近誰的專輯還不錯?</p> <p>Ss:.(Silence)</p> <p>S7:鄧麗君。</p> <p>T: 鄧麗君不是做古了嗎?(1-C-E-P)</p> <p>S8:她不會發專輯了!</p> <p>Ss:[Laughter]</p>	<p>I.Direction</p> <p>II.Direction</p>	<p>I. — Eliciting students' knowledge of word use.</p> <p>II — Eliciting students' various answers to the open-ended questions</p>

	<p>T: 還有沒有? Jay 的專輯?(1-C-NE-P)</p> <p>S10: 阿妹的!</p> <p>S11: 周杰倫!</p>		
1-C-E-P(A)	<p>I. T: OK, this one? Ss: Sick. T: OK, what is "sick"? Ss: 生病的。</p> <p>II. T: 怎麼了, 好第一題 Who watches S22: ↳ TV on Wednesday night. T: Wednesday 是哪一天? S22: 星期三。</p> <p>III. T: In Chinese? S4: "壁虎" T: ↳ "壁虎" 喔? Ss: 牛肉。</p>	<p>I.Direction</p> <p>II.Direction</p> <p>III.Negotiation</p>	<p>I. — Checking students' understanding</p> <p>II. — Checking students' understanding</p> <p>III. — Checking students' understanding</p>
1-C-E-P(B)	<p>I. T: ↗Mike, 請問他們在幹嘛? Ss: 點名。 T: 點名怎麼說? Ss: ...hum?? [Silence] T: Have---a---roll—call.</p> <p>II. T: 在課堂當中叫 in class, 那 in the class? Ss: ↳ 教室。 T: 教室嗎? S6: 在班級。 T: ↳ 對對對, 在這個班級。</p> <p>III. T: 什麼要劃起來? Ss: "Make excuses" T: 什麼叫 Make excuses? Ss: 找藉口。</p>	<p>I.Direction</p> <p>II.Direction</p> <p>III. Direction</p>	<p>I. — Checking students' understanding</p> <p>II. — Checking students' understanding</p> <p>III. — Checking students' understanding</p>
1-C-NE-NOM(A)	<p>I. T: OK, how much bread do we need? 那你告訴我那你先告訴我這句話是什麼意思? Ss: [wait for 2 sec.] 我們需要多少麵包?</p> <p>II. T: 那 "Homesick" 是什麼? S5: 想家。 S6: 在家生病。 T: 在家生病 shishiang? (What is "sick</p>	<p>I.Direction</p> <p>II.Direction</p>	<p>I. — Eliciting students' interpretation</p> <p>II. — Eliciting students' prediction of new information</p> <p>III. — Eliciting</p>

	<p>in home” ?)</p> <p>T:今天不小心我們 S12 不小心考上了台大法律系，要去台北讀書…</p> <p>III. T:OK,補充,如果要說一雙筷子怎麼說?</p> <p>Ss:A</p> <p>T: ↳A pair of Chopstick.</p> <p>Ss: ↳of Chopstick.</p> <p>IV. T:OK,然後他買了很多東西? 你看他買了什麼東西? 買了什麼?</p> <p>S18:Orange.</p> <p>S20: ↳Pizza.</p> <p>S21: ↳Tomato.</p> <p>Ss: ↳Apple juice.</p> <p>V. T:OK, S16,what are you doing?</p> <p>S16: What are you doing? 你在幹麻?</p> <p>T:↘, S16, 你英文進步勒。</p> <p>S16:Maionegonla.(Don't say that.)[Become blushed]</p> <p>VI. T:好那 “have”也是?</p> <p>Ss:擁有。</p> <p>T:OK,那有什麼不一樣?</p> <p>S1:就是不一樣!</p> <p>S10: ↳一個是有一個是擁有。</p>	<p>III. Direction</p> <p>IV. Direction</p> <p>V. Direction</p> <p>VI. Direction</p>	<p>students' prediction of new information</p> <p>IV. — Eliciting students' prediction of new information</p> <p>V. — Noticing students' attention to the class</p> <p>VI. — Eliciting alternative answers</p>
1-C-NE-NOM(B)	<p>I. T:↗Mike,請問他們在幹嘛?</p> <p>Ss:點名。</p> <p>T:點名怎麼說?</p> <p>Ss:...hum??[Silence]</p> <p>T:Have---a---roll—call.</p> <p>II. T:還有哪個瑕疵?</p> <p>Ss:Excu..excu</p> <p>T: ↳↘ excu-ses ↘是複數對不對? 請問發生什麼事情了? Who is he?(1-C-NE-P)</p> <p>Ss:Mike.</p> <p>T:OK, Mike. Why?</p> <p>Ss:...(pause for a whole)</p> <p>S4:不知道。</p> <p>S5: ↳不知道。</p> <p>T:對,yeah, you don't know.不</p>	<p>I.Direction</p> <p>II.Direction</p>	<p>I. — Eliciting students' interpretation of the text.</p> <p>II. — Eliciting students' opinions towards peer's performance/ Eliciting students' interpretation of the text./ Eliciting different opinions of</p>

	<p>知道。</p> <p>T:In class 跟 in the class 有什麼不同?</p> <p>Ss: in class 是在教室/是在課堂 ...in the class 是...(Discussion)</p> <p>T:In---the---class 在?(1-C-NE-P)</p> <p>Ss:課堂/教室。</p>		word use.
1-F-E-P(A)	<p>I. T:OK, them.</p> <p>T:什麼格?主格、所有格、受格?</p> <p>Ss:受格。</p> <p>T: 誰的受格?</p> <p>Ss: They...I..</p> <p>T:ʎhun?</p> <p>Ss:They.</p> <p>II. T:那兩杯怎麼說?</p> <p>Ss:Two glasses of water.</p> <p>T:OK,glasses 要加?</p> <p>Ss:E-S.</p>	<p>I.Direction</p> <p>II.Direction</p>	<p>I. — Checking students' understanding</p> <p>II. — Checking students' understanding</p>
1-F-E-P(B)	<p>I. T: 誰這一課的主要文法?</p> <p>Ss:頻率副詞。</p> <p>T: 頻率副詞擺的位置? 等一下,十四(some of the students stood up, and the teacher found a student)來頻率副詞擺的位置?</p> <p>Ss:BE 後助後動前。</p> <p>II. T:OK,簡答需要使用頻率副詞時?(1-F-NE-P)要擺的位子?</p> <p>Ss:Be 動</p> <p>T: 主詞跟? Be 動詞或助動詞的?</p> <p>Ss: 中間。</p> <p>T:主詞跟什麼? Be 動詞、助動詞的? 中間。</p> <p>Ss: 中間。</p>	<p>I.Direction</p> <p>II.Direction</p>	<p>I. — Checking students' understanding</p> <p>II. — Checking students' understanding</p>

2-C-E-P(A)	<p>I. T:所以你用 when 問的時候一定有什麼? Ss:時間。 T:有時間。可以嗎? T:OK,跟我唸。</p> <p>II.T:OK,有人說阿,他的單字阿他會比較害羞,因為他不敢唸怕唸錯,可是你要唸出來我才知道你會不會唸,對不對? Ss:對[less louder and lower tone] T:對不對? Ss:對 - 阿! [louder and higher tone]</p>	<p>I.Direction</p> <p>II.Direction</p>	<p>I. — Checking students' understanding towards teacher's explanation</p> <p>II. — Checking students' understanding towards the teacher's explanation</p>
2-C-E-P(B)	<p>I.[In preparation of the drama activity] T:你們已經練好了嗎? Ss:好啦好了。 T:已經好了的那組請舉手? Ss:(In a discussion)(Laughter) 兩個人而已 T:只有你們兩個好啦嗎?你們是同一組的嗎?好,那就你們這組,來!這個改編版的,好了嗎? S13:老師這個講桌要不要搬? T:講桌要搬嗎?好好好! S14:老師那個還有一組? T: 沒有我們只有一組!好不好,因為其他人不在,等下一節。</p>	<p>I.Negotiation</p>	<p>I. — Checking students' preparation for activities or the classroom settings</p>
2-C-E-NOM(A)	<p>I. T:那"Homesick"是什麼? S5:想家。 S6:在家生病。 T: 在 家 生 病 <i>shishiang</i>? (What is "sick at home"?) T:今天不小心我們 S12 不小心考上了台大法律系,要去台北讀書...</p>	<p>I.Direction</p>	<p>I. — Checking students' understanding or opinions of students' own interpretation</p>

2-C-E-NOM(B)	<p>I. T:Late for English class. 有沒有, 你們有沒有覺得每次上補習班的課都不想去上課?(2-C-NE-P) S8:沒有我覺得補習班很好玩耶! T:阿很好玩!哪一定是上英文課的對不對?(2-C-NE-P) S9:沒有數學。 T:數學也很好玩喔? Ss: (Laughter).</p> <p>II.T :好啦, 下個單字, 我們好像學過了對不對?(1-C-NE-P) Ss:自由的! T: 我們在哪學過?(1-C-NE-NOM) Ss :那個(Laughter) T: ↳Take me Ss: ↳Heart T: ↳and set me“Free”. Ss: ↳me “Free”. S10: 還有那個 “自由女神”! T: ^,自由女神怎麼說? Ss:Free....free.. T:跟這個單字完全沒有關係!</p>	<p>I.Negotiation</p> <p>II.Direction</p>	<p>I. — Checking students’ opinions</p> <p>II — Checking students’ knowledge of vocabulary.</p>
1-F-NE-NOM(A)	<p>I. T:我們之前用 how many 耶? 那為什麼這裡…因為 bread 怎樣? S: ↳ 不可數。</p>	<p>I.Direction</p>	<p>I. — Encouraging students to self-construct grammatical knowledge</p>
1-F-NE-NOM(B)	<p>I. T:來這裡出現了一個我們常用的文法是什麼? Ss: (Silence) T: ↳這一課的主要文法? Ss:頻率副詞。</p> <p>II. T:來有關頻率副詞的其他表達方式, 來其他表達方式。什麼樣的表達方式? 頻率副詞除了這些以外呢</p>	<p>I.Direction</p> <p>II.Direction</p>	<p>I. — Inducing students to recall the learned grammatical knowledge</p> <p>II. — Inducing students to</p>

	另外一個就是…。不是單純的頻率副詞。而是使用片語，就是說一個禮拜一次？(1-C-NE-P)怎麼樣？(1-C-NE-P) S1:One Ss: One		recall the learned grammatical knowledge
2-F-NE-NOM	I. T:那可數不可數你們可不可以各舉一個例子？不可數呢？ Ss: (No particular answer)	II.Direction	I. — Giving students chances to come up with examples of the grammatical knowledge.
2-F-NE-NOM			

APPENDIX 3

課堂教師問卷

- (1)請問您在課堂上會使用英語、國語或台語問學生問題嗎？大概在何時會用何種語言問呢？
- (2)您時常問學生課本上的問題嗎？還是課本主題延伸至學生生活的問題呢？其原因為何？
- (3)在您的課堂中，學生會回問您的問題嗎？可否舉一個簡單的例子呢？您通常都如何回應學生發問的問題呢？
- (4)您常在問學生問題後，學生並沒有真正回答您要的答案嗎？您都會如何處理呢？
- (5)您較常問全班問題或是問班上其中一位同學問題呢？其原因為何呢？
- (6)您覺得問問題能增加班上的互動嗎？請簡述其原因。

—問卷到此結束，感謝您的參與—

An Integrated Freshmen English Instruction through Technology

Pei-ling Wang (王佩玲)

National Kaohsiung University of Applied Sciences

peiling@cc.kuas.edu.tw

The purposes of this study are: (1) to demonstrate a way of using technology to incorporate listening, speaking, reading and writing components into Freshmen English instruction; (2) to investigate students' perceptions of integrated instruction and classroom activities; (3) to explore how much effort students make to learn English after class; (4) to examine the effectiveness of the holistic teaching method by comparing students' pre- and post-test scores; and (5) to determine the difference between high ability and low ability students' learning achievements. A total of 212 freshmen (96 low level and 108 high level students) were quasi-randomly selected as the participants of the survey study. Instruments included a questionnaire and a standardized English exam. The teacher researcher used leveled textbooks and Readers for different levels of classes. Students needed to regularly surf the e-learning website provided by the textbook publisher (located at http://203.64.90.116/z_liveorg7/), and the supplemental materials on the University Learning website (located at <http://140.127.113.194>). The results of the study indicate 91% of the students evaluated the instruction positively. Watching movies was students' favorite activity, and writing a reflection report of a novel was the least popular. Surprisingly, over 60.8% of students spent less than one hour per week on the English Learning website. While the significant difference in the gain scores of the t-test reveals that integrated instruction is effective ($p=0.00$), low level students had much more improvement than high level students, whose scores even showed slight regression.

INTRODUCTION

According to the study investigating Tunghai University students' English language ability changes from 1998 to 2005 (Chen, 2005), there has been a decline in the level of university students' English grammar and reading ability. Although students' overall listening ability has improved over the past seven years, their intermediate listening skills fell again after 2003. From the perspective of inferential statistics, this case study's findings cannot be applied to all university students in Taiwan; however, they do suggest that Taiwanese university students' English ability is getting worse.

Among the many studies exploring the reasons for Taiwanese university students' poor English ability, some researchers (e.g. Liou, 2008) claimed that students spent too little time reading, they lacked opportunities for using English outside class, and their linguistic knowledge (phonemic rules, syntactic rules, and vocabulary) was limited. Others (e.g. Sheu, & Wang, 2007; Shu & Wang, 2006) indicated that a majority of students were lacking in motivation to study and had low expectations of success.

Since numerous studies have shown that the degree of students' motivation significantly influenced their effort to study and further facilitated their learning achievements, it is vital that teachers apply interesting teaching methods and materials to motivate their students. Unfortunately, the mismatch between teachers' focus of instruction and students' preferences has already been found in many classrooms. This might also account for students' common low learning interests (Nunan, 1995).

Traditionally, Taiwanese university teachers adopt a discrete-skills approach in their teaching. They offer courses such as Freshmen Reading, and Sophomore Listening and Speaking. Many teachers believe that students need to have adequate linguistic knowledge

and knowledge of text structures before they acquire listening and speaking competence. Therefore, they often emphasize analyzing the roots of words and memorizing vocabulary items instead of learning for authentic communication. Some teachers even argued that it is instructionally impossible to concentrate on more than one skill at a time (Oxford, 2001). However, does segregated-skill instruction effectively enhance students' English overall English performance? Reports of the Educational Testing Service (2007) revealed that Taiwanese university students' English language skills in terms of reading, writing, speaking and listening are all disappointing. Obviously, teachers need to consider other methodologies such as the whole language approach in their teaching.

Proponents of the whole language approach (e.g. Doake, 1994; Dudley-Marling, 1995; Freeman & Freeman, 1994; Newman & Church, 1990) stress that human language learning starts from the whole to the parts. Teachers need to provide natural reading, writing, speaking and listening situations where students can practice discrete language components such as spelling, vocabulary, and pronunciation. Moreover, this approach encourages students to cooperate with others, to create, and to take risks during task-based instruction.

While the whole language approach has been very popular in western countries such as New Zealand and Canada, Taiwanese teachers need to know how to actually implement integrated-skill instruction in a large class (50 or more students) within limited class time. Certainly, it is very challenging for teachers to design classroom activities which involve every student practicing multiple language skills at one time. Fortunately, with the aid of multimedia, it is possible with computer-assisted instruction to put the theory of the whole language approach into practice in the classroom.

The main purpose of this study is to demonstrate a way of using computer technology to integrate listening, speaking, reading and writing components into Freshmen English instruction. Research questions include (1) What are students' perceptions of the integrated instruction and classroom activities? (2) How much effort do students make to learn English after class? (3) What is the effect of the integrated instruction on students' learning achievements? and (4) Does significant difference exist between high ability and low ability students' learning achievements?

In order to help readers understand the theoretical backgrounds of the study, the researcher reviewed studies on the whole language approach and computer-assisted language learning as follows.

LITERATURE REVIEW

Studies on Whole Language Approach

According to Goodman (1986), language ability develops naturally as a consequence of experiences with language. He argued that the teaching of language as isolated skills was inappropriate and not likely to succeed, because the focus of segregated-skill instruction is not on learning language for authentic communication. Goodman's theory was the theoretical basis for the whole language approach, and continued to receive support from other scholars.

The whole language theorists (e.g. Norris & Hoffman, 1993; Reutzel & Hollingsworth, 1988) urged that students should be given the opportunity to use, to explore, and to make sense of all language arts (listening, speaking, reading, and writing) in contextually meaningful and cooperative activities, since language learning is a social process and besides, oral and written languages are all inter-related. Others (e.g. Anderson, 1984) stressed that if language is taught as isolated skills, it will be hard for the human brain to memorize those pieces of information for a long time. Moreover, some scholars (e.g. Reutzel & Cooter, 1992) believed that when teachers allocate most classroom time to explaining discrete language skills, students won't have much time to experiment with the target language. They asserted that language learning involves relating new information to prior knowledge, so

learning tasks should be student-centered and process-oriented.

Several studies (e.g. Rosberg, 1995; Moosavi, 2007) have shown the effectiveness of the integrated teaching method in western countries, and the subjects of those studies included elementary students, high school students, and university students. In Rosberg's study, elementary schools students were asked to read from a variety of genres and authors, predict the plots, record their reflections in journals, compare their own and team members' writing styles with the authors' styles, adapt the stories and then share their work in the form of role play. Rosberg believed that this literature-based teaching "promoted metacognitive awareness because there were more opportunities for reflection and cognitive linkages" (p.5). Moreover, this instruction also helped develop imagination, increase vocabulary, experience the process of writing, and foster communicative competence. Most importantly, children became mature physically, intellectually, socially, and emotionally when they participated in group activities.

In comparison, although some studies on integrated instruction have been performed in Taiwan, most research was conducted with elementary and high school students (e.g. Cheng, 1998; Chern, 2000; Tu, 2004; Wang, 2003). At present, research focusing on integrated instruction at universities is still scarce. Among these studies, almost all subjects were not from technical universities but from general universities such as Providence University (Chiang, 2005), Taiwan Normal University (Chu, 2005), Cheng Kung University (Lai, 2005), and Tunghai University (Sim, 2005). So far, only one study has been conducted in a technical university (Eyerman, 2005). Generally, these studies demonstrated that students benefited from integrated instruction; however, some results were not in line with studies in western countries. For example, in Lin's study (2007), integrated instruction was implemented to 2,707 freshmen for two semesters. Students improved in all measured areas: grammar, reading, and listening. The improvement was especially significant in listening ability. Moreover, low ability students improved most while high ability students improved least. In contrast, in EI-Koumy and Abdels' (2000) study, 96 Egyptian university students were divided into a skills-based group and a whole language group for 15 weeks of listening instruction. The result indicated that there was no significant difference in listening ability among low-level students in two groups, while high-level students in the whole language group had great improvement. EI-Koumy and Abdel concluded that whole language approach and the discrete-skills approach could complement one another. In other words, the whole language approach to teaching listening cannot work without teaching basic skills such as recognition of individual sounds, reduced forms, stress and intonation patterns.

Although the whole language approach has gained popularity since 1980, it has also received some criticism. Opponents argue that this approach focuses on meanings and neglects accuracy; however, they believe that accuracy is an essential element in language development (Eldredge, 1991; Goldenberg, 1991). In addition, some scholars are concerned that students might not be able to regulate and direct what they have learned (Freeman & Freeman, 1992). Moreover, some claimed that this approach is very time-consuming (Danehower, 1993). Finally, others assert that it is difficult to implement a whole language approach due to factors such as teachers' inadequate professional knowledge, increased workloads, and the lack of teaching resources and equipment in schools (Sanacore, 1995).

Studies on Computer-Assisted Language Learning

The topic of computer-assisted language learning (CALL) has received considerable attention since the 1960's. According to Warschauer (1996), CALL has gone through three phases: Behaviorist CALL (1960's- 1970's), communicative CALL (the late 1970's-1980's), and interactive CALL (1990's-present). The design of Behaviorist CALL, which was based on behaviorist theories, provided students with abundant drills and practice. Students could repeatedly practice at their own pace without worrying that computers may get bored or

angry with them. In this stage, the computer functioned as a tutor, presenting exercises and non-judgmental feedback. Next, the design of communicative CALL, based on the communicative approach, put emphasis on using the target language rather than on analysis of the language. It provided communication exercises in a non-drill format, such as paced reading, writing, discussions, and language games. In this stage, the computer was used as stimulus and a tool. Finally, the design of integrative CALL, based on constructivism, whole language theory and socio-cultural theory, provided authentic discourse through the use of multimedia and Internet to help learners fit into the real world. In this stage, the designs of text, graphics, sound, animation and video in multimedia programs are even more attractive and interesting. Also, students have greater control over their own learning including the content, sequence and pace of learning. With modern technology, students can practice not only reading and listening skills but also writing and speaking. For example, using email, synchronous chat or video conferencing could allow students to interact with real people and obtain immediate feedback. Automatic speech recognition and visual sound waves technology are also very helpful (Moras, 2001).

According to Liu and Chen (2008), the application of technology in language classrooms has been utilized in different ways; for example, it could be a tool that facilitates language acquisition or provides remedial instruction to learners with limited language proficiency. Many studies on CALL have shown that the application of computer technology has a positive effect on students' learning interest (Hung, 2001), motivation (Chang & Lehman, 2002), self-concept (Stepp-Greany, 2002), self-reflection (Chen, 2004), confidence (Chu, 2004), critical thinking (Chiu, 2005), cultural awareness (Cheng, 2004) academic outcomes in learning listening skills (Tsai, 2003), speaking (Chen, 2004; Chiu, 2005), reading (Chen & Tseng, 2006; Ho, 2003), and writing (Chan, 2003; Lu, 2002).

While the benefits of CALL have been widely accepted by educators, there are still some criticisms of CALL instruction. The first is the limitation of the technology. For example, the computer cannot effectively evaluate students' speaking practice and the machine's pronunciation is not human-like. Moreover, the cost and availability of the Internet are another problem in many developing countries (Ehsani & Eva, 1998). The second is the stability and quality of CALL software. Teachers may rely on commercial sources; however, these sources may not be pedagogically sound. The third is technical difficulties which would interfere with the learning process. Lack of computer expertise might make both teachers and students reluctant to use this technology (Thelmadatter, 2007).

Furthermore, some research even showed that too many multimedia features interfere with students' learning (e.g. Mayer, Heiser, & Lonn, 2001), and others argued that while CALL might supplement face-to-face instruction, it should not replace it. For example, Huang's study (2004) showed that 100% CALL implementation was the least effective instruction model in teaching the English alphabet to second graders, compared with three other models: 100% lectures, 5/7 lectures supplemented by 2/7 CALL, and 5/7 CALL supplemented by 2/7 lectures. Interestingly, the group taught using 5/7 lectures supplemented by 2/7 CALL improved most.

Liu and Chen (2008) reviewed the literature on CALL in Taiwan from 2000 to 2006. They found that most of the studies focused on reading and writing skills, and only a few studies addressed listening and speaking skills. In addition, they complained that "many studies lacked a thorough description of the procedures, including information on the subjects, materials, technologies, treatments, tests, and statistical analyses used in the particular investigation" (p.32).

In an attempt to clarify the inconclusive results of previous studies on the whole language approach and CALL, this study provided information of various multimedia features, and also measured the outcomes of integrated English instruction through

technology. The results of the study might shed some light on the procedures for integrating listening, speaking, reading, and writing components into English teaching and learning. They could also help instructional designers develop effective CALL programs.

METHODOLOGY

Subjects

A total of 212 non-English majored freshmen at Kaohsiung University of Applied Sciences (KUAS) were quasi-randomly selected from the researcher's students as the subjects of the study. These participants came from four intact Freshmen English ability-grouping classes, including two low-level classes (96 students) and two high-level ones (108 students) based on their English scores in the 2007 Entrance Examination for the Technical University.

Instruments

Instruments included (1) a self-designed questionnaire consisting of 4 students' background information questions, 10 four-point Likert Scaled questions, 3 multiple-choice questions, and 1 open-ended question (See Appendix), and (2) the 2007 English Entrance Examination for Technical University, which was used as the pre/post test of the study. The test consists of vocabulary (30%), dialogues (20%), grammar (30%), and reading comprehension (20%). The internal reliability for the Likert Scaled questions was estimated with Cronbach's alpha. The result showed that the questionnaire was reliable ($\alpha=0.80$).

Procedures of Data Collection

The researcher collected students' English scores on the Technical University Entrance Examination from the Office of Academic Affairs at KUAS, so she could record students' performance prior to the Freshmen English instruction. Next, she conducted English instruction with the aid of multimedia technology to the subjects for one semester. Furthermore, the subjects took the post-test and answered the questionnaire on the last two weeks of the semester. The procedures for implementing instruction are described as follows.

English Instruction through Technology

Materials

The teacher-researcher used two textbooks in the class. For high-level class, she used OFF WE GO 3 (Live ABC Company) and THE FIRM (Penguin Readers, level 5). For low-level classes, she used OFF WE GO 1 and THE CLIENT (Penguin Readers, level 4). The components of OFF WE GO include a textbook, an interactive CD-ROM/MP3, and an interactive web site (http://203.64.90.116/z_liveorg7/). Penguin Readers contain a storybook and a tape.

On-line Learning and Resources

There were three sources of on-line learning in this course. The first one was the KUAS Learning Platform, an interactive website which was constructed by the university. The researcher regularly posted course information and supplemental materials on this platform (located at <http://140.127.113.194>) (See Figure 1). Prior to the class, students were asked to download the list of vocabulary items with K.K. phonetic symbols and Chinese/English definition, an explanation of grammatical rules, related articles, and guided discussion questions for the Penguin Reader (See Figure 2). Individual student's surfing time on the on-line materials was calculated. After the class, students needed to send their homework to the website. Whenever they had English problems, they could either go to the discussion board and ask for help from classmates, or send emails to the teacher. This on-line learning source allowed partial learner control with teacher guidance.

The second source of on-line learning was an interactive CD-ROM, provided by the LIVE ABC COMPANY. The CD-ROM offered a wealth of self-study activities, which

allowed students to preview and review the content of the OFF WE GO textbook either online or offline. When students worked online, their learning hours were recorded. The CD-ROM had many functions such as reading aloud at a fast/slow speed, repeated reading, recording, speech recognition, video of real-life situational conversations, subtitles in Chinese, dictionary, and reading/listening/speaking/writing exercises (See figures 3-7).

The third source of online learning was an e-learning website (LIVE ABC Internet English College, located at http://203.64.90.116/z_liveorg7/), also provided by LIVE ABC COMPANY. This website provided a web-based English proficiency diagnosis assessment, self-learning courses, electronic resources (a picture dictionary, dialogues, and an interactive theater), GEPT and TOEIC practice tests, games, discussion boards, and students' on-line learning records. In this learner-controlled learning environment, learners decided their own learning contents, sequence and pace of learning (See Figures 8-9).

教材開放	科目	科目名稱	章次	章名	更新日期	線上教材				討論區		
						編輯	檔案明細	大小	連結	討論	開放	給老師
<input checked="" type="checkbox"/>	ENG	實用英文	1	Introduction	未設定	<input type="button" value="編輯"/>	<input type="button" value="檔案明細"/>	62,464	<input type="button" value="連結"/>	<input type="button" value="討論 (0)"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Figure 1: KUAS Learning Platform

3. **push** [puʃ] verb [T] INFORMAL 【俚】販賣(毒品)。

to sell illegal drugs:。

He was arrested for pushing drugs to schoolchildren.

pusher [ˈpuʃə] noun [C] (ALSO **drug pusher**)。

毒品販賣者。

someone who sells illegal drugs.

4. **junk** (DRUG) noun [U] MAINLY US SLANG.

Figure 2: On-line teaching materials

Unit 8: Reading Fountain *How about a movie?*

Stella,

How have you been recently? **Still busy balancing classes and part-time jobs?** Hey, how about seeing a movie with me this Saturday?

I read a movie review recommending a Japanese film. The movie, which won an award at a film festival last month, will probably be very good. It's not a big, Hollywood production. The review says **史黛拉:**

最近在忙什麼啊? 依然奔波於課業與打工中嗎? 這禮拜六去看場電影好不好?

文章推薦了一部上個月才在影展得獎的日本片, 應該滿不錯的。雖然不是好萊塢的大片, 但影評說這

Figure 3: Reading/Listening/Speaking Exercises Sample

concession stand
[kən'seʃən][stænd]
n. 販賣部; 飲食部

Figure 4: Vocabulary

Unit 8 Lounge Chat : Part 1

Drag each word into the blank. Listen to the complete sentence by clicking on the speaker icon.

Q. Where are you?

A. I'm at the entrance

at the emergency exit in front of the entrance under the show-time monitor
at the box office in front of the emergency exit at the concession stand
in the rest room in front of the rest room
in front of the poster in front of the concession stand
under the preview monitor

Figure 5: Click and Drag Exercise



Figure 6: Video on Real-life Situational Conversations

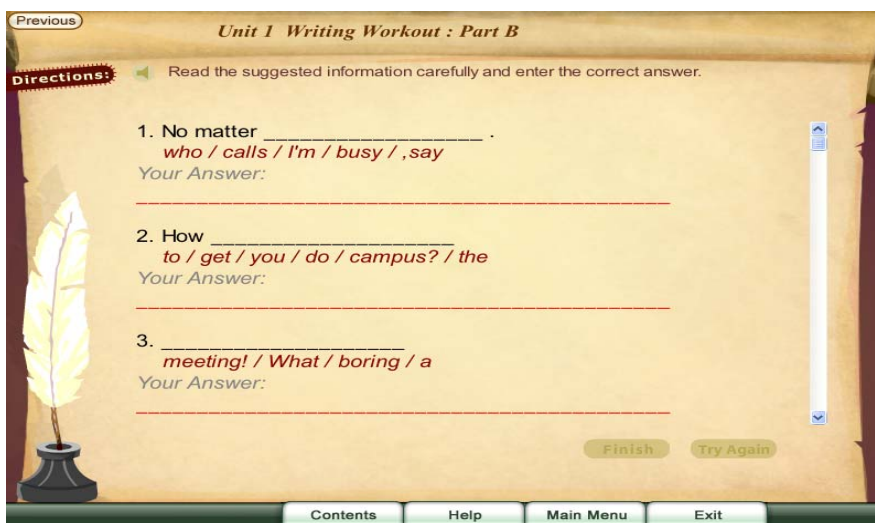


Figure 7: Writing Exercises Sample



Figure 8. LIVE ABC Internet English College



Figure 9: Interactive Theatre

Integrated Classroom Instruction

The class goal was to train students' English listening, speaking, reading, and writing skills. Most of the materials presented in class were electronic files such as Word documents, Power Point files, and CD/DVD programs, which were designed to give students various visual and audio inputs and help them to reinforce the language learned in class. The instructional language was mainly English, while Mandarin was also used when it was necessary.

Both textbooks were taught in turn; that is, one week the class studied OFF WE GO, and another week the class studied the leveled reader. When the class content was OFF WE GO, a typical lesson plan involved the teacher first asking several questions regarding the text to warm-up students' background knowledge. Then the teacher introduced some new vocabulary items. Third, the whole class listened to the text CD. After that, the teacher explained the text and raised some related issues. Students in groups discussed the issues, and presented their answers orally. Sometimes, groups wrote down their opinions, and then groups exchanged their papers and conducted peer correction. Furthermore, while the class was watching a video on the CD-ROM, the teacher would pause the disc and ask students some questions. Other classroom activities such as group games, singing songs and role-playing were also included.

As for the leveled reader teaching, students were divided into groups, one for each chapter of the book. Each group was responsible for looking up unfamiliar vocabulary in their assigned chapter, and then they posted the meanings, with some example sentences, on the KUAS website before they came to the class. Moreover, the teacher would post several guided questions on the website as well. In class, groups in turn presented the summary of their assigned chapter and then students discussed the guided questions together.

In addition to the textbooks, students in groups were required to choose another leveled reader as their outside-reading from a book list provided by the teacher. They orally presented their story to the class two weeks before the final examination. Students had to pay attention to their peers' presentation, because questions about its content would be included in the final examination.

Course Assessments

Students' achievements were evaluated based on their participation in online learning and English performance in listening, speaking, reading, and writing. In detail, the total scores consist of an oral presentation of the outside-reading (10%), a written reflection report (10%), reading comprehension quizzes (15%), listening comprehension quizzes (15%), on-line learning activities (10%), mid-term exam (20%), and final exam (20%). The mid-term

examination included a reading aloud of the texts (20%) and a paper-and-pencil test (80%). The final examination was a paper-and-pencil test, and test questions came from the students' group presentations (50%) and the learning materials in class (50%).

Procedures of the Data Analysis

After the data were collected, descriptive statistics were used for the first research question: "What are students' perceptions of the integrated instruction and classroom activities?" and the second research question: "How much effort do students make to learn English after class?" For the third research question: "What is the effect of the integrated instruction on students' learning achievements?", a paired t-test was applied. Finally, an independent sample t-test was applied to answer the fourth question: "Does significant difference exist between high ability and low ability students' learning achievements?"

RESULTS AND DISCUSSION

Students' Perceptions of the Integrated Instruction

In general, 91% of the students believed this integrated instruction was helpful to them. Moreover, 81.1% of the students agreed that this course enhanced their interest in learning English. It seems that most students didn't feel anxious in this course, since only 27.8% of the students experienced pressure. In addition, although listening and reading comprehension quizzes were regularly held every other week, the majority of students did not perceive them as a heavy burden. This response was probably, was due to the fact that those quizzes were not too difficult. Another reason might be that the teacher was kind and gentle in her attitude toward students. Actually, only 7 students thought the teacher was very severe. This easy-going teaching style might be different from that of their high school teachers, who are often very strict with students.

As for the teaching materials, most of the students felt that materials were neither too difficult nor too easy (73.6%), and fitted in with their needs (78.3%). Furthermore, 82.6% of the students believed KUAS Learning Platform website was helpful, and 74.5% of the students thought the CD-ROM of OFF WE GO was useful. Finally, 90.5% percent students perceived their English had improved due to the training of this course.

In short, the majority of students gave positive evaluation to this course in terms of instruction, the teacher, and the learning materials. Therefore, this study suggests that an integrated instruction is worth advocating to other teachers (See Table 1).

Table 1 *Students' Perceptions of the Integrated Instruction*

Description	SA	A	DA	SD
	N (%)	N (%)	N (%)	N (%)
1. This course is helpful to my English learning.	35(16.5%)	158(74.5%)	18 (8.5%)	0
2. This course makes me become interested in learning English.	35(16.5%)	137(64.6%)	38(17.9%)	0
3. I have much pressure in this course.	16(7.5%)	43 (20.3%)	125(59%)	27(12.7%)
4. I think there are too many quizzes in class.	16(7.5%)	53(25%)	120(56.6%)	23(10.8%)
5. Teachers are too severe.	1(0.5%)	6(2.8%)	79(37.3%)	125(59%)
6. Teaching materials are appropriate for my English ability.	20(9.4%)	136(64.2%)	48(22.6%)	7(3.3%)
7. Teaching materials fit in with my needs.	20(9.4%)	146(68.9%)	37(17.5%)	7(3.3%)
8. KUAS Learning Platform is helpful to me.	33(15.6%)	142(67%)	30(14.2%)	7(3.3%)
9. The CD-ROM of OFF WE GO is useful.	36(17%)	122(57.5%)	36(17%)	14(6.6%)
10. I think my English has improved due to this course.	70(33%)	122(57.5%)	18(8.5%)	2(0.9%)

SA= Strongly Agree A= Agree DA=Disagree SD= Strongly Disagree

Students' Favorite Classroom Activities

Regarding the rankings of students' favorite classroom activities, the item "Watching DVD films" was ranked as the most popular, "Christmas celebration" was the students' second highest choice, and "Watching the video of OFF WE GO" was ranked as the third favorite (See Table 2). Apparently, students liked the video-assisted English teaching, no matter whether the content of the selected video was relevant to the content of the textbook or not. Therefore, teachers could utilize videos in class to enhance students' learning interests. It would be even better if teachers could design some activities relevant to the content of each film before/after watching the film. For example, teachers could ask students to predict the plot of the film, discuss the characters of the actors/actresses, or write a reflection paper of the film. Since writing papers was ranked as students' least favorite activity in this study, teachers may need to give students more guidance in the process of writing or put students into groups. In this way, writing a collaborative paper with peers might be more fun and also easier to achieve.

Students' second favorite activity – a Christmas celebration event - was held to help students understand holiday culture. The design of this activity originated from Prof. Chern Chiu-Lan's presentation in the 2007 Happy Summer English Camp at St. John's University. In this activity, students played the role of Secret Santa, and had to write a card to one classmate. On the card, they described themselves in English, and the class had to guess and identify the Santa from their description. During the event, the teacher played Christmas songs and treated the class to candy. Therefore, the class was filled with joy. Students' positive evaluation of this activity supported previous studies which highlighted the effectiveness of incorporating holiday celebration events in teaching culture (Cheng, 2006; Herron, Cole, & Corrie, 1999; Kramsch, Cain, & Murphy-Lejeune, 1996; Krasner, 1999).

Table 2 *Students' favorite classroom activities*

Activities	Mean	First favorite	Second favorite	Third favorite	Rank
		N (%)	N (%)	N (%)	
Watching DVD films	2.5167	132(62.3%)	19(9.0%)	19(9.0%)	1
Christmas celebration event	1.2889	15(7.1%)	79(37.3%)	29(13.7%)	2
Watching the video of textbook-OFF WE GO	0.7833	10(4.7%)	33(15.6%)	45(21.2%)	3
Group presentation of the story book	0.6278	10(4.7%)	20(9.4%)	43(20.3%)	4
Listening to teacher's lectures using Power Point files	0.5611	10(4.7%)	20(9.4%)	31(14.6%)	5
Writing a reflection report of the story book	0.2278	3(1.4%)	10(4.7%)	12(5.7%)	6

Students' Efforts to Learn English after Classes

Surprisingly, approximately 60% of the students spent less than one hour using the website resources. Although level III students spent a little longer than level I students, both groups did not work hard enough. Students' answers to the open-ended question, "What are your suggestions and comments on this course?" might shed some light on their passive self-learning. Some students commented that they had neither internet access for online learning nor the printer to print out the documents on the website. Some said that reading electronic articles on screen was tiring. And others complained about the unstable quality of the CD-ROM. Many of them claimed that the CD-ROM couldn't be played on their computers or even crashed their computers. After consulting the LIVE ABC Company, the researcher understood that when students' computers lacked some necessary programs to run the CD-ROM, they wouldn't be able to use the CD-ROM. Therefore, how to help students successfully install these programs in their computers might be the pre-requisite of

web-learning. It would be very helpful if there was a clear installation manual or if the company could send a computer technician to the class to solve students' problems at the beginning of the program.

Table 3. *Students' Efforts to Learn English after Class*

Efforts		Total	<1 hour		1-3 hr (s)		3-5 hrs.		5-7 hrs.		> 7 hrs.	
		Mean	N	%	N	%	N	%	N	%	N	%
The weekly time I spent on KUAS Learning Platform was....	I	1.459	61	62.2	30	30.6	6	6.1	1	1.0	0	0
	III	1.460	68	59.6	40	35.1	3	2.6	2	1.8	0	0
	T	1.459	129	60.8	70	33	9	4.2	3	1.4	0	0
The weekly time I spent on the CD-ROM of OFF WE GO was....	I	1.469	61	62.2	29	29.6	7	7.1	1	1.0	0	0
	III	1.557	62	54.4	42	36.8	7	6.1	1	0.9	1	0.9
	T	1.516	123	58	71	33.5	14	6.6	2	0.9	1	0.5

I= Level I II=Level II T= Total

The Effect of the Instruction on Students' Learning Achievements

Table 4 indicates that level I students had significant improvements, although there was evident deviation among this group of students. In contrast, level III students generally showed slight regression. This finding was surprising, as it partly contradicted a previous study which revealed that the whole language approach was effective only for high ability listeners (El-Koumy & Abdel, 2000). Some tentative interpretations for this result are as follows. First, since the task-based instruction in this study encouraged cooperation and assistance among peers, students could work on assignments together and experience a sense of achievement. This is a feeling which students with low English proficiency ability might seldom experience in their learning. Thus, level I students became more confident in the classroom. As for higher level students, they might have been very self-confident and used to relying on themselves. Therefore, they probably did not highly value peer cooperation and then became less motivated in group work. This finding supports Wu's study (2003), which demonstrated that the effect of cooperative learning on freshmen's English reading comprehension was especially significant in low-level students. Similarly, the present study also echoes Hsiao's (2006) observation about low achievers in cooperative learning. In her study, the low achievers liked English mainly because there was less pressure to learn, more chances to develop social relationships, and build confidence with other classmates. Moreover, according to Kan's opinion (2006), when students are more confident, they are more likely to take risks and learn more.

Another possible explanation for this result might be due to the fact that level III student's classes were cancelled twice because they clashed with a national holiday and a school anniversary. As a result, they had only 16 weeks of classes and had less writing practice and grammar explanation, compared with level I participants.

Furthermore, a more likely explanation may rest in the nature of the pre/post-tests, which are mainly reading comprehension tests. Since the goal of the integrated English instruction is to train students' listening, speaking, reading, and writing skills, students' post-test scores might not have reflected learning gains in listening, speaking and writing.

Table 4. *Students' English Learning Outcomes*

Level	N.	Pre-test			Post-test			Learning Outcomes		
		Mean	S.D.	t	Mean	S.D.	t	Mean	S.D.	t
I	98	57.77	14.07	-20.45***	73.63	16.62	-7.355***	15.86	19.28	7.90***
III	99	88.04	4.33		87.13	7.53		-0.91	8.58	

***= p<0.001

CONCLUSIONS AND RECOMMENDATIONS

Several conclusions and recommendations for future studies can be drawn from this study. First, the results support the view that integrated English instruction is beneficial to technical university students. It helped to enhance not only students' learning interests but also students' English performance. Although only low ability students showed improvements in the study, the majority of students perceived their English had improved due to the course. Since this study only measured students' reading comprehension ability, future studies could adopt an English test which measures students' listening, reading, speaking, and writing abilities. If it is difficult to implement a more integrated test, further research could ask students to separately evaluate their improvements in discrete language components.

Furthermore, the findings of the present study suggest that integrated English instruction might not be so effective when applied to high ability students; further research is needed to investigate why low ability students seem to benefit more from this kind of instruction than high ability students.

Finally, an interesting finding of this study is that while students felt online learning was an excellent vehicle for their English study, lots of students spent less than one hour on the electronic resources provided by the course. Therefore, future studies could investigate the problems students encounter in the process of e-learning and find some solutions to these problems. For example, researchers could explore whether technical problems and access to network resources are barriers in online self-learning, and when these problems occur, what methods are available to handle or prepare for network difficulties.

REFERENCES

- Anderson, G.S. (1984). *A whole language approach to reading*. Lanham: University press of America.
- Chang, M. M., & Lehman, J. D. (2002). Learning foreign language through an interactive multimedia program: an experimental study on the effects of the relevance component of the ARCS model. *CALICA Journal*, 20 (1), 81-98.
- Chang, Y. C. (2004). *An automatic collocation writing assistant for Taiwanese EFL learners based on NLP technology*. Unpublished doctoral dissertation, National Tsinghua University, Taiwan.
- Chen, Z. F., & Tseng, H. C. (2006). Effects of implementing the computer assisted problem-posing strategy on English vocabulary learning achievement and affective reaction. *Journal of National Taipei University of Education*, 19 (1). 89-118.
- Cheng, C. L. (2006). *Integrating English learning websites into listening comprehension instruction at college level: An evaluative study*. Unpublished master's thesis, Tamkang University, Taiwan.
- Cheng, C. W. (2006). *A study of the teaching of holiday culture across different grade levels in English class in municipal elementary schools in Taipei City*. Unpublished Master's thesis. National Taipei University of Education, Taiwan.
- Chen, C.Y. (2005). *A decade of changes in the English ability of university freshmen in Taiwan*. Unpublished Master's thesis. Tunghai University, Taiwan.

- Chen, K. (2004). *An investigation on the impact of ASR software feedback on EFL college students' pronunciation learning*. Unpublished doctoral dissertation, National Taiwan Normal University, Taiwan.
- Cheng, Y. Y. (1998). *Adapted whole language teaching in an English class in junior high school*. Unpublished Master's thesis. Kaohsiung Normal University, Taiwan.
- Chern, C. L. (2000). Junior high school students' attitudes toward integrated English teaching activities. *English teaching*, 24(3), 69-76. (NSC 87-2411-H-029-007)
- Chiang, Y. H. (2005). Freshman English at Providence University. Paper presented at the symposium on Freshman English for Non-majors Program in Taiwan, Taichung, Taiwan.
- Chiu, T. L. (2005). *Effects of online conversation materials with the support of speech recognition technology on college EFL learners*. Unpublished Master's thesis. National Tsinghua University, Taiwan.
- Chu, H. S. (2005). Freshman English at National Taiwan Normal University. Paper presented at the symposium on Freshman English for Non-majors Program in Taiwan, Taichung, Taiwan.
- Chu, Y. M. (2004). *Development of an English learning passport system for primary school using speech technologies*. Unpublished Doctoral dissertation, National Central University, Taiwan.
- Danehower, V. (1993). Implementing whole language: Understanding the change process. *Schools in the Middle*, 2(4), 45-46
- Doake, D. (1994). The myths and realities of whole language: An educational movement at risk. In Alan D. Flurkey and Richard J. Meyer (Eds.), *Under the Whole Language Umbrella: Many Cultures, Many Voices* (pp. 125-157). Urbana, Illinois: National Council of Teachers of English.
- Dudley-Marling, C. (1995). Whole language: It's a matter of principles. *Reading and Writing Quarterly: Overcoming Learning Difficulties*, 11(1), 109-117.
- ETS (2007). Taiwan TOEIC Score Statistics Report. Retrieved July 18, 2008 from <http://www.toeic.com.tw/stat.htm>
- El-Koumy, A., & Abdel, S. A. (2000). Effects of skills-based versus whole language approach on the comprehension of EFL students with low and high listening ability levels. *ERIC Document*, ED 449670.
- Eldredge, J. (1991). An experiment with a modified whole language approach in first grade classrooms. *Research and Instruction*, 30, 21-38.
- Ehsani, F., & Eva, K. (1998). Speech technology in computer-aided language learning: Strengths and limitations of a new CALL paradigm. *Language Learning and Technology*, 2 (1), 45-60.
- Eyerman, L. (2005). Freshman English at Ming-chuan University. Paper presented at the symposium on Freshman English for Non-majors Program in Taiwan, Taichung, Taiwan.
- Freeman, D., & Freeman, Y. (1994). Whole language principles for bilingual learners. In Alan D. Flurkey and Richard J. Meyer (Eds.), *Under the Whole Language Umbrella: Many Cultures, Many Voices* (pp. 240-263). Urbana, Illinois: National Council of Teachers of English.
- Freeman, D., & Freeman, Y. (1992). *Whole Language for Second Language Learners*. Portsmouth, NH: Heinemann.
- Goldenberg, C. (1991). Learning to read in New Zealand: The balance of skills and meaning. *Language Arts*, 68, 555-562.
- Herron, C., Cole, S. P., Corrie, C., & Dubreil, S. (1999). The effectiveness of video-based curriculum in teaching culture. *The Modern Language Journal*, 83(4), 518-533.
- Ho, C. F. (2003). *The effects of computer-based cooperative learning on the vocabulary acquisition*. Unpublished Master's thesis. National chiayi University, Taiwan.
- Hsiao, Y. C. (2006). *Effects of tracking on the motivation and English achievements of junior*

- high EFL students in Taiwan. Unpublished Master's thesis. National Kaohsiung Normal University, Taiwan.
- Huang, M. C. (2004). *The effectiveness of incorporating CALL into English alphabet teaching*. Unpublished Master's thesis. Taipei Teacher's College, Taiwan.
- Hung, S. M. (2001). *The effects of constructive reading courseware on EFL reading instruction for junior high school students*. Unpublished Master's thesis. National Kaohsiung Normal University, Taiwan.
- Kan, Y. F. (2006). *A study on the correlation among the cognitive Style, mathematics attitude and mathematical problem-solving performance of grade 5th students in elementary schools*. Unpublished Master's thesis. National Kaohsiung Normal University, Taiwan.
- Kramsch, C., Cain, A., & Murphy-Lejeune, E. (1996). Why should language teachers teach culture? *Language Culture and Curriculum*, 99-100.
- Krasner, I. (1999). The role of culture in language teaching. *Dialog on Language Instruction*, 13(1-2), 79-88.
- Lai, Y. S. (2005). Freshman English at National Cheng Kung University. Paper presented at the symposium on Freshman English for Non-majors Program in Taiwan, Taichung, Taiwan.
- Liang, Y. M. (2003). *The study on the effects of auditory multi-media instrument on the development of English oral abilities for Taiwanese EFL elementary-school children- the application of Talkworks*. Unpublished Master's thesis. Taipei Teacher's College, Taiwan.
- Lin, H. T. (2007). *An investigation of the effectiveness of the Tunghai method to freshman English for Non-English Majors*. Unpublished Master's thesis. Tunghai University, Taiwan.
- Liu, P. L., & Chen, C. J. (2008). Trends and issues in CALL research in Taiwan, 2000-2006: A glance. *Studies in English Language and Literature*, 21, 25-35.
- Lu, K. Y. (2002). *Effect of e-mail exchange on EFL senior high school students' writing*. Unpublished Master's thesis. National Tsinghua University, Taiwan.
- Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constrains on multimedia learning; When presenting more material results in less understanding. *Journal of Educational Psychology*, 93 (1), 187-198.
- Moosavi, S. A. N. (2007). An integrated holistic method for teaching introductory German to community college students. *Community College Journal of Research and Practice*, 31, 797-812.
- Moras, S. (2001). Computer-Assisted Language Learning (CALL) and the Internet. Retrieved March 8, 2009 from <http://www3.telus.net/linguisticsissues/CALL.html>
- Newman, J., & Church, S. (1990). Commentary: Myths of whole language. *The Reading Teacher*, 44(1), 20-26.
- Norris, J. & Hoffman, P. (1993). *Whole language intervention for school-age children*. San Diego, California: Singular Publishing Group, Inc.
- Nunan, D. (1995). Closing the gap between learning and instruction. *TESOL Quarterly*, 29 (1), 133-158.
-
- Oxford, R. (2001). Integrated skills in the ESL/EFL classroom. ERIC Digest.
- Sanacore, J. (1995). What to Look for in Whole Language Content Area Classrooms. *ERIC Document*, ED381761.
- Sheu, C. M., Wang, P. L. (2007). A Study on Teachers' Perceptions towards Between-class Ability Grouping in Freshman Practical English Class-----A Case Study of KUAS. *Journal of Shu-Te University*, 9 (1), 41-61.
- Shu, L. N., Wang, P. L. (2006). A Study on English Graduation Threshold Policy and Its Effect on Students Learning Motivation. *Journal of National Kaohsiung University of*

Applied Sciences, 35, 51-67.

- Sim, J. (2005). Freshman English at Tunghai University. Paper presented at the symposium on Freshman English for Non-majors Program in Taiwan, Taichung, Taiwan.
- Stepp-Greany, J. (2002). Students perceptions on language learning in a technological environment: Implications for the new millennium". *Language Learning and Technology*, 165-180.
- Thelmadatter, L. (2007). The computers are coming ... are here!.. *TESOL Greece Newsletter*, 95.
- Tsai, P. H. (2003). *An analysis of the effects of multimedia on English listening comprehension of junior high school students*. Unpublished Master's thesis. National Kaohsiung Normal University, Taiwan.
- Tu, Y. F. (2004). *The application of the whole language approach in instruction of English poems in senior high School*. Unpublished master's thesis. Kaohsiung Normal University, Taiwan
- Reutzel, D., & Cooter, R. J. (1992). *Teaching children to read: From basal to books*. New York: Merrill/MacMillan.
- Reutzel, D., & Hollingsworth, P. (1988). Whole language and the practitioner. *Academic Therapy*, 23(4), 405-416.
- Rosberg, M. (1995). Integrated Approaches to Learning, *ERIC Document*, ED 389684.
- Wang, C. Y. (2003). *Language experience approach: A pilot study on its effects on Taiwanese junior high school students' reading ability*. Unpublished Master's thesis. National Taiwan Normal University, Taiwan.
- Warschauer, M. (1996) [Computer-assisted language learning: An introduction](#). In Fotos S. (ed.) *Multimedia language teaching*. Tokyo: Logos International.
- Wu, H. Y. (2003). *A study of an application of cooperative learning to freshman English teaching*. Unpublished Master's thesis. Tzu-Chi University, Taiwan.

APPENDIX

Questionnaire

第一部份：學生基本資料

1. 學院：工學院 商管學院 人文學院
2. 系別：機械系 土木系 化材系 工管系 企管系
財稅系 會計系 國企系 人資系 文化系
3. 性別：男 女
4. 分級：I III

第二部份：對本學期大一英文上課的感想

(4=非常同意；3=同意；2=不同意；1=非常不同意)

- | | | | | |
|--|---|---|---|---|
| 1、此課程對我的英語學習有幫助 | 4 | 3 | 2 | 1 |
| 2、此課程讓我對學習英語更有興趣 | 4 | 3 | 2 | 1 |
| 3、我覺得壓力大 | 4 | 3 | 2 | 1 |
| 4、我覺得考試太多 | 4 | 3 | 2 | 1 |
| 5、我覺得老師太兇 | 4 | 3 | 2 | 1 |
| 6、我覺得上課使用的教材適合我的程度 | 4 | 3 | 2 | 1 |
| 7、我覺得上課使用的教材符合我的學習需求 | 4 | 3 | 2 | 1 |
| 8、我覺得 KUAS 學習平台對我的英語學習有幫助 | 4 | 3 | 2 | 1 |
| 9、我覺得 OFFWEGO CD-ROM 對我的英語學習有幫助 | 4 | 3 | 2 | 1 |
| 10、我覺得經過一學期的訓練，我的英文進步了 | 4 | 3 | 2 | 1 |
| 11、我每週使用老師的線上學習平台頻率 | | | | |
| <input type="checkbox"/> 少於1小時 <input type="checkbox"/> 1-3小時 <input type="checkbox"/> 3-5小時 <input type="checkbox"/> 5-7小時 <input type="checkbox"/> 7小時以上 | | | | |
| 12、我每週使用教科書 LIVE ABC 互動光碟的頻率 | | | | |
| <input type="checkbox"/> 少於1小時 <input type="checkbox"/> 1-3小時 <input type="checkbox"/> 3-5小時 <input type="checkbox"/> 5-7小時 <input type="checkbox"/> 7小時以上 | | | | |
| 13、請挑出最喜歡的前三項上課活動 | | | | |
| (以 3, 2, 1 表示；3=最喜歡，2=其次，1=次之) | | | | |
| <input type="checkbox"/> 看 LIVE ABC 影片光碟 <input type="checkbox"/> 聽 LIVE ABC PowerPoint 講解 <input type="checkbox"/> 看電影 | | | | |
| <input type="checkbox"/> 故事書分組報告 <input type="checkbox"/> 寫故事書讀後心得 <input type="checkbox"/> 耶誕慶祝活動 | | | | |
| 14. 您對這門課的建議為何? | | | | |

An Analysis of Collocational Errors in Chinese Learner English Corpus

Feng-fan Wei (魏峰帆), Tsuo-lin Chiu (邱作麟)

National Taiwan Normal University

ivan90883@hotmail.com
stanneko@gmail.com

Vocabulary teaching and learning have gradually received attention in recent years. Among different aspects of vocabulary knowledge, collocation, which refers to word chunks that occur together frequently, is often neglected in vocabulary teaching. Learners tend to have difficulty producing appropriate collocations. Research has shown that explicit teaching of collocation could improve learners' collocational competence. However, there are numerous collocations and language teachers often hesitate about what collocations should be taught in classes. Fortunately, studies of learner corpora have received attention in recent years, which may provide suggestions for choosing collocations that should be attended to in language classes. The present study adopts error analysis approach to investigate collocational error types and causes of Chinese Learner English Corpus (CLEC). The corpus contains around one million words of English collected from Chinese learner of English with various levels of English proficiency. Two of the most frequently produced lexical collocation errors, verb-noun and adjective-noun collocation errors, were chosen for analysis. The result revealed that L1 transfer and use of words which share similar semantic features with target words are the most frequently produced error types.

INTRODUCTION

Vocabulary teaching and learning have gradually received attention in recent years. Among different aspects of vocabulary knowledge, collocation is often a neglected aspect in vocabulary teaching (Farghal & Obiedat, 1995). Learners of English tend to have problems producing correct combinations of two or more words (Bahns, 1993). It is therefore necessary to highlight the importance of collocation teaching and learning.

Collocation

With increasing emphasis on lexical approach, teaching of collocation has become crucial. The definitions of collocation are many. Firth (1957: p.196) defined collocation as "the company words keep together." Lewis (2000) stated that collocation is "the way in which words co-occur in natural text in statistically significant ways (p.132)." Based on Lewis's definition, collocation resulted from native speakers' habitual expressions in natural context. Nation (2001) provided two criteria for determining what could be called a collocation. Collocations are regarded as items which "frequently occur together and have some degree of semantic unpredictability (p.317)." Benson, Benson, and Ilson (1986) further divided collocations into two major categories: grammatical collocations and lexical collocations. A grammatical collocation consists of a noun, an adjective, or a verb, plus a grammatical structure such as an infinitive or clause. A lexical collocation, on the other hand, is a phrase combining two lexical components, such as verb-noun, adjective-noun, and adverb-adjective collocations.

The importance of collocation in language learning

A number of arguments have been proposed to support the importance of collocational knowledge. First, learning collocations could reduce processing time. Carter and McCarthy (1988) suggested that learning collocations could speed up language processing. They maintained that when learners produced speeches, they could use collocations as “pre-packaged building blocks” without thinking about how to restructure the language they wanted to convey. Hill (2000) also held similar opinions. He remarked that collocation allowed learners to think quickly and communicate more efficiently.

Second, the teaching and learning of collocations were important to second/foreign language learners because collocations were not always predictable. Lewis (2000) stated that many collocations are non-literal and idiomatic. For example, native speakers usually used “open a meeting” instead of “start.” Without collocation knowledge, learners may have difficulty interpreting or producing idiomatic collocations. Hill (2000) argued for the importance of collocations from a pedagogical aspect. He maintained that teaching collocations enabled students to learn words in chunks instead of separate items. If learners could correctly chunk the words, they would understand a text more easily and improve their pronunciation because they could read words in chunks, not individually.

Studies on collocation errors

Perceiving the importance of collocation, researchers in Taiwan have conducted studies on learners’ collocation errors (Liu, 1999; Chen, 2002; Hsueh, 2002; Tang, 2004). The results showed that verb-noun and adjective-noun collocations were the most frequent produced errors.

Liu (1999) investigated collocation errors in college students’ writing. She collected 127 copies of final examination papers and 94 copies of students’ composition for analysis. In terms of lexical collocation errors, the result showed that verb-noun collocation error occurred most frequently. Students produced collocations such as **pay an effort* instead of *make an effort*. With respect to the sources of errors in L1 collocations, errors were mainly due to (1) false concepts of de-lexicalized verbs such as *make* and *do* (e.g. **do plans* instead of *make plans*), (2) the misuse of synonym (e.g. **call at his parents* instead of *call on his parents*), and (3) negative transfer (e.g. **eat vitamins* instead of *take vitamins*). In addition, a small number of errors were due to word coinage (e.g. **see sun-up* instead of *see the sunrise*), and approximation (e.g. **fell the exam* instead of *failed the exam*). Overall, the study found that the number of grammatical errors were larger than that of lexical errors. In all the types of errors, negative transfer was a main source of collocational errors.

Tang (2004) also conducted a study focusing on college students’ collocation errors. In addition to written production, she also examined students’ oral collocation errors. The study revealed that verb-noun and adjective-noun collocations were the most frequent produced lexical collocation errors. The students tended to use verbs with similar pronunciation. For example, “*Community with” was used instead of “communicate with.”

Wu (2005) investigated EFL learners’ errors of VN collocations specifically. Data were collected from writings of 178 English majors in four academic years. Four error categories, which were (1) lexical transfer, (2) phonological similarity, (3) grammatical irregularity, and (4) semantic choice, proposed by Lombard (1997), were adopted for coding learners’ VN miscollocations. The results showed that incorrect semantic choice was the main cause of learners’ errors.

Different from above studies, Chen (2000) and Hsueh (2002) analyzed senior high school students’ collocation errors. Chen collected 90 exam papers from 30 students and analyzed their collocation errors. It was found that adjective-noun and verb-noun were the most frequent lexical collocation error types. However, the causes of errors were not discussed. The researcher suggested that due to the lack of empirical evidences, determining

the causes of errors was quiet challenging. Hsueh (2002) collected 261 English compositions from 87 senior high school students. She analyzed adjective-noun and verb-noun collocation errors. In all adjective-noun collocation errors, spelling error was the highest error type. Direct translation error and use-of-wrong-derivative error occupied the second and third place respectively. She also examined the relationship between learners' proficiency level and their error types. Students of lower proficiency level were likely to use direct translation strategy to convey meaning than students of mid level. Another feature was that lower level students tended to make errors related to the use of nouns, while mid and high level students tended to use wrong adjectives.

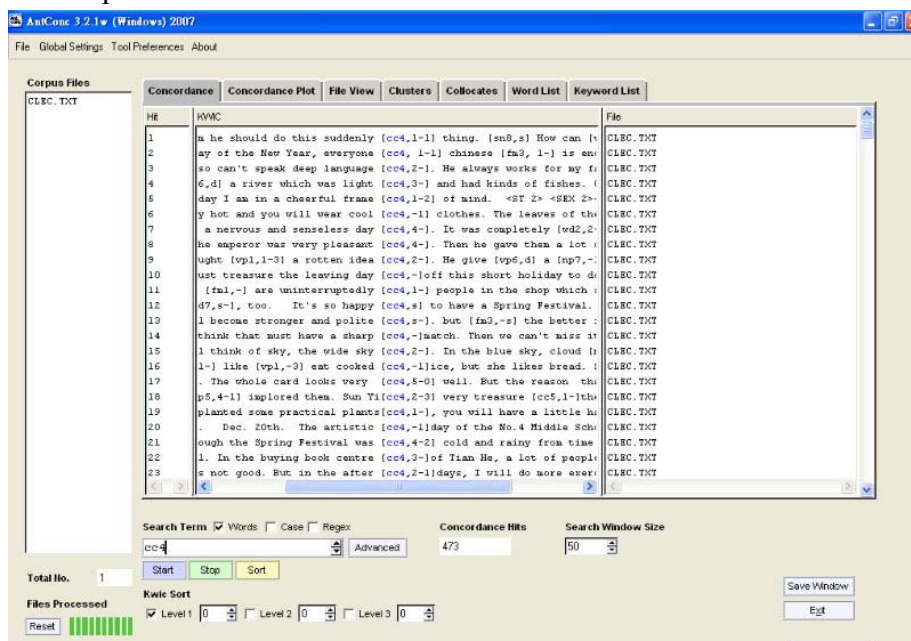
The current study attempts to investigate the VN and AN collocation errors collected from Chinese Learner Error Corpus (see next section for details). On the basis of the results from previous analysis, the present study tries to gain further insights into the patterns of uses of VN and AN collocations by Chinese EFL learners from both quantitative and qualitative perspectives.

METHOD

Data source

Examples of EFL learners' VN miscollocaitions were collected from Chinese Learner Error Corpus (CLEC). CLEC contained around one million words of English compositions collected from Chinese learners of English with various levels of English proficiency, ranging from secondary school students, non-English-major university students, to English-majors in China. The corpus was error-tagged according to an error marking scheme of 61 types of errors. These include various lexical, grammatical, semantic and sentence level errors (information retrieved from <http://langbank.engl.polyu.edu.hk/corpus/clec.html>). Moreover, concordance tool, AntConc, was used for extracting the specific type of errors needed for the study. AntConc was a freeware concordance program capable of generating concordance lines and concordance distribution plots. It was used to analyze word clusters, word frequencies, collocates, and key words. Figure 1 demonstrates the use of AntConc for retrieving target collocation.

Figure 1 Example of collocations taken from AntConc



Classification system of collocation errors

VN collocation errors. There were a total of 1570 phrases of VN collocation errors found in CLEC. The current study randomly chose 196 out of 1570 phrases for analysis. The categorizations of error types by Lombard (1997) as used in Wu (2005) were adopted for quantitative analysis of VN miscollocations. After quantitative analysis, certain recurring/significant error patterns were presented with examples from the corpus for further discussion.

AN collocation errors. The classification system was adapted from the categorization used by Liu (1999) and Hsueh (2002). Liu analyzed general collocation errors and identify seven sources of errors, including overgeneralization, ignorance of rule restrictions, false concept hypothesized, use of synonyms, negative interlingual transfer, word coinage, and approximation. Hsueh modified Liu's classification, deleting sources that did not occur in her data and further adding four more error types, which were false concept errors, incorrect-form-of-compound-word errors, spelling errors, and use-of-wrong-derivative errors, to her category system. The present study combined and modified classification used by Liu and Hsueh. Seven types of error were used for categorization, which were (1) use of wrong part of speech, (2) L1 transfer, (3) approximation, (4) use of sunonym, (5) ignorance of rule restrictions, (6) false concepts, and (7) miscellaneous.

RESULTS AND DISCUSSION

VN collocation errors

Among 196 VN collocational errors retrieved from Chinese Learner Error Corpus, 24 (12%) were wrong uses of noun while 172 (88%) were wrong uses of verb. Liu (2002) has found that 96% of VN miscollocations were due to wrong choice of verb. The result from the current study was somehow consistent with Liu (2002) in which the majority of VN micollocaitons were verb errors. The following two examples demonstrate errors in noun collocates.

“He smashed his *first* into the street where Jack lived”
“it should meet your own *interesting*”

Excluding the 24 noun errors, the remaining 172 errors in verb collocates were further examined according to Lombard (1997)'s categorization of error types. The first type of errors examined in the current study was semantic choice. According to Lombard, semantic choice refers to “the use of near synonym in place of the appropriate word for the given context” (p. 85). There were 81 out of 172 (47%) errors found in this category, which was the most frequently error found in the current data bank. Examples of this type of errors are given below:

“do our best to *win* greater improvement”
“whatever jobs they take and *make* responsibility for other perpons”

The second type of error examined was lexical transfer. Lexical transfer refers to “the reliance on the native language in selecting lexical items which become members of English language collocations” (Lombard, 1997, p. 60). The result showed that 38 out of 172 (22%) errors fall into this category, which was the second most frequently made errors. In fact, while coding on error in this category, it was sometimes difficult to define whether the error is due to L1 transfer or semantic choice. This ambiguity would be discussed later in qualitative analysis. Examples of errors by lexical transfer are given below:

“they want to find a job to *win* more salary”
“And they could *use* their minds better.”

The third type of error examined was grammatical ambiguity. Lombard (1997) claimed that it was the error in “word orders, word forms, prepositions, genitive structures, and so forth that do not conform to the rules of standard written American English” (p. 76). 27 out of 172 (27/172, 16%) errors belong to this category. The following examples demonstrate errors of this type:

“should devote our lives to the jobs we *take up*”
“For them, the job they now *carry on* is the one they favourite”

The fourth type of error was phonological similarity, and it was the least frequently made errors found in the current data. It refers to “confusion encountered by the student at the phone and syllable level in selecting appropriate words found in English language collocations” (Lombard, 1997, p. 68). Only 2 out of 172 (1%) errors were found in this category. The following examples show these two:

“We can *say* many examples around us”
“It will *wind* all our country”

The rest 25 out of 172 (14%) are miscellaneous errors. They were either combinations of other grammatical errors or not belonging to the four main categories. Examples are as follows:

“people still *farm* the farm by hands”
“*inrich* my spare time in my life”

The quantitative analysis in the current study followed the same error categorization by Lombard (1997) as Wu (2005) did. Two interesting results are found from the comparison of the two studies (for comparison, see Table 1 below). First, there are over 70% of errors in Wu (2005) found in Semantic Choice while only 2.96% of the errors found in Lexical Transfer. This might be possibly due to different definitions adopted in the two studies concerning these two categories or due to different criterion for rating. Second, the number of errors in Phonological Similarity is equally small in both studies. This might raise an interesting question why EFL learners whose mother tongue is Chinese rarely made this kind of error. Third, compared with Wu (2005), there are more errors found in grammatical irregularity in the current study (16% vs. 2.96%). It is possibly due to different data source. The data in Wu (2005) are writings from English majors in four academic years in one university while the data in the current study is from an error corpus in which data collected is relatively more diverse in learner proficiency. Thus, there are relatively more grammatical errors found in the current one.

Table 1. Comparison of the results with Wu (2005)

Error type	The current study	Wu (2005)
Semantic Choice	47%	71.69%
Lexical Transfer	22%	2.96%
Grammatical Irregularity	16%	2.53%
Phonological Similarity	1%	1.55%
Miscellaneous	14%	21.27%
Total	100%	100%

Some significant findings from the process of analyzing the data are discussed in the following. The first pattern found is that many wrong verb collocates are still in lexical-semantic relation with the right verb collocate. This finding is consistent with Liu (2002). As demonstrated in the examples below, the right verb for example 1 should be “encounter”, example 2 should be “increase”, example 3 should be “change”, and example 4 should be “adapt to”. In fact, these errors are categorized into the category of semantic choice, and the wrong verb collocates chosen by learners are often synonyms, hypernyms, or troponyms of the correct ones. Grammatically speaking, those verb collocates are correct. It is just the matter of appropriateness. This finding reflects the formulaic nature of language by Wray (2000) in which the sense of native-like is often not rule-based but pre-conceived.

- (1) “for a new job you will *meet* the new difficulties”
- (2) “*enhance* their chance of success”
- (3) “would rather *transfer* from one job to another than be fixed on one job.”
- (4) “have the skill to *fit* the competitive society”

The second finding emerged from the process of categorizing the errors. In some cases it is difficult to decide whether the error belongs to semantic choice or lexical transfer. The research speculated that the ambiguity of distinguishing these two might be the source of discrepancy between the current study and Wu (2005). In example 5, “study knowledge” in Mandarin can be translated into “XUE XI ZHI SHI (學習知識)”. In example 6, “choose my life” in Mandarin is “XUAN ZE REN SHENG (選擇人生)”. In example 8, “catch a success” is “ZHUA ZHU CHENG GONG (抓住成功)”. In those cases, should we attribute students’ errors to confusion of lexical semantic relation or influence of L1? Since the collocations can be translated into Chinese equivalent, sometimes it is hard to define their error source.

- (5) “to *study* more knowledge and accumulate much experiences”
- (6) “How to *choose* my life after graduating is very important to me.”
- (7) “Only by working hard can we *acquire* the triumph”
- (8) “to *catch* a success”
- (9) “can’t *learn* any new knowledge, so they often change their job”
- (10) “can get great achievement and *learn* knowledge”

The third finding is also emerged from the process of categorizing the errors. In certain cases we do not know whether it is wrong noun collocate or verb collocate. As shown in the following example, “reach the effect” is tagged as an error. Should we change it into “achieve the effect” or “reach the goal”? Perhaps it may depend on the examination of the relation between the clause with other clauses.

- (11) “If we do it quickly but don't *reach* the effect, we must do it again in order to reach”

There are several findings which are consistent with previous error analysis on EFL learners, such as overuse of de-lexicalized verbs “make” or “do”. In the present study, 17 out of 81 errors in semantic choice are errors of use of de-lexicalized verbs. It seems to be a common error by Chinese EFL learners. Also, there are recurring errors on the verb collocates of “knowledge” and “ability”. Learners tend to make errors like “show ability” and “learn knowledge”. As mentioned, these errors could either be attributed to confusion of the senses of verb or interferences from L1.

AN collocation

There were 473 adjective-noun collocation errors tagged in CLEC. The results of the study revealed that among these collocation errors, L1 transfer was the most frequent error types. The second was the approximation, and the third was the use of synonym. Table 2 presents the findings.

Table 3. *The frequency of each type of adjective-noun collocation errors*

Error type	Number of errors	Percentage
L1 transfer	136	29%
Approximation	117	25%
Use of synonym	96	20%
False concepts	27	6%
Ignorance of rule restrictions	25	5%
Use of wrong part of speech	25	5%
miscellaneous	47	10%

L1 transfer was the most common adjective-noun errors in CLEC. Examples are presented below:

- (1) In a word, Chinese is an *red language*.
- (2) After a *boiling discussion*, we tell a joke.
- (3) however *fresh hand* he is, he'll win in the end.

In the first example, the learner intended to say that Chinese is a *popular language*; however, he/she used **red language* instead. This erroneous expression might derive from a Chinese expression, DANG HONG, which means popular. Similar phenomenon also occurred in verb-noun collocation errors. Liu (1999) found that students produced errors such as **listen his advice* and **wait your phone call* in verb-noun collocations. Wu (2005) also found that lexical transfer was the main source of errors in verb-noun collocations. This result revealed that students tended to rely on their native language to express themselves when they did not know the target expressions in L2. It was suggested that language teachers should help students aware that different languages may have different expressions. Direct translation strategies might result in unacceptable collocations.

Approximation was a kind of paraphrase in which learners used words that shared similar semantic features with the target expressions. It was also a type of error learners often made in CLEC. Examples are showed below:

- (1) to get the good marks in the *middle exam*
- (2) But if you *health* is very *weak*, lots of things can't be done

(3) I did not want a *lazy job*.

Students used **middle exam* to express the idea of *mid-term exam*. The example might reveal that learners did not know the correct collocation and therefore, used a word which had similar meaning with the target expression. This type of error may result from lack of interaction with native speakers or lack of input. Foreign language learners usually have fewer opportunities to communicate with English native speakers. Thus they have few chances to negotiate meaning with native speakers, which may help learners to acquire appropriate expressions.

In addition to L1 transfer and approximation, use of synonym was also a common error types. Examples are as follow:

- (1) they have a *large knowledge* in a field
- (2) Now I have *large pressure*.
- (3) On the other hand, *great pollution* make fresh water...

Examples showed that students failed to differentiate synonymous words and tended to overuse certain adjectives, such as *large*. Language teachers were suggested to teach students the differences of synonymous words explicitly. One way was to use data from native corpus to help students identify the differences among words (Granger & Tribble, 1998). Teachers could also introduce online concordance tools to students. Using the tools could help learners use words accurately (Li, 2005).

Examples of errors resulting from false concepts are provided below:

- (1) I would surely receive a *delicious scolding* of my parents.
- (2) GUFL enjoys a autiful and *tasteful environment*
- (3) wines made from *industrious alcohol*

In the first example, students used “*delicious*” to modify “*scolding*.” This type of error seemed to indicate that student did not know the correct meaning of words. In order to avoid this type of error, teachers could encourage students to consult dictionaries if they were not sure about the meaning of a certain word.

Ignorance of rule restrictions was a less frequent error types, occurring 25 times in CLEC. Examples were given below:

- (1) the moon and the ground were *perpendicular*
- (2) I have learned *many knowledge* about China
- (3) we shall learn to find fun from *little things* like children

Examples showed that students did not know the usage of some words. For instance, the second and third examples revealed that students did not know many could only be followed by countable nouns while little could only be followed by uncountable nouns. It was suggested that teachers should not only demonstrate the meaning of a word but also teach students how to use it.

Use of wrong was also a less frequent error types in CLEC. Examples are shown below:

- (1) Your kids choose the floor as a *skate field*.
- (2) I read a new book. It is an *interest book*.
- (3) Now, I am a *graduating student*

Students might choose wrong word forms. In the second example, the learner used “*interest*” instead of “*interesting*.” In order to help students, teachers should not only teach word families but also teach students the differences among those related words.

Errors in the category of miscellaneous were errors unable to classify into other categories. This was mainly because of the inappropriate tagging in CLEC. Some errors were tagged as adjective-noun collocation errors; however, those errors did not belong to adjective-noun collocation errors. For example, the line “*we should notice [cc4,-2] the government*” was tagged as cc4 (the tag for adjective-noun collocation error in the corpus). The error should be verb-noun collocations error instead of adjective-noun collocation error. The other kind of inappropriate tagging was the rater of the corpus treated acceptable collocations as unacceptable. For example, “*immense impact*” was error tagged. However, the expression might be acceptable. The expression was actually used by native speakers. Such results revealed that the error tagging in CLEC might not be so accurate.

CONCLUSION AND IMPLICATIONS

The current study analyzed verb-noun and adjective-noun collocation errors from CLEC. The results showed that the majority of errors fall in the categories of semantic choice and lexical transfer, with relatively fewer errors made in grammatical irregularity and phonological similarity. A further investigation of the data reveals that learners’ selections of verb collocates is not at random. The wrong verb collocates chosen are often in lexical semantic relation with the right verb collocates. In terms of adjective-noun collocation errors, the result showed that L1 transfer was the most frequent error types in the corpus. Students tended to use their first language to convey meaning in L2 when they did not know equivalent L2 expressions. Future research might look at errors made by learners of different proficiency levels to further understand learners’ error patterns.

Based on the results of present study, a number of pedagogical implications concerning the vocabulary instruction are proposed. First, in addition to grammatical rules, teachers should also emphasize the fixed and routine aspects of language in class. Second, vocabulary is suggested to be presented to students with a number of their collocates in the example sentences. Or, teachers can demonstrate to students useful online resources such as concordancer as a reference tool for catching the norm of formulaic usages. Third, collocations which are not direct translation from students’ L1 should be paid more attention. Teachers can explicitly direct students’ attention on frequently-made collocational errors. Bahn (1993) suggested that teachers teach collocations that are difficult for learners with a particular L1 background. The result of the present study might be helpful for teachers to decide what collocations should be taught in class.

REFERENCES

- Bahns, J. (1993). Lexical collocations: a contrastive view. *ELT Journal* 47(1), 56-63.
- Benson, M., Benson, E. & Ilson, R. (1986). *The BBI combinatory dictionary of English: A guide to word combinations*. Amsterdam: John Benjamins.
- Carter, R., & McCarthy, M. (1988). *Vocabulary and language teaching*. New York: Longman.
- Chen, P. C. (2002). *A corpus-based study of the collocational errors in the writings of the EFL learners in Taiwan*. Unpublished master’s thesis, National Taiwan Normal University, Taipei, Taiwan.

- Farghal, M., & Obiedat, H. (1995). Collocations: A neglected variable in EFL. *IRAL*, 33(4), 315-333.
- Firth, John Rupert. (1957). *Papers in linguistics 1934-1951*. London: Oxford University Press.
- Granger, S. & Tribble, C. (1998). Learner corpus data in the foreign language classroom: form-focused instruction and data-driven learning. In S. Granger (Ed.), *Learner English on Computer*. London & New York: Addison Wesley Longman.
- Hill, J. (2000). Revising priorities: from grammatical failure to collocational success. In M. Lewis (Ed.), *Teaching collocation: further developments in the lexical approach* (pp. 49-50). London: Language Teaching Publications.
- Hsueh, S. C. (2003). *An analysis of lexical collocational errors in the English compositions of senior high school EFL students*. Unpublished master's thesis, National Kaohsiung Normal University, Kaohsiung, Taiwan.
- Lewis, M. (2000). Language in the lexical approach. In M. Lewis (Ed.), *Teaching collocation: further developments in the lexical approach* (pp. 155-184). London: Language Teaching Publications.
- Liu, C. P. (1999). An analysis of collocational errors in EFL writings. *The proceeding of the eighth international symposium on English teaching* (pp. 483-494). Taipei: The Crane Publishing Co. Ltd.
- Liu, L. E. (2002). *A corpus-based lexical semantic investigation of verb-noun miscollocations in Taiwan learners' English*. Unpublished Master's thesis, Tamkang University, Taipei, January.
- Lombard, R. J. (1997). *Non-native speaker collocations: A corpus-driven characterization from the writing of native speakers of Mandarin*. Unpublished doctoral dissertation, University of Texas at Arlington, Texas.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Tang, Y. T. (2004). *A study of the collocation errors in the oral and written production of the college students in Taiwan*. Unpublished master's thesis, National Taiwan Normal University, Taipei, Taiwan.
- Wu, L. H. (2005). *A study of English verb-noun collocational knowledge of technological university English majors in Taiwan*. Unpublished master's thesis, National Kaohsiung First University of Science and Technology, Kaohsiung, Taiwan.

A Holistic Approach to Fluency: Seeing the Wood for the Trees

Giles Witton-Davies
National Taiwan University
giles@ntu.edu.tw

Oral fluency is a notoriously difficult notion to pin down, and it is often therefore equated with speech rate. However, while fluent speech is usually faster than disfluent speech, it also shows other characteristics: namely, smoothness, efficiency, and rhythm. Research into fluency has demonstrated the role of speed in fluent speech (e.g. Lennon 1990), but has less often been successful in showing how the other features may be assessed. This paper looks at how a global analysis of fluency can be arrived at, by considering not only speech rate, but also pausing patterns, length of run, and repair. Pausing patterns, in particular the location of pauses within or between clauses, contribute to smoothness; the length of runs between pauses is also a sign of how well speech flows; finally, the amount of repair a speaker needs reflects the efficiency of speech.

The data for this study comes from Taiwanese university students majoring in English, using narrative tasks performed in their first and final years at university. The learner speech, ranging from intermediate to advanced level, has been transcribed and analyzed for a variety of measures of fluency, thus allowing the role of each of these to be evaluated. In the end it is the interaction of all of these fluency variables that defines the fluency of each speech sample. The picture of fluency that emerges is one of improvement for those students who were less fluent in the first year, but of stagnation for the more fluent speakers. This calls into question the success of university-level teaching, where perhaps a more holistic teaching approach is required for students to achieve higher levels of fluency.

INTRODUCTION

Aims

This study analyses the oral fluency of a group of university English students and native-speakers of English. After examining the overall scores for typical measures of fluency, and the correlations between these, the study moves on to focus on individual rather than average scores. Fluency measures are considered in detail for two high fluency, one middle and two low fluency students, along with two more and two less fluent native-speakers of English, in order to see what lies behind the impressions of fluency or disfluency that they give. The idea is to see how much students at each level have in common. Are all fluent speakers fluent in the same way, and all disfluent speakers disfluent in the same way? Or are there many different ways to be fluent or disfluent? Another question to be asked is whether these learners improve the oral fluency in their four years at university. Before looking at the data, I will consider the concept of fluency and how it has been defined, and review some previous attempts at measuring it.

What is fluency?

Fillmore (1979) identified four senses of fluency when referring to native-speakers speaking their own language. The first is the ability to speak quickly and without hesitation - to "fill time with talk". The second refers to skill in discussing difficult topics coherently, using complex language. The third is the conversational skill of being able to come up with something appropriate to say in any situation. The fourth is the ability to be creative with

language. The value of these senses of fluency in the context of foreign language learning is limited because they were not designed with the foreign language speaker in mind.

However, consideration of Fillmore's four types of fluency reveals that there is something in common to all four definitions – the idea of speed and lack of hesitation. This is just as important for senses 2-4 as for sense 1, where it is most obvious. A speaker who can speak on difficult subjects, use complex language, and who can think of appropriate, creative or original things to say, is only considered fluent if he/she can do these things at least reasonably quickly. Thus in the end it is speed that is at the root of all four of these types of fluency, but the actual speed will vary according to the context, speaker, and the complexity of what is being said. Fillmore's analysis is perhaps most helpful in the way it shows fluency to be a judgement made of spoken performance that uses different criteria for different speaking tasks and contexts – we do not expect the same kind of fluency from a disc-jockey as from a university professor, or from a social worker as a poet.

Lennon's (1990) discussion of the concept of fluency is more helpful in discussing second language fluency. He talks of two senses of fluency - a broader and a narrower sense. The former is the everyday idea of someone being fluent in a foreign language, where fluent means simply proficient. The latter is the more specialist sense of fluent in language teaching and applied linguistics, where fluency is usually seen as one aspect of speech (although the term may be used with other skills such as reading and writing). Fluency in this sense is considered as distinct from, and potentially orthogonal to, grammatical and lexical accuracy or complexity. Lennon (2000) uses the terms higher-order and lower order fluency to make the same distinction.

Subsequent studies of fluency have generally defined fluency in the narrow, or lower order sense, taking it to relate to temporal aspects of speech such as speed, pausing and repair. This approach is developed by Schmidt (1992) and Towell et al. (1996) into a psycholinguistic view of fluent oral language production, where the basic language skills required for the formulation and articulation of utterances have been proceduralised to allow for efficient and effortless production, allowing attention to be directed to other aspects of speech production and interaction. Fluency here is seen as a psycholinguistic skill lying behind speech production, yet it can still only be observed through the speech that is produced, and by using temporal measures to gauge speed, pausing and repair.

Studies of task performance have looked at the interaction between fluency, accuracy and complexity in spoken production, taking fluency to be represented by one or more of the following variables: the rate of speech; pause time, frequency, and location; and the amount and extent of repair. Tavakoli and Skehan (2005) label fluency in each of these areas speed, breakdown and repair fluency. Here, fluency is seen as competing with accuracy and complexity, with the possibility of trade-offs between these different aspects of speech: a speaker may become more accurate by speaking more slowly, or speak faster by using simpler or less accurate language.

Studies of fluency

Two approaches have been taken to studying fluency, both of them involving a comparison of more and less fluent speakers. One way of doing this is to record students before and after a period of study abroad in an L2-speaking country. The same learner can then be compared at time 1 and time 2 (after the stay abroad), and measurements made of (usually temporal) variables to see which have changed. Lennon (1990), Freed (1995) and Towell et al (1996) took this approach. The other way of distinguishing between fluent and disfluent performance is to have samples rated by a panel of judges, and to analyse only samples that have been agreed by all judges to be of high or low fluency. Riggensbach (1991), Derwing et al (2004) and Kormos and Denes (2004), used this method. The purpose of analysis in both approaches is the same: to discover which measures of fluency (e.g. of the

speed of speech) correlate best with differences in the level of fluency that have been independently established.

Following on logically from the definitions of fluency discussed above, the measures of fluency used in these studies were mainly temporal variables. These include measures of the speed of speech and of pausing. The main non-temporal measures in common use are ones of repair, which address quality of the actual language that is produced. This is important because a rapid stream of nonsense would not be called “fluent” by most people – it has to make sense, and to be produced smoothly. Speed is measured by articulation rate (AR), speech rate (SR) and pruned speech rate (PSR). Pausing is assessed through pause time ratio (along with its converse, phonation time ratio), average length of pause (ALP), mean length of run (MLR), and measures of pause location. Repair is either counted (e.g. number of repetitions or false starts) or the number of words involved in repair can be counted. Some of the most effective measures of fluency are ones that combine two or more aspects of fluency – SR measures speed and pausing, while pruned SR measures speed, pausing and repair.

These studies, despite some variation in results, have shown that SR and MLR are the most consistent indicators of fluency. PTR and ALP correlate with fluency only in some studies, and other variables, such as AR, pause frequency, and filled pauses and repair measures, are not found to be reliable measures of fluency. Thus speed of speech and frequency of pausing seem to be reliable indicators of fluency, while pause time does not correlate with fluency in all studies. Meanwhile, filled pausing and repair have a complex relationship with fluency, and can be viewed in two ways: they may be seen simply as disfluencies that impair the flow of speech, or alternatively they may be perceived as a means of achieving fluency, or at least a way of avoiding excessive silent pausing.

Pruned speech rate was used by Lennon (1990) and Derwing et al. (2004), and found to be highly effective as a measure of fluency. Pruned speech rate is based firstly on articulation rate, but takes into account pause time by dividing words by total time rather than simply speaking time, and also monitors repair by removing all words that are “repaired” (i.e. they are repeated, reformulated or abandoned as false starts) from the total word-count. In this way speakers who rely on excessive amounts of repair or pausing will not score as well as those who repair and pause less while achieving a similar speech rate – differences that AR and SR may not detect, if speech is produced at an equal rate. In view of the proven reliability of pruned speech, and its attractiveness as an all-round measure of pausing, it is perhaps surprising that other studies have not tested this measure.

However, the fact that certain temporal or other variables have not been found to predict fluency in these studies does not necessarily rule out the possibility that they contribute to fluency or disfluency. The failure to detect a relationship between repair and fluency for example, may be due to the way in which it is quantified. The usual way of doing this is to count the number of instances of each kind of repair, but this does not take into account the length of each instance (whether it is one or many words) or the relationship between speed and repair for a particular speaker: it may be that a given person allows for the possibility of more repair as a means of speaking faster, or conversely speaks slowly in order to avoid the need for any repair. Mean scores and correlations between them will not reveal this kind of complex interaction. Only by looking at individual data can such patterns be detected.

Pause location is not treated in most of the above studies, probably because it is more complicated to measure. It is intuitively likely that pausing should be mainly at clause junctures in fluent speech, while there would be more pausing within clauses in disfluent speech. However, as Goldman Eisler (1968) pointed out, this is more true of prepared or read-aloud speech than of spontaneous speech, where she found only 55% of pauses occurred at grammatical junctures. Hawkins, analysing children’s spontaneous narratives, categorised

all pauses according to 47 different locations, in order to see at which locations pauses were most frequent. He found that 66% of pauses, and 75% of pause time, were at clause boundaries. As all speakers, even fluent ones, have to pause, and pausing time may be as much a question of speaking style and personality as of language proficiency, these results suggest the possibility that pause location may be more important than pause time or frequency in assessing fluency.

METHOD

Data

The learner data used here comes from 17 students (11 female and 6 male) at a university English department in Taiwan, and was recorded in the first and final years of their degree courses. All were L1 Mandarin or Taiwanese speakers. The tasks performed were a narrative monologue and a discussion, but here I will focus on the narrative task, where each student produced one story based on a jumbled set of pictures and told it to a partner. Ten minutes of preparation time were allowed after students were given the materials for both tasks. The mean length of narratives was 2 minutes, and the same task was given to each student in year one and in year 4. The original year one recordings came from two classes, comprising 41 different students, but as attendance in year 4 was voluntary, the number of students for whom two recordings were available was only 17. Students worked in pairs in order to make the tasks as communicative as possible, with one group of three in year 4 because there was an odd number of students.

The native-speaker data came from 6 British people aged 20-30 (3 male, 3 female). They were recorded producing two narratives each (of 2 min on average in length) based on the same picture stories used by the students, and their narratives were transcribed and analysed in the same way as the students' data. This made it possible to compare non-native-speaker and native-speaker speech for all the fluency variables used.

Transcriptions were made of the spoken narratives following Foster et al. (2000), and pauses (both filled and unfilled) were measured using a free software programme called Transcriber. This enabled calculation of articulation rate (AR), speech rate (SR), pruned speech rate (PSR), pausing time (PsTR and PTR), pause frequency and mean length of run (MLR), average pause length (ALP), and the percentage of pruned speech. Silent pauses were coded according to their location – using Hawkins three overall categories: between clauses (external pause), within clauses but between grammatical constituents (constituent pause), and within grammatical constituents and so also within clauses (word pause). An additional category of pause was added – repair pauses, which are pauses immediately before a repair. A pause after a word or phrase and its repetition would be a repair pause. Thus pause time and frequency could also be calculated according to location, as well as on an overall basis. Two useful broader categories were defined – external pauses (between clauses) and internal pauses (including constituent, word and repair pauses).

In this paper, the word “pause” will always be used to refer to silent pauses. When filled pauses are mentioned, they will always be termed “filled pauses”. Where I refer to both silent and filled pauses, I will again specify this by referring to “silent and filled pauses”.

Analysis

The first analysis of the data was a quantitative analysis of all temporal variables. Pearson correlations were calculated between all fluency variables in order to see to what extent speed, breakdown and repair fluency (Tavakoli and Skehan 2005) were related. However, care needs to be taken not to read too much into correlations between complex variables which may be tapping the same basic measure. For example, SR and PTR are very likely to correlate because they both depend on pause time. Therefore correlations between “pure” variables, such as AR, were preferred to those between “mixed” variables, such as SR.

The following measures of fluency are used for both quantitative and qualitative analysis:

1. AR or articulation rate: the number of words produced per minute of actual speaking time, with pause time removed. This measures the speed of speech without taking pausing into account.
2. SR or speech rate: the number of words produced per minute of speaking time including pause time. This measures the speed of speech once pausing is taken into account, combining AR with PTR.
3. PSR stands for pruned speech rate. This the speech rate in words per minute once repair has been eliminated. It can be seen as representing a combination of SR, PTR and the percentage of pruned speech (see no. 4 below), or combining measures of speed, pausing and repair.
4. Percentage of pruned speech: the percentage of words remaining after all 'repaired' words (words that are abandoned as false starts, repeated to no rhetorical effect or reformulated) have been removed. This is a measure of the amount and extent of repair.
5. PTR or phonation-time ratio: the percentage of time actually spent speaking, as opposed to pausing. This is the simplest measure of overall speaking and pausing time. The converse of PTR is pause-time ratio.
6. MLR or mean length of run: the mean number of words between silent pauses of over 0.25 sec in length. This shows the average length of chunks of speech that are produced, and so probably indicates how far ahead a speaker can plan the formulation of an utterance. MLR is also a measure of pause frequency, and in fact is equal to the number of words divided by the number of pauses plus one.
7. Pruned MLR: the mean number of pruned words between silent pauses, once repair has been removed.
8. ALP - average length of pause - was calculated by dividing total pause time by the number of pauses.
9. External pauses - pauses between clauses - were scored in several different ways. External pause time was analysed as a proportion of both total time (ePTR or external pause-time-ratio), and pause time (external pause time as a percentage of pause time, ep%pt). External pause frequency was calculated as words per external pause (wpep), while the number of external pauses was divided by the total number of pauses to give the proportion of external pauses (ep%p#).
10. Internal pauses consist of pauses within clauses but between constituents (constituent pauses or cp), within constituents (word pauses or wp), and immediately before repair but after the repaired words – e.g. between a word and its repetition (repair pauses or rp). These were analysed for time and frequency in the same way as external pauses.
11. Filled pauses, or fp, are used to fill time through sounds usually written as er, uh, um, erm, etc. Also included here is laughter (where it takes up measurable amounts of speaking time) and throat clearing noises, which sometimes appear to be used in the same way as filled pauses are.

In addition to the quantitative analysis, a more qualitative analysis was carried out on a representative sample of the data. This was felt to be useful as a means of investigating the variation beneath the averages – how each speaker performed on the different measures and how one aspect of performance interacted with another to produce a particular result. The aim here was to see whether (dis)fluent speakers were basically alike, or whether they achieved their (dis)fluency in different ways. Riggensbach (1991) has claimed that while fluent speakers are alike in their fluency, disfluency is caused by many different weaknesses and problems. The qualitative analysis aims to find out whether this is so, and to investigate the

possible reasons for different speakers' scores on each measure.

Table 1: Pearson Correlations between fluency variables for non native-speaker data, 17 subjects with 2 stories each, 34 stories in total.

		AR	SR	percentPS	PTR	wpp	meanp	epvPtime	epvPno	ipvPtime	ipvPno	wpfp	fPTR	sfPTR
AR	Correlat	1	.732**	.375*	.136	.194	-.219	.241	.356*	-.241	-.356*	.201	-.171	-.220
	Sig.		.000	.029	.444	.272	.214	.171	.039	.171	.039	.255	.332	.211
SR	Correlat	.732**	1	.597**	.770**	.666**	-.744**	.441**	.583**	-.441**	-.583**	-.112	.125	-.772**
	Sig.	.000		.000	.000	.000	.000	.009	.000	.009	.000	.527	.482	.000
percentP S	Correlat	.375*	.597**	1	.528**	.403*	-.507**	.546**	.675**	-.546**	-.675**	.140	-.018	-.582**
	Sig.	.029	.000		.001	.018	.002	.001	.000	.001	.000	.430	.920	.000
PTR	Correlat	.136	.770**	.528**	1	.814**	-.871**	.425*	.537**	-.425*	-.537**	-.334	.362*	-.914**
	Sig.	.444	.000	.001		.000	.000	.012	.001	.012	.001	.054	.035	.000
wpp	Correlat	.194	.666**	.403*	.814**	1	-.484**	.527**	.551**	-.527**	-.551**	-.163	.435*	-.692**
	Sig.	.272	.000	.018	.000		.004	.001	.001	.001	.001	.357	.010	.000
meanp	Correlat	-.219	-.744**	-.507**	-.871**	-.484**	1	-.250	-.437**	.250	.437**	.353*	-.186	.852**
	Sig.	.214	.000	.002	.000	.004		.154	.010	.154	.010	.040	.292	.000
epvPtim e	Correlat	.241	.441**	.546**	.425*	.527**	-.250	1	.903**	-1.000**	-.903**	-.014	.239	-.360*
	Sig.	.171	.009	.001	.012	.001	.154		.000	.000	.000	.936	.174	.036
epvPno	Correlat	.356*	.583**	.675**	.537**	.551**	-.437**	.903**	1	-.903**	-1.00**	-.028	.192	-.497**
	Sig.	.039	.000	.000	.001	.001	.010	.000		.000	.000	.874	.276	.003
ipvPtime	Correlat	-.241	-.441**	-.546**	-.425*	-.527**	.250	-1.000**	-.903**	1	.903**	.014	-.239	.360*
	Sig.	.171	.009	.001	.012	.001	.154	.000	.000		.000	.936	.174	.036
ipvPno	Correlat	-.356*	-.583**	-.675**	-.537**	-.551**	.437**	-.903**	-1.000**	.903**	1	.028	-.192	.497**
	Sig.	.039	.000	.000	.001	.001	.010	.000	.000	.000		.874	.276	.003
wpfp	Correlat	.201	-.112	.140	-.334	-.163	.353*	-.014	-.028	.014	.028	1	-.648**	.076
	Sig.	.255	.527	.430	.054	.357	.040	.936	.874	.936	.874		.000	.669
fPTR	Correlat	-.171	.125	-.018	.362*	.435*	-.186	.239	.192	-.239	-.192	-.648**	1	.034
	Sig.	.332	.482	.920	.035	.010	.292	.174	.276	.174	.276	.000		.847
sfPTR	Correlat	-.220	-.772**	-.582**	-.914**	-.692**	.852**	-.360*	-.497**	.360*	.497**	.076	.034	1
	Sig.	.211	.000	.000	.000	.000	.000	.036	.003	.036	.003	.669	.847	

** Correlat is significant at the 0.01 level .

* Correlat is significant at the 0.05 level .

RESULTS

Correlations between fluency measures

Table 1 shows Pearson correlations between the fluency measures. Initially, as stated previously, it is best to focus on correlations between “pure” measures, which include the following: AR for speed, PTR for pause time, MLR for pause frequency, and percentage of pruned speech for repair.

AR has only a small correlation with the percentage of pruned words (pw%cw), which is a measure of repair. This suggests that faster speakers tend to use less repair, perhaps because their knowledge of the language is more proceduralised and so they can formulate their utterances without hesitation (as hypothesised by Schmidt 1992 and Towell et al 1996). Less fluent speakers, in contrast, may need to think about their grammar and lexis and modify it as they go along, often in mid-clause. AR does not correlate significantly with PTR or MLR, showing that speed of speech is largely independent of the amount and frequency of pausing. In other words, some fast speakers may pause more, while some pause less, with slow speakers showing similar variation. AR does correlate strongly with SR, but this is clearly circular – SR is based on AR combined with PTR, so such a correlation is unsurprising.

PTR and MLR correlate strongly with each other, as is to be expected from two measures of pausing - MLR measuring pause frequency, and PTR measuring pause time. They also, and less predictably, have a medium sized correlation with the percentage of pruned speech, indicating that speakers who pause more generally tend to rely more on repair. Repair and pausing are independent features of speech, so this correlation is interesting. However, it may simply be an indication that more proficient speakers both pause less and repair less, and it may still be that for individual speakers there is a trade-off between repair and pausing – to discover this it would be necessary to compare samples by the same speaker at the same stage, but under differing task conditions.

Pause location is more difficult to measure than pause length, time or frequency. One measure of it is words-per-internal pause, which is a measure of pause frequency for pauses that occur within, rather than between, clauses. Another marker of pause location is the percentage of internal pauses (as a percentage of all pauses), a measure which reflects the ability of a speaker to formulate whole clauses at a time – a typical feature of fluent speech according to Pawley and Syder (2000).

A third gauge of pause location is the external and internal pause time ratio (ePTR and iPTR), which give the proportion of total time taken up by either external or internal pausing. Unlike the percentage of external or internal pause time measures mentioned above, these two measures are independent of each other. And while iPTR is partly dependent on repair (through repair pauses), ePTR is not.

Correlating any of these measures of internal pausing with measures of repair could be misleading because one kind of internal pause is the repair pause, so more repair will tend to affect both measures – increasing repair pause frequency and decreasing the percentage of pruned speech. It is nevertheless interesting to see that the correlation of all the main fluency variables (AR, SR, PTR, MLR) with the percentage of external pausing is positive, while the correlation with percentage of internal pausing is negative. As those measures, when higher, generally indicate fluency, the correlation suggests that, in general, a greater proportion of internal pausing means less fluency, while a higher proportion of external pausing indicates more fluency.

One other correlation of interest is the small-to-medium sized one between filled pausing (filled pause ratio or fPTR) and two measures of silent pausing – PTR and MLR. The interest lies in the positive nature of the correlation – indicating that more filled pausing means less silent pausing (both overall pause time and pause frequency). This shows that speakers who produce more filled pauses do not also have more (silent) pauses, as might

easily be the case, but rather they tend to pause less. In other words, filled pauses can serve as a means by which speakers keep talking and avoid the need to pause (silently). In a way this is a straightforward exchange of one kind of pause (a silent pause) for another (a filled pause). But the point to notice is that it could have been otherwise. Silent pauses could have increased together with filled pauses, or there could have been no relationship between the two types of pause.

Overall scores

Tables 2 and 3 show the scores for five learners and four native-speakers. In table 2, NS1 and NS2 are the most fluent, while NS3 and NS4 are the least fluent, native-speakers. In table 3, S3 and S4 are the least fluent non-native-speakers, while S11 and S15 are the most fluent, with S13 somewhere in between. In keeping with the results for the complete data, it can be seen even from this more limited sample that the native-speakers score higher than the learners, and the fluent learners score higher than the less fluent ones, on average and for most measures. This is true for AR, SR, percentage of pruned speech and filled-pause-time ratio. However, for all of these measures, some fluent learners perform better than some native-speakers, showing that there is a large overlap between native-speaker and fluent non-native-speaker performance.

The biggest differences are between measures of the percentage of internal pauses and internal pause time. Learner averages are more than double those of native speakers, while fluent learners pause less internally than the more disfluent learners. Even the less fluent native speakers pause less internally than the more fluent non-native speakers, providing one possible means of discriminating between these two groups, who otherwise overlap on all other fluency measures.

Equally interesting is the fact that scores on certain measures - PTR, ALP, ePTR – are very similar for both native speakers and non-native speakers, and the mean value of silent-plus-filled pause-time ratio is equal for the two groups. This means that fluent speakers may pause as much as less fluent speakers, and that overall pause time is likely to be equal for the two groups. The difference with fluent speech is that pausing is less frequent (so runs are longer), and internal pausing makes up a greater proportion of pauses and pause time in relation to external pausing. The lack of a clear difference for PTR and silent-plus-filled PTR is particularly surprising, as table 1 shows a strong correlation between PTR and SR. Moreover, SR does seem to be an indicator of fluency as is shown by the higher mean for native-speakers and fluent learners as compared non-native-speakers and less fluent learners. The problem, presumably, with such apparent contradictions, is that each case is different, and averages cover up much of the individual variation along with its causes and possible explanations. More pausing does, in some instances, mean less fluency, while in other cases fluency can be maintained by compensating for a high rate of pausing through fluency in other aspects of speech - e.g. with a rapid AR, or a slow, deliberate style of speech with plenty of external pausing but much less internal pausing.

It is therefore necessary to look at the data of individual cases in order to see whether what is suggested by the overall data is borne out on the level of the individual. It may also be possible, on the basis of individual data, to infer why each speaker performs in the ways/he does. This may include reasons not related to language ability or proficiency, but rather to context and the speaker's interpretation of the task.

Table 3: Native-speaker data.

name	AR	pAR	SR	PSR	pw%cw	fPTR	sfPTR	PsTR	PTR	MLR	pMLR	%ip#	%ipt
NS1a	258.2	206.5	198.9	159.1	80.0	7.33	30.3	22.9	77.06	10.69	8.55	22.4	20
NS1b	233.7	195.0	172.8	144.2	83.4	7.02	33.0	26.0	73.97	6.49	5.42	26.5	26.8
NS2a	229.6	213.4	174.2	161.9	92.9	2.19	26.3	24.1	75.89	8.48	7.88	10.2	16.7
NS2b	224.1	214.5	166.6	159.5	95.7	1.05	26.7	25.7	74.34	6.59	6.31	15.5	16.1
NS3a	173.2	165.1	107.5	102.5	95.3	3.33	41.3	37.9	62.08	4.03	3.84	22.7	35.5
NS3b	159.8	152.0	102.7	97.6	95.1	3.98	39.8	35.8	64.22	3.74	3.55	34.7	40.5
NS4a	193.5	176.1	103.4	94.1	91.0	0.00	46.6	46.6	53.45	4.17	3.79	32.1	45.5
NS4b	206.0	203.0	125.5	123.6	98.6	0.00	39.1	39.1	60.91	6.27	6.18	10.0	20
mean	209.8	190.7	144.0	130.3	91.5	3.11	35.4	32.3	67.7	6.31	5.69	21.77	27.6

Table 4: Learner Data

	AR	R	pA	SR	pSR	pw%cw	fPTR	sfPsTR	PsTR	PTR	MLR	pMLR	%ip#	%ipt
S4A	155.4	8	78.	89.8	45.6	50.7	7.6	49. 79	42.2	57.81	3.24	1.53	65.8	69.4
S4B	185.7	117.9		1282	81.4	63.5	3.4	34.40	31.0	69.01	3.97	2.54	59.6	60.3
S5A	170.8	117.6		81.1	55.8	68.8	0.2	52.72	52.5	47.46	3.12	2.06	48.2	56.3
S5B	172.6	127.0		104.3	76.7	73.6	6.2	45.82	39.6	60.41	3.24	2.36	66.3	61.8
S11A	199.0	184.8		140.7	130.7	92.9	0.2	29.524	29.3	70.71	4.83	4.46	31.7	37.5
S11B	180.2	166.9		135.5	125.5	92.6	3.5	28.333	24.8	75.15	4.84	4.31	46.3	41.9
S15A	187.6	162.2		150.3	129.9	86.4	5.4	25.258	19.9	80.10	6.23	5.38	28.5	31.6
S15B	169.0	135.6		128.8	103.3	80.3	7.1	30.903	23.8	76.18	4.90	3.94	36.7	41.9
S13A	139.1	106.8		124.6	95.7	76.8	11.7	22.098	10.4	89.58	8.46	6.51	28.9	34.4
S13B	148.4	117.4		112.1	88.7	79.1	9.3	33.75	24.5	75.51	4.20	3.37	61.0	61.0
mean	170.8	131.5		119.5	93.3	76.5	5.5	35.3	29.8	70.2	4.7	3.6	47.3	49.6

AR = articulation rate, pAR =pruned AR, SR =speech rate, pSR =pruned SR, pw%cw =percent of pruned speech, fPTR =filled pause-time ratio, sfPTR =all pause time ratio, PsTR =pause time ratio, PTR =phonation time ratio, MLR =mean length of run, pMLR =pruned mean length of run, %ep# =external pauses as % of all pauses, %ept =external pause as % of all pause time

Most fluent learners in year 1

The most fluent learners in year 1 did not improve their fluency over the four years at university, in so far as can be seen from their narratives. In the case of S15, there was a slight but definite decline on all measures of fluency, including speed, repair and pausing. In the case of S11, the decline was less marked and limited to speed (AR, SR and PSR) while MLR and repair rate (percentage of pruned speech) remained stable and PTR actually increased, with less pausing although there is an increased internal pausing ratio.

S15 was the fastest speaker in year one, and in the recordings she sounds as fluent as the temporal measures suggest. The apparently reduced fluency in year 4 should be seen in this context – she was already very fluent, with not very much scope for further improvement in terms of purely temporal fluency, at the beginning of her university career, so even with the decline she is still of well above average fluency. She seems to be speaking more carefully in year 4, and to be more concerned to make her narrative interesting and expressive. She may be a case where what Lennon (2000) calls “higher order fluency” would be more appropriate to judge her by than “lower order” fluency measures. This would mean taking the complexity and sophistication of her language into account. However, this is beyond the scope of the present study, and we have to recognise that in purely temporal terms, and on the basis of this evidence, this student has not improved, but rather regressed slightly.

S11 was a fast speaker in year 1, who spoke very naturally and spontaneously, with little apparent effort at adding much detail or sophistication to her narrative. Her ability to speak rapidly and effortlessly made her seem highly fluent, and in year 4 her performance changed less than that of S15. Her speed was somewhat slower (AR, SR and PSR were all down), but she paused less (though silent-plus-filled pause-time is the same). Both MLR and repair rate improved. This student was an L1 bilingual speaker of Chinese and Japanese, whose formal and written Chinese was more limited before coming to university. She said after her year 4 interview that she felt her English had deteriorated while at university, while her Chinese had improved. In fact, she seems to have remained more-or-less at the same level of fluency.

S13 was an unusual case in that she had an extremely high PTR rate, with very little silent pausing, while her AR and SR were slightly above average. In year 1, her MLR was higher than that of most of the native speakers, at a mean of 8.46 words per run. It seems that this was partly achieved through relying on high rates of repair and filled pausing – measures for both are above average. Yest even her pruned MLR (with filled pauses and repairs removed) was higher than for any other learner, at 6.51 words per run. It appears that this student simply did not need to pause much while speaking. Her personality was very lively, and my impression is that she spoke in the same way in Chinese, while somehow she managed to transfer her speaking style into English.

In year 4, S13 did not perform at the same level, pausing more (though with fewer filled pauses) and producing much shorter runs (MLR= 4.20, pruned MLR = 3.37). As a result, while her AR was higher, her SR and pruned SR were lower. In both narratives, S13’s range and accuracy of grammar and lexis are limited, well below the level of S11 and S15. It may be that her year 1 performance was the result of effective use of preparation time, while in year one there was thought to be no particular need to perform well.

Least fluent learners at T1

The least fluent speakers in year 1 were the ones who made the most progress in terms of the fluency measures used here. However, it is interesting to see both where

they were most disfluent in year 1, and where they improved most over their four years at university.

S3 had a very high repair rate high in year 1, with about half of his words lost in repair of one kind or another. This is the feature above all others that made this speaker disfluent. AR seems reasonably high, SR is lower, but it is in pruned SR that we see the true level of disfluency displayed. PTR was high, but it was the frequency of pausing that was more significant here, reducing the flow with an MLR of 3.24 and pruned MLR of 2.54. This is indicative of an inability to plan ahead, and a tendency to plan one word at a time. As suggested by Pawley and Syder (2000), fluency comes with the ability to formulate whole clauses at a time. When a speaker can only think one or two words ahead, and when corrections and repetitions are frequent, speech will never be fluent. S3 improved on every measure in year 4, the overall change perhaps being best reflected in the almost doubling of pruned speech rate (from 45.6 to 81.4 words per minute).

S4 also gave an impression of disfluency, but the measures reveal a rather different picture. In his case, measures of speech rate are much higher than for S3, and his AR in year 1 was above average. Nor was the repair rate high, as it was for S3. The problem in his case was the occurrence of frequent and long pauses, as reflected by a low PTR (more than half the time was spent pausing) and also a low MLR of 3.12 words per run. It seems this learner thought long and deeply about each utterance, yet was only able to produce one to three words each time he spoke before more planning time was required.

High fluency native-speakers

Just as a closer look at learner data is revealing, so it is in the case of the native speaker performances. Not all of the native speakers spoke rapidly or even fluently on other measures such as those of pausing and MLR.

NS1 was one of the fastest speakers in terms of both AR and SR in all of the native and non-native speaker narratives. Her MLR was also high, especially for the first narrative. However, pruned AR and pruned SR are both considerably lower, though still high. It seems that this person achieved her high speed of speech through a style of speech that allowed for plenty of self correction through repair. Her repair rate was the highest for all native speakers, and it seems no coincidence that high speech rate and high repair rate go together here. Avoiding repair would have required slower, more measured speech. This speaker was more concerned with keeping going with the story than with accuracy and economy of words. In fact, she made her story very colourful and entertaining, and gave the impression that she was chatting to friends and relating an anecdote, as she might have done in a coffee shop or pub.

NS2 is another fluent speaker, even by native speaker standards. He has a slower AR and SR, but higher pruned AR and SR than NS1, because he relies less on repair. The rate of pruned speech is very high, with less than 10% of words “wasted” through repetitions, false starts and reformulations. This is especially impressive in view of the high speech rate. MLR also high, and the overall performance is of a similar or even higher level of fluency to that of NS1. These narratives are suggestive of a dramatic theatre performance, aimed at entertaining the audience and maintaining their interest.

Low fluency native speakers

The rather slow, halting performance by two native speakers was initially rather surprising, as I had expected all their narratives to be more fluent on most measures than all but the most fluent learners. The fact that this was not the case may of course mean simply that not all native-speakers are equally fluent in their own L1. However, it may also be a sign that they are interpreting and performing their task in different

ways, and speaking in a style appropriate for that perceived task. These slower native-speakers may have spoken as they did as a result of a completely different interpretation of the narrative task. If this is the case, these narratives do not reflect each speaker's maximum fluency, and therefore the same may also be true of learner performances.

NS3 spoke slowly, with a low AR and SR, and with frequent and long pauses. Her strength, which partly compensated for his, was her almost total lack of repair and filled pauses (only 5% of her speech was lost through repair). This, I believe, is highly revealing, and shows that this speaker was tackling this narrative task in a very different way to NS1 and NS2. Rather than seeking to entertain and be expressive, NS3 was more concerned with accuracy and economy of language, seeking to relate the story with as much precision as possible. Where NS1 seemed to be chatting with friends, and NS2 appeared to be performing, NS3 narrated the story as someone might speak when giving evidence in court.

NS4 was similar to NS3, with a very slow, careful delivery. His second story achieved a pruned speech rate of 98.6%, meaning there was almost no repair or filled pausing. Curiously this narrative was rather faster than the first, but also shorter. Both stories were of below average speed, and MLR was higher in the second than the first. S4, to continue the comparisons with real life contexts, appeared to be taking a language test, where errors were to be avoided at all costs, while the quality of the narrative in terms of complexity, sophistication or entertainment was not an issue.

These interpretations of the native speaker performances are speculative, but the results themselves suggest some such interpretation is required, if we are to understand such wide variations in performance between L1 English speakers.

DISCUSSION

The lesson to be drawn from this study, and in particular from the qualitative review of individual fluency measures, is that we should avoid hasty over-generalisations about what contributes to or detracts from fluency. Fluent speakers are fluent in a variety of ways, and disfluent speakers similarly show their disfluency differently one from another.

Fast speakers will often be judged as fluent, but not necessarily – not, for example, if the speech is repetitive, and full of reformulations and false starts. At the same time, slow speakers may also be judged fluent, if slow speech is deemed appropriate for the context, if the speaker expresses herself with precision, and if pauses appear to be used for the purpose of thought and improved self-expression rather than due to problems with basic grammar and vocabulary. There are different ways to be fluent – fast speech being one, but speaking with a regular, steady rhythm being another. A third way is to produce short, rapid runs with long pauses in between, while a fourth is to produce slow but elegant and well chosen speech.

Similarly, we have seen that disfluent speakers may be disfluent because they speak slowly (in terms of AR) but that it is just as likely to be because they pause excessively, and in particular because these pauses occur too often within clauses; or they may keep pausing for short pauses within clauses and within phrases, thus interrupting what would normally be the chunks typical of spoken language; or else lack of a grasp of basic grammar and vocabulary may lead to continuous false starts, repetitions and reformulations, thus making smooth speech impossible.

It is of great interest to observe that native-speakers do not always out-perform non-native-speaker in tasks of the kind employed here. The low scores on most fluency measures achieved by NS3 and NS4 should sensitise teachers to the

possibility that students may sometimes perform below their competence, not because of any inability to speak more fluently, but because they decide to attach more importance to other considerations that limit fluency.

As for progress made at university, the learners in the sample studied here do not seem to have improved greatly over their four years at university. This may be because there are fewer language classes in the last two years of their degree course, with more emphasis being placed on the study of literature. One possible remedy, if a remedy is considered desirable, might therefore be to schedule more language skill classes in years 3 and 4. Another, more radical suggestion would be to encourage, or even require, students to spend one year of their course abroad, in an English speaking university. This is usual in the UK for foreign language students, and a year of immersion in the L2 environment would almost certainly be more effective than adding extra speaking classes here in Taiwan. Indeed, the many students who participate in exchange programmes with foreign universities are doing just that, and seem to benefit greatly from the experience.

REFERENCES

- Derwing, T., Rossiter, M., Munro, M. and Thomson, R. (2004). "Second language fluency: judgements on different tasks." *Language Learning* 54.4: 655-679.
- Fillmore, C. (1979). "On Fluency. Reprinted in In Rigggenbach, H. (ed.). *Perspectives on Fluency*. University of Michigan: Ann Arbor,. 43-60.
- Foster, P. and Tonkyn, A. and Wigglesworth, G. (2000). "Measuring spoken language: a unit for all reasons. *Applied Linguistics*, 21.3: 354-375.
- Freed, B. (1995). "What makes us think that students who study abroad become fluent?" In B. Freed (Ed.) *Second language acquisition in a study abroad context*. "Philadelphia, PA: John Benjamins, 123-148
- Goldman Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. London: Academic Press.
- Kormos, J and Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System* 32: 143-164.
- Lennon, P. (1990). "Investigating fluency in EFL: a quantitative approach." *Language Learning* 40.3: 387-417.
- Pawley, A. and Syder, F. (2000). The one-clause-at-a-time hypothesis. In H. Rigggenbach (ed.) *Persepectives on Fluency* (pp. 163-199). Ann Arbor: The University of Michigan Press.
- Rigggenbach, H. (1991). Towards an understanding of fluency: a microanalysis of non-native speaker conversation. *Discourse processes* 14: 423-441.
- Schmidt, R. (1992). "Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition* 14.4: 357-85.
- Tavakoli, P. and Skehan, P. (2005). 'Strategic planning, task structure, and performance testing.' In R. Ellis (ed.) *Planning and Task performance in a second language*, Amsterdam: John Benjamin, 239-273.
- Towell, R., Hawkins, R, and Bazergui, N. (1996) . "The development of fluency in advanced learners of French. *Applied Linguistics* 17.1: 84-119.

Comic Strip as a Supplement of Reading: Effects of Presentation Sequence¹

Hui-ju Wu (吳蕙如), Hsi-chin Chu (朱錫琴)

Changtai Junior High School, Changhua County
National Taiwan Normal University

wu.huiju@msa.hinet.net

hcchu@ntntu.edu.tw

This study investigates the effects of presentation sequence of comic strips on text memory and inference generation in the recall of junior high school students in Taiwan. The participants of this study were four classes of junior high ninth graders. One class was assigned as Control Group, receiving no comic strips when reading a 237-word narrative. The other three classes formed as three Experimental Groups, reading the same narrative text with comic strips presented in different sequencing: Pre-reading Group, During-reading Group, and Post-reading Group. After reading, students wrote a recall of the story.

The recall data were coded in three dimensions: recall of idea units, text-based inference and reader based inference. ANOVA analyses comparing group differences on the three dependent measure yielded the following findings. First, for recall, only During-reading Group performed significantly better than Control Group. Besides, During-reading Group also recalled significantly more story content than Post-reading Group. This indicated the positive effect of dual coding of text and imagery on memory. Second, for reader-based inference, all three Experimental Groups outperformed Control Group although no significant difference was found among the three groups. It is postulated that supplying visual before, during, or after reading text may help readers to expand the text representation by integrating personal experience/knowledge. Third, for text-based inference, again all three Experimental Groups made significantly more units than Control Group. However, among the three groups, Pre-reading Group and Post-reading Group generated more text-based inferences than During-reading Group. Dual coding might leave no mental space for higher-level inference generation.

INTRODUCTION

In EFL reading material design, many kinds of text supports, text, visual, or acoustic, are proposed to motivate readers and facilitate their reading comprehension. Among the types of support, visual aids are commonly used, especially at elementary and secondary level. Visuals are everywhere in the EFL elementary and secondary teaching and test materials. The pervasiveness of such issue is self-evident.

Psycholinguistic theories in text processing, those of transmediation (Siegel, 1995), repetition hypothesis (Gyselinck & Tardieu, 1999), and dual coding (Paivio,

¹ This paper is drawn from the M.A. thesis of the first author, which was modified by the second author, the supervisor of the thesis, into the present version.

1986; Sadoski & Paivio, 2001) also address the potential contribution of visuals to text understanding. Empirical evidence speaks to us that when visuals are relevant to the verbal text, mental referential connections between text and visual modes will facilitate comprehension and memory because important information is processed twice (Glenberg & Langston, 1992; Gyselinck & Tardieu, 1999; Mayer & Sims, 1994; Moore & Scevak, 1997).

Yet in L2 research, few studies are done to corroborate the proclaim. How the visual aids, such as comic strips, play a role in EFL reading comprehension is scarcely investigated. To date, Liu (2005) found its impact on Low-proficiency ESL readers in their text memory when reading difficult text. Still, whether the effect of presenting comic strips along with text would hold for EFL readers and whether the sequencing of its presentation, before, during, or after reading text affects reading comprehension has yet to be explored.

Thus, the understanding of the issue of whether supplementing texts with visuals, such as comic strips, and of whether the presentation of such should be given prior to, during, or after reading might assist practitioners in field in designing teaching materials, instructional procedure, and test materials. It is with these aims in mind, this study was designed and conducted. In addition to test their the effects on text memory, as is commonly investigated in previous studies, the effects on another two levels of comprehension outcome, text-based inference and reader-based inference, are explored as well to learn whether the visual presentation would boost higher level reading. Specifically, the research questions are:

1. What is the effect of the sequencing of comic strips presentation on EFL students' reading comprehension as indicated by recall?
2. What is the effect of comic strips on EFL students' inference generation as indicated by recall?
3. What is the effect of the sequencing of comic strips presentation on EFL students' inference generation as indicated by recall?

METHODS

Selecting Texts with Comic Strips

To select suitable texts for the participants, four steps were taken.

First, two narrative texts with comic strips were chosen from a commercialized readers. The readability of these texts (3.1 and 3.3 on Flesch-Kincaid Grade Level) fall within the range of ninth graders' current textbook readings: 2.6 to 6.8. Second, these texts were further modified slightly on proper names, tense, sentence structure to make the texts more natural, comprehensible and coherent, with the length and readability unchanged.

Finally, to make sure that students had no previous exposure to the content and to determine a text with the difficulty level suits target students' comprehension ability, a pilot study was conducted on two classes of 38 ninth graders other than the experimental participants. All students claimed they have never been exposed to the two texts before, and rated the difficulty of the texts, 2.8 and 1.5 (1: too difficult to 5: too easy) respectively. The former text, 237 words in length, and 3.3 in readability, was therefore chosen. The comic strip consisted of nine pictures, without verbal stimulus. Further examination by four English teachers based on four criteria confirmed their relevance to the text.

Participants

144 junior high ninth-graders in four classes (36-38 each) participated in this study. Each of the four classes was assigned to Control group, Pre-reading group,

During-reading group and Post-reading group respectively. They were homogeneous in language proficiency, which was supported by ANOVA analysis on three Mock Basic Competence English Tests indicating no difference among groups, $F(3, 145) = .370, p > .05$.

Data Collection Procedures

The experiment was conducted in one class period. As shown in Table 1, Control group read the text without the comic strip for ten minutes while the other three Experimental groups read the text with comic strip but with different sequencings of its presentation in the meantime. Both Pre-reading and post-reading groups read the two materials separately while During-reading group read with two materials presented simultaneously.

Table 1. *The Experiment Procedure*

	Control	Pre	During	Post
Reading Phase	Text-only 10 mins.	(1)Comic-3mins. (2)Text-7mins.	Text + Comic 10 mins.	(1) Text-7 mins. (2) Comic-3mins.
Questionnaire Phase	5 mins.	5 mins.	5 mins.	5 mins.
Recall Phase	15 mins.	15 mins.	15 mins.	15 mins.

Note. Control = Control Group; Pre = Pre-reading Group;
During = During-reading group; Post = Post-reading Group

After the reading session, students filled in a questionnaire as a means to disrupt their short-term memory for text before their performed recall in Chinese.

Scoring

Scoring of the recall data involves two rounds. First, pausal unit system suggested by Johnson (1970) and Bernhardt (1991) was employed to score for memory for text. After a template was established, two raters scored the data for those units matching the template independently, yielding an inter-rater reliability of .97.

Second, the two raters gleaned for the unmatched units and rated for text-based inferences – the meaning made to fill in the textual gaps, and reader-based inferences – the meaning drawn from readers' knowledge and experience to reason (Barry & Larzate, 1998; Chu, 2002). The inter-rater reliabilities for text-based and reader-based inferences were .88 and .92 respectively.

FINDINGS

Effects on Recall

ANOVA analysis was performed on the percentage of recalled units to compare the four groups. Table 2 shows the means recall score of the four groups. There was a significant difference among the four groups on the written recall, $F(3, 140) = 3.891, p < .05$

Table 2. *Descriptive Statistics for the Written Recall of Four Groups*

	GROUP (N = 144)			
	Control (n = 35)	Pre (n = 35)	During (n = 37)	Post (n = 37)
Mean	15.20 (11.27)	19.08 (9.7)	23.08 (9.19)	18.29 (9.22)

Note. 1. Control = Control Group; Pre = Pre-reading Group; During = During-reading group;
Post = Post-reading Group
2. Standard Deviations are presented in the parenthesis.

Since there was an overall difference, post hoc pair-wise analysis was thus performed. The results (Table 3) indicated that the major source of significance comes from the difference between two pairs: During Group vs. Control Group, and During Group vs. Post Group. Thus comic strips presentation makes a difference in text memory only when it is read along with text. To present comic strips after text processing, the impact is similar to that on Control group, i.e., not presentation at all.

Table 3. *LSD Pairwise Comparison for the Written Recall of Four Groups*

	GROUP (N=144)			
	Control	Pre.	During	Post.
Control (n = 35)		-3.88 (2.36)	-7.88 (2.33)*	-3.09 (2.33)
Pre. (n = 35)	3.88 (2.36)		-3.99 (2.33)	.78 (2.33)
During (n = 37)	7.88 (2.33)*	3.99 (2.33)		4.78 (2.29)*
Post. (n = 37)	3.09 (2.33)	-.78 (2.33)	-4.78 (2.29)*	

* $p < .05$

Note. 1. Control = Control Group; Pre = Pre-reading Group; During = During-reading group;
Post = Post-reading Group
2. Standard Deviations are presented in the parenthesis.

Effects on Text-based Inference

ANOVA analysis was performed on the raw units tabulated as text-based inference in the recall data to compare the four groups. Table 4 shows the mean number of inference for the four groups. There was a significant difference among four groups, $F(3, 140) = 16.20, p < .005$.

Table 4. *Descriptive Statistics for the Text-Based Inference of the Four Groups*

	GROUP(N = 144)			
	Control (n = 35)	Pre (n = 35)	During (n = 37)	Post (n = 37)
Mean	.45 (.65)	1.82 (1.18)	1.18 (.98)	2.04 (1.27)

Note. 1. Control = Control Group; Pre = Pre-reading Group;
During = During-reading group; Post = Post-reading Group
2. Standard Deviations are presented in the parenthesis.

Post hoc comparison (Table 5) indicated that the text-based inferences generated by each of the three experimental groups differs from those by the Control Group, indicating that as long as comic strips is presented, regardless of the sequence, students would generated more text-based inferences.

In addition, Post Group and Pre Group also significantly outperformed During Group in the generation of text-based inference, indicating that students generate more text-base inferences when shown comic strips without text.

Table 5. *LSD Pairwise Comparison for the Text-Based Inference of Four Groups*

	GROUP (N = 144)			
	Control	Pre.	During	Post.
Control (n = 35)		-1.37 (.25)*	-.73 (.24)*	-1.58 (.24)*
Pre. (n = 35)	1.37 (.25)*		.63 (.24)*	-.21 (.24)
During (n = 37)	.73 (.24)*	-.63 (.24)*		-.85 (.24)*
Post. (n = 37)	1.58 (.24)*	.21 (.24)	.85 (.24)*	

* $p < .05$

Note. 1. Control = Control Group; Pre = Pre-reading Group; During = During-reading group;

Post = Post-reading Group

2. Standard Deviations are presented in the parenthesis.

Effect of Comic Strips on Reader-based Inference

ANOVA analysis was performed on the raw units tabulated as reader-based inference in the recall data to compare the four groups. Table 6 shows the mean number of inference for the four groups. There was a significant difference among four groups, $F(3, 140) = 5.721, p < .05$.

Table 6. *Descriptive Statistics for the Reader-Based Inference of the Four Groups*

	GROUP(N = 144)			
	Control (n = 35)	Pre (n = 35)	During (n = 37)	Post (n = 37)
Mean	.42 (.55)	1.27 (1.46)	1.22 (.99)	1.36 (1.10)

Note. 1. Control = Control Group; Pre = Pre-reading Group; During = During-reading group;

Post = Post-reading Group

2. Standard Deviations are presented in the parenthesis.

Follow-up pair-wise comparisons among the four groups are presented in Table 7, which indicated that all three experimental groups outperformed the Control Group in the production of reader-based inferences. Thus, students generated significantly more reader-based inferences when reading with comic strips than without, no matter at which temporal order the comic strip was presented.

Table 7. *LSD Pairwise Comparison for the Reader-Based Inference of Four Groups*

	GROUP (N=144)			
	Control	Pre	During	Post
Control (n = 35)		-.84 (.25)*	-.80 (.25)*	-.93 (.25)*
Pre (n = 35)	.84 (.25)*		.04 (.25)	-.09 (.25)
During (n = 37)	.80 (.25)*	-.04 (.25)		-.13 (.25)
Post (n = 37)	.93 (.25)*	.09 (.25)	.13 (.25)	

* $p < .05$

CONCLUSION

To sum up, effects of the comic strip and its sequencing of presentation have been shown on students' recall and generation of text-based and reader-based inferences.

First, for recall, only students in During-reading Group outperformed students in Control Group and Post-reading Group. To boost memory for text, comic strips need to be read with the presence of text.

Second, for both types of inference, all three Experimental Groups generated more than Control Group. This shows that supplementing the text with comic strips might help the generation of inferences boosting higher level of reading, be it presented prior to, during, or after reading.

Third, both Pre-reading and Post-reading Groups performed better than During-reading Group in the generation of text-based inference. It is conjectured that the presence of comic strips and text at the same time might inhibit the generation of text-based inference. This finding coupled with the first finding that comic strips presented during reading may increase text memory may be interpreted in terms of capacity constraint in working memory. Processing comic strips and text simultaneously may enhance memory, attesting the effect of dual coding on the one hand; on the other hand, limited space is allocated to the higher-level processing, inference making.

In conclusion, comic strips presented before reading might help activate relevant schema for later text-base processing producing a rich model of text processing; those presented after reading might extend the text processing after reading is done. In other words, readers continue to expand already stored story content by anchoring the established schema, generating an elaborated text representation. Yet, simultaneous provision of comic strips during reading, hence dual coding, might more or less constrain the text-base elaboration. In other words, dual coding might leave no mental space for higher-level inference generation.

REFERENCES

- Barry, S. & Lazarte, A. A. (1998). Evidence for mental models: How do prior knowledge, syntactic complexity, and reading topic affect inference generation in a recall task for nonnative readers of Spanish? *Modern Language Journal*, 82, 176-193.
- Bernhardt, E. B. (1991). *Reading development in a second language*. Norwood, NJ: Ablex.
- Chu, H. (2002). Inference generation in L2 text memory: Effects of rhetorical convention, content familiarity, and grade level. *English Teaching and Learning*, 26(3), 59-78.
- Glenberg, A. M., & Langston, W. E. (1992). Comprehension of illustrated text: pictures help to build mental models. *Journal of Memory and Language*, 31, 129-151.
- Gyselinck, V., & Tardieu, H. (1999). The role of illustrations in text comprehension: What, when, for whom, and why? In H. van Oostendorp & S. R. Goldman (Ed.), *The construction of mental representations during reading* (pp.195-218). Mahwah, NJ: Lawrence Erlbaum.
- Johnson, R. E. (1970). Recall of prose as a function of the structural importance of the linguistic units. *Jornal of Verbal Learning and Behavior*, 9, 12-20.
- Liu, J. (2004). Effects of Comic Strips on L2 Learners' Reading Comprehension. *TESOL Quarterly*, 38(2), 225-243.

- Mayer, R. E. & Sims, V. (1993). Role of spatial ability in learning with computer-generated animations. Paper presented at the annual meeting of the American Educational Research Association, Atlanta.
- Moore, P. J. & Scevak, J. J. (1997). Learning from texts and visual aids: a developmental perspective. *Journal of Research in Reading*, 20(3), 205-223.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. New York: Oxford University Press.
- Sadoski, M. & Paivio, A. (2001). *Imagery and text: A dual coding theory of reading and writing*. Mahwah, NJ: Lawrence Erlbaum.

Multiple Instruction Strategies for Increasing Learners' English Vocabulary Retention

Su-yueh Wu (吳素月)
National University of Tainan
suyuehwu@mail.nutn.edu.tw

There has been a revival of interest in vocabulary teaching and learning because it causes a challenge for student with limited vocabulary to access the meanings of the texts. Learning is remembering; however, most researchers stress that it is common for us to forget what we first learn. Various areas of vocabulary research highlight the effect of certain teaching strategy on increasing students' vocabulary retention, such as explicit instruction, teaching in meaningful contexts, using films or any media. In particular, the purpose of the present study was to examine the effect of integrating various strategies into vocabulary teaching on students' vocabulary retention. Three classes of freshmen participated in various contexts for vocabulary learning. One class was exposed to one film with the explanation of new words, another class was exposed to explicit instruction and dialogue making practice, and the other class was exposed to one film with the explanation of new words, explicit instruction and dialogue making practice. Without preparation, students took a vocabulary pop quiz, which could provide students with the act of retrieving new words. The other data was collected from class observation notes and one questionnaire. The results showed that students exposed to the most activities got the best result in the vocabulary quiz; in contrast, students with single exposure got the lowest score. Additionally, students were highly enthusiastic about film watching. The questionnaire also revealed students' attitudes toward various teaching strategies. The findings demonstrated the significant effect of integrating a variety of teaching strategies on students' vocabulary retention. Consequently, the study highlights that the importance of multiple instruction strategies in vocabulary teaching because students can repeat new words in different, meaningful contexts. Future research and pedagogical suggestions will be discussed in the paper as well.

INTRODUCTION

The role that vocabulary knowledge plays in second and foreign language acquisition has lone been neglected; nevertheless, vocabulary is currently receiving increased emphasis in teaching practice as well as in research (Constantinescu, 2007). Previous researchers stressed that learners need a critical mass of vocabulary to get them over the threshold of the second language (Thornbury, 2008, p. 31). For instance, it is frustrating for learners to find the right word to fit the intended meaning when their store of words is limited (Thornbury, 2008, p. vi). Most learners acknowledge the importance of vocabulary acquisition as well (Thornbury, 2008, p. 13). In addition, vocabulary size targets are essential because they indicate the number of lexical items necessary to function in English. About 5000-6000 are for un-simplified reading texts or undergraduate study, about 8000-9000 are for newspapers, novels, and academic texts and about 10,000 are for academic texts (Nation., & Gu, 2007). Huang (2003) also suggested that English teachers in senior high instructional level should prioritize their students to acquire more English frequency words and general academic words to achieve better reading comprehension. There is strong relationship between vocabulary size and reading comprehension (Chao & Chen, 2005). Therefore, as we

enter the 21st century, acquisition of vocabulary has played a more important role in learning a second language (Lewis, 1993). With this shift in vocabulary emphasis, the language teachers are faced with the challenge of how best to help their students store and retrieve words in the target language (Sokmen, 2004). However, vocabulary learning strategies has attracted a noticeable lack of attention (Schmitt, 2004, p. 199).

Obviously, the learner needs not only to learn a lot of words but also to remember them. That implies that exercises and learning strategies that involve a deeper engagement with words should lead to higher retention than 'shallower' activities (Schmitt, & McCarthy, 2004, p. 3). That means instructional conditions should be arranged to provide opportunities for a maximum amount of processing of the words. Students should need to manipulate words in varied and rich ways. Sokmen (2004) stressed that providing a variety of classroom vocabulary exercises will expose students to possible strategies that they may discover feel right for them. In conclusion, the main purpose of the study was to examine the effect of applying various strategies in vocabulary teaching on students' vocabulary retention. Furthermore, it would demonstrate the more activities teachers apply in classroom teaching; the better students can retain vocabulary they learn.

LITERATURE REVIEW

Teaching Vocabulary in Multiple Contexts

It's effective for teachers to present new words in multiple contexts in which students can have more opportunities to learn and use them meaningfully and interestingly because repetition is considered to be an important aid to solidify new words in long-term memory. Memory clearly plays a key role in vocabulary learning. Thornbury (2008) interpreted that learning is remembering because vocabulary knowledge is largely a question of accumulating individual items. There are few short cuts in the form of generative rules: it is essentially a question of memory. Successful vocabulary learning clearly involves more than simply holding words in learners' mind for a few seconds (Thornbury, 2008, p. 23). In a word, vocabulary is not acquired in quick doses

When we respect this axiom that learning is remembering, the review and recycling of new language items will be critical if they stand a chance of becoming readily accessible in long-term memory. Moreover, forgetting mostly occurs shortly after the lesson and then the rate of forgetting diminishes. To avoid this lexical vanishing act, one solution is to follow the 'principle of expanding rehearsal' (Koprowski, 2006).

To achieve vocabulary rehearsal, researchers agree that repeating is an essential strategy for storing and memory. If words were repeated more times in a course-book, then students learnt them (O'Dell, 2004, p. 276); therefore, students should be required to manipulate words in varied and rich ways (Beck, McKeown, & Omanson, 1987). Thornbury (2008) also stressed that words need to be presented in their typical contexts; thus learners can get a feel for their meaning, their register, their collocations, and their syntactic environments. Based on Koprowski's (2006) experience, repetition and retrieval can be best achieved by organizing fun, competitive, and motivating vocabulary games and activities that adhere to the expanding rehearsal. In a word, it is effective for teachers to offer learners multiple exposures to words and opportunities to retrieve words from memory repeatedly.

Vocabulary Explicit Teaching

Vocabulary can be learned intentionally through explicit instruction in particular words. Explicit teaching can be a very good first introduction to a word; after this, the

context encountered when reading can lead to new knowledge of its collocations, additional meanings, and other higher-level knowledge. In addition, explicit teaching that can be repeated exposure from reading will help to consolidate the meaning(s) first learned (Schmitt, & McCarthy, 2004, p. 3). For most adult learners, direct vocabulary instruction is beneficial and necessary, due to the fact that they are not able to acquire the mass of vocabulary just by their own language practice. Learners can be taught explicitly how to increase their own vocabulary by teaching them appropriately vocabulary learning strategies in contrast to simply letting them learn vocabulary in their own way (Brown & Perry, 1991). Another similar statement is that better learning will take place when a deeper level of semantic processing is applied because the words are encoded with elaboration (Craik and Lockhart, 1972). In the light, providing ‘elaborative rehearsal,’ richer levels of encoding, will result in better learning (Sokmen, p. 242). In conclusion, current research would suggest that it is worthwhile to add explicit vocabulary to the usual inferring activities in the L2 classroom practice.

Teaching Vocabulary through Media

Research has shown that contemporary popular films are a valuable resource in classroom teaching. The study of TV and other mass media, new interactive media, and popular culture is important not only because of their profound influence and pervasiveness, but because of the ways that media easily become “naturalized” (Luke, 1999). Shepherd (1992) also stated that children have vast amounts of information from media sources, verbal, written and visual. It is the child’s task to make sense of it all, to construct reality from this information. Costanzo (2004) offers high school and college teachers a relevant way to engage their students through a medium that students know and love. He combines developments in pedagogy with many aspects of film study—film scholarship, the nature of movies themselves, American culture, globalization, and the connection with literary texts. Another study is that using silent films could facilitate the students’ critical thinking and writing skills (Kasper & Singer, 2004). In Chiu’s (2006) study, the sitcom provided students with a good source of vocabulary and the students learned more than 30 vocabulary words from each episode. Her result demonstrated that using sitcom activity could help the students become more active in language learning. Utilizing popular video materials is effective in eliciting students’ creative, fluent and remarkably expressive writing based on the targeted vocabulary and grammatical structures from video materials (Wolf, 2006). Wu (2004) found that students really enjoyed watching movies and video and they also got exposure to natural language in a non-threatening setting in exploring their further discussion and writing. All in all, previous research shows that the growing worldwide interest in film presents efficient and exciting teaching opportunities.

Dialogue Production

Offering learners interaction activities with other learners can improve their language learning (Schmitt, 2004, p. 205). Thorbury (2008) suggested that learners need to be actively involved in the learning of words and they would learn words effectively. While making dialogues with other learners, learners not only can be active but also have interaction opportunities with others. In addition, when students are asked to manipulate words, relate them to other words and to their own experiences, and then to justify their choices, these word associations are reinforced (Sokmen, 2004, p. 242). In this sense, constructing dialogues with other learners can be a helpful, effective teaching activity in which they can integrate new words to their prior learning. Furthermore, Nation (1990) stressed that dialogues have the advantage

of putting words directly into productive vocabulary. A similar but less structured technique, role-playing, is an option for more spontaneous oral practice of vocabulary. Dialogues can provide situations for students to practice ordinary conversation and offer them ample practice with basic speaking skills in context. Dialogues can be also viewed as short plays and used for students to act out rather than simply read aloud. The dialogues the students write function as basic communication at all levels (Scott & Ytreberg, 1990). In the light, putting pupils into pairs for the role-play in the daily dialogues is an effective way of oral practice for various ages and levels (Scott & Ytreberg, 1990). Role-plays are also excellent activities for learners' speaking in the relatively safe environment of the classroom before they must do so in a real environment (Nunan, 1999). Hence, constructing dialogues can offer learners valuable opportunities to manipulate the words they have learned and practice their oral skill in meaningful, personalized and non-threatening contexts.

METHODOLOGY

Participants and Procedures

Three classes of freshmen participated in various contexts for vocabulary learning. Class 1 had 52 students, Class 2 had 51 students and Class 3 had 55 students. However, only 45 students in Class 1, 48 students in Class 2 and 44 students in Class 3 could be counted for the data collection because some students didn't take tests and presence rate during the study were low as well.

Firstly, the present researcher chose two films she could wholeheartedly share with the students in the study because the movies must be appealing and engaging to your students: motivation is the key element in this activity (Wolf, 2006). In addition, as Tatsuki (2000) stressed that the teacher should like the film as well! The reason why the researcher chose two films was that few students from Class 2 and Class 3 participated in Class 1 as well. The film for the Class 1 was "You've Got Mail," and the other film for the Class 3 was "While You Were Sleeping."

Secondly, it took five weeks for students to finish watching DVD and videotape. Class 1 was exposed to one film with the explanation of new words. That meant the present researcher introduced and explained new words from each episode of "You've Got Mail" to students immediately after they watched. Another class of 51 (Class 2) was exposed to explicit instruction and dialogue making practice for five weeks. The researcher gave students one handout covering the new words from "While You Were Sleeping" every week. After she introduced the new words in various aspects such as word structure, syntactic pattern of the word in a phrase and sentence, meaning, lexical relations of the word with other words and common collocations. After this, she asked students to make their own dialogues by using the new words they just learned and to present their dialogue randomly. The other class of 55 (Class 3) was exposed to one film with the explanation of new words, explicit instruction and dialogue making practice. It meant that the researcher introduced and explained new words shortly after students watch each episode. Then in the following week, she passed out handouts to students, did explicit teaching and asked students to make their own dialogues in pairs.

Data Collection and Data Analysis

A delayed testing over one month later was carried out. Without preparation, three classes of students took a vocabulary pop quiz covering 10 new words they learned. They got to write down the Chinese translations of these ten English words. Students' test performance was coded by using the number of right Chinese translations. One-way ANOVA was utilized to compare the difference of their test

performance among these three classes. The other data was collected from class observation notes and one questionnaire. Class observation notes and one open-ended questionnaire were analyzed by using the qualitative approach.

RESULTS

The main purpose of the study was to examine the effect of applying various strategies in vocabulary teaching on students' vocabulary retention. Moreover, it could demonstrate the more activities teachers apply in classroom teaching; the better students can retain vocabulary they learn. The result showed that there was a significant difference among three classes. By using Post Hoc Tests to do multiple comparisons among these three classes, there was significant difference in each comparison between Class 1 and Class 2, Class 1 and Class 3, and Class 2 and Class 3. Therefore, the results showed that students exposed to the most activities got the best result in the vocabulary quiz; in contrast, students with single exposure got the lowest score. The results appear in Table 1 and Table 2.

Table 1. *All Means and Standard Deviations of Test Performance among Three Groups*

Class 1		Class 2		Class 3	
N=45		N=48		N=44	
M	SD	M	SD	M	SD
2.90	2.08	5.15	2.63	6.30	1.71

Table 2. *Significance of Test Performance among Three Groups*

Variance Source	SS	MS	F	P
Test Performance	222.68	111.34	21.25	.000

Moreover, students were highly enthusiastic about film watching. The questionnaire also revealed students' attitudes toward various teaching strategies. Students favored integrating three activities most, watching one film with the explanation of new words, explicit instruction and dialogue making practice. The class that was exposed to explicit instruction and dialogue making practice got better performance than the class watching one film with the explanation of new words; however, most of the students showed less positive toward these two activities than the class watching film only. Some students even thought it was boring to have explicit teaching and dialogue making only without film watching.

DISCUSSION

The findings demonstrated the significant effect of integrating a variety of teaching strategies on students' vocabulary retention. In the light, the study highlights that the importance of integrating multiple instruction strategies in vocabulary teaching because students can repeat new words in different, meaningful contexts. The more activities teachers apply in the vocabulary teaching; the better their students can retain what they learn. The finding is consistent with the findings of previous research. To achieve effective vocabulary learning, instructional conditions should be arranged to provide opportunities for a maximum amount of processing of the words (Beck et al., 1987; Koprowski, 2006; O'Dell, 2004; Thornbury, 2008); therefore,

students can repeatedly learn and manipulate words in varied and rich ways (Beck, McKeown, & Omanson, 1987). Moreover, this assures that organizing fun, competitive and motivating vocabulary games and activities can achieve best repetition and retention (Koprowski, 2006) because students don't appreciate boring repetition. For instance, the movie is appealing and engaging to the students in the study and their motivation is the key element in learning vocabulary. Another instance, dialogue making can stimulate students to express personally relevant meanings during interaction with the other learners; therefore, memory of new words can be reinforced in a rich and personalized context. Last, explicit instruction is effective; nevertheless, detailed handouts only cannot draw students' interest to pay attention to vocabulary learning. For future pedagogical and research suggestions, teachers and researchers can apply different instruction strategies, such as integrating oral, aural, visual, kinesthetic instruction activities in vocabulary teaching.

REFERENCES

- Beck, I. L., M. G. McKeown, M. G., & Omanson, R. C. (1987). The effects and use of diverse vocabulary instruction techniques. In M. G. McKeown and M. E. Curtis (Eds.), *The Nature of Vocabulary Acquisition* (pp. 147-163). Hillsdale, NJ: Erlbaum.
- Brown, T. S. & Perry, F. L. (1991). A comparison of three learning strategies for ESL vocabulary acquisition. *TESOL Quarterly*, 25, 655-670.
- Chao, K.C. & Chen, P.C. (2005). *Reading Comprehension: Vocabulary Size and Syntactic Competence*. 14th International Symposium and Book Fair on English Teaching.
- Chiu, C-Y. (2006). Using situation comedy (sitcom) to enrich learners' vocabulary. *The Internet TESL Journal*, XII (8). <http://iteslj.org/Techniques/Chiu-Sitcom.html>
- Constantinescu, A. I. (2007). Using technology to assist in vocabulary acquisition and reading comprehension. *The Internet TESL Journal*, XIII (2). <http://iteslj.org/Articles/Constantinescu-Vocabulary.html>
- Costanzo, W. V. (2004). *Great films and how to teach them*. Illinois, Urbana: NCTE.
- Craik, F. & Lockhart, R. (1972). Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671-684.
- Ellis, N. C. (2004). Vocabulary acquisition: word structure, collocation, word-class, and meaning. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, Acquisition and Pedagogy* (pp. 122-139). New York: Cambridge University Press.
- Huang, C.C. (2003). *Vocabulary Knowledge, Content Knowledge and Reading Comprehension, Senior High Students*. 12th International Symposium and Book Fair on English Teaching.
- Kasper, L. F., & Singer, R. (2004). *Unspoken content: Silent film in the ESL classroom*. Illinois, Urbana: NCTE.
- Koprowski, M. (2006). Ten good games for recycling vocabulary. *The Internet TESL Journal*, XII (7). <http://iteslj.org/Techniques/Koprowski-RecyclingVocabulary.html>
- Leeser, M. J. (2008). Pushed output, noticing, and development of past tense morphology in content-based instruction. *The Canadian Modern Language Review*, 65 (2), 195-220.
- Lewis, M. (1993). *The lexical approach*. Hove and London, England: Language Teaching Publications.

- Luke, C. (1999). Media and cultural studies in Australia. *Journal of Adolescent & Adult Literacy*, 42, 622-626.
- Nation, P. (1990). *Teaching and learning vocabulary*. New York: Newbury House.
- Nation, P., & Gu, P. (2007). *Focus on vocabulary*. Sydney: NCELTR.
- Nunan, D. (1999). *Second language teaching & learning*. Boston: Heinle & Heinle.
- O'Dell, F. (2004). Incorporating vocabulary into the syllabus. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, Acquisition and Pedagogy* (pp. 258-278). New York: Cambridge University Press.
- Schmitt, N. (2004). Vocabulary learning strategies. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, Acquisition and Pedagogy* (pp. 199-227). New York: Cambridge University Press.
- Schmitt, N., & McCarthy, M. (Eds.). (2004). *Vocabulary: Description, Acquisition and Pedagogy* (5th ed.). New York: Cambridge University Press.
- Scott, W.A., & Ytreberg, L. H. (1990). *Teaching English to children*. New York: Longman.
- Shepherd, R. (1992). Elementary media education: The perfect curriculum. *English Quarterly*, 25, 2-3.
- Sokmen, A. J. (2004). Current trends in teaching second language vocabulary. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, Acquisition and Pedagogy* (pp. 237-257). New York: Cambridge University Press.
- Sternberg, R. J. (1987). Most vocabulary is learned from context. In M. G. McKeown and M. E. Curtis (Eds), *The Nature of Vocabulary Acquisition* (pp. 89-105). Hillsdale, NJ: Erlbaum.
- Tatsuki, D. H. (1998). ESL/EFL lessons using movies. *The Internet TESL Journal*, IV (3). <http://iteslj.org/Lessons/Tatsuki-Movie/SimplePast.html>
- Thornbury, S. (2008). *How to teach vocabulary*. Harlow, England: Longman.
- Wolf, G. S. (2006). Using video to develop writing fluency in low-proficiency ESL students. *The Internet TESL Journal*, XII (8). <http://iteslj.org/Techniques/Wolf-UsingVideo.html>
- Wu, S-Y. (2004). *The application of group discussion in EFL classrooms*. Taipei: Tung Hua.

WriteAhead: An Abstract Writing Assistant System for Academic Writing

Ping-che Yang (楊秉哲), Hsin-chin Liou (劉顯親), Jason. S. Chang (張俊盛)
National Tsing Hua University
maciacalrk@gmail.com

Various online writing assistance programs have been developed through efforts in the areas of Natural Language Processing such as online thesaurus, automatic essay scorer, My Access, Criterion, web-based concordancers, word bank, and electronic portfolios. Their effectiveness when applied in language classrooms may vary depending on their functional rigor and pedagogical designs. In this paper, we introduce how we develop a computer program that can provide suggestions when academic writers are composing abstracts online. A computer system, WriteAhead, which provides learners with writing suggestions from a specific discipline on a web-based platform, in order to assist their abstract writing. The method involves automatically building domain-specific corpora of abstracts from the web via domain names and related hyponyms as query expansions, and automatically extracting proper names and grammatical collocations as writing suggestions from the corpora. When WriteAhead is running, according to learners' phrases or sentences, a set of corresponding writing suggestions can be actively triggered to present. This prototype abstract writing assistant system, WriteAhead, can enhance interactions between learners and system for writing abstracts in an effective and contextualized way. For assessment of WriteAhead, we conducted an experiment with graduate students of two groups who wrote their abstracts with and without the system. Surveys and interviews were also conducted for the assessment. Findings and implications will be presented in this conference to reveal the effectiveness of the system on different aspects of student writing.

INTRODUCTION

Using computer-assisted language tools in language teaching is a trend in TESOL, especially teaching English for Specific Purposes (ESP). As to English for Academic Purposes (EAP), research has conducted fruitful results concerning the specific genre of research articles (RA) in two aspects: on one hand, Swales (1990) implemented genre analysis approach to study the macrostructure of RA, such as information structures and rhetorical functions of various stages; and implemented text analysis approach to study the linguistic features of RA, such as tense, voice, or modals. Besides, a number of studies have begun to explore the association of the lexico-grammatical features with the rhetorical functions of particular stages or the writing strategies. On the other hand, studies (Gledhill, 2000; Howarth 1996; Nattinger & DeCarrico, 1992) have proposed a phraseological approach to the teaching and learning of EAP writing; in other words, identifying and presenting recurrent lexico-grammatical patterns in a genre. This is regarded as suitable for RA, since it is a highly conventionalized and specialized genre. In Dong (1998), the study on non-native graduate students' research articles showed that the major perceptions of assistance that graduate students estimated from advisors' suggestions are collocation and grammatical usage; Cobb *et al.* (2001) showed the improvement of vocabulary acquisition using a language tool including a concordancer, which might

imply a potential need of language tool for such assistance. This paper presents an abstract writing assistant system to help graduate students to write their abstracts, especially on automatically providing writing suggestions of word completion and collocation. The following paragraphs will introduce related work and differences with this paper; the methods to develop the system and preliminary findings.

RELATED WORK

Since Scott (1998) developed Wordsmith Tools, a series of related studies for ESP have been published, and corpus-based concordancers have played an important role on pedagogical practice. Based on Scott's work, Curado Fuentes (2001) indicated that the corpus genre and size might influence the effectiveness when applying to a certain disciplines for ESP, and that the importance of updating contents of the corpus due to rapid changes with time. Besides, the availability of corpora from specific disciplines varies, which increases the complexity when building a multi-discipline task. The proposed system applied a novel web-based approach, which enabled to retrieve data of specific disciplines from the web, and automatically build and update corpora with time.

Another issue respecting concordancers is that the benefit of using a concordancer lies in the user's knowledge of how to use it, especially how to give a correct query to the concordancer in order to have a better searched result. In 2006, Google launched its new feature called *Google suggestion* for search engine, which will automatically display a list of possible query results for user as query suggestions. The proposed system adopted this feature, aiming at automatically providing users with possible suggestions such as words or phrases while users write a partial segment of a sentence.

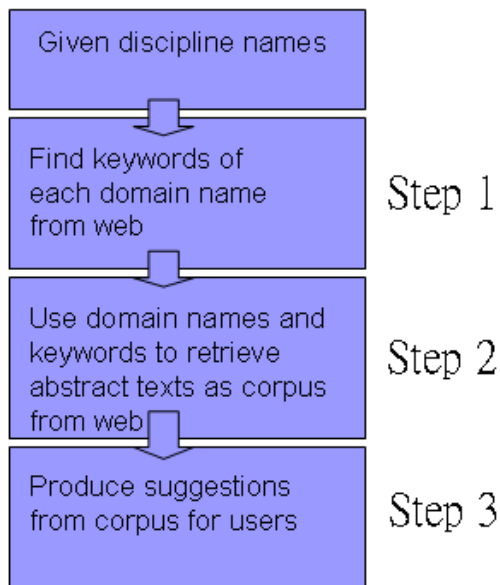
In addition to concordancers, previous research has designed prototype writing systems for EAP. In China, Yang (2000) developed the EAME (Electronic Abstract Made Easy) for writing abstracts, using question-driven framework to generate abstract automatically by asking users a series of blank-filling questions concerning users' research. In the contrary, our proposed system did not ask users questions; instead, our proposed system took segments of users' keyboard input as questions, attempting to give suggestions as possible as users' expected answers. In Brazil, Genoves et al. (2007) developed the Argumentative Zoning for English Abstracts (AZE) authoring tool to identify the structure of research articles based on the work of Teufel & Moens (2002). Our proposed system also integrated this part to provide words of different structural.

In a historical view, as for related EAP work in EFL setting in Taiwan, Kuo (1999, 2002) has worked on the phraseology in scientific research articles with the help of computer programs; however, the use of the in early days, due to shortcomings of software and financial limitations in most EFL setting in Taiwan, it was less possible to have access to use such language tools for EAP. However, with the emergence and popularity of Internet technologies, instructions aided by web-based tools and materials are gradually feasible in practice. Sun (2007) proposed a web-based Scholarly Writing Template (SWT) system for graduate students from different disciplines, and an empirical observation that graduate students tend to consult Google Scholar to check collocations or word usages when writing their research articles. The proposed paper adopted this observation and implemented it in the method section.

METHOD

We introduce a method for actively providing learners with referable content-based writing suggestions of a discipline-specific corpus. The Figure 1 illustrates the whole procedure of such method. The method can be divided into four steps: In the first step, each discipline name was queried from web to retrieve its relative keywords. For example, the discipline name “Computer Science” returned keywords such as “Artificial Intelligence”, “Natural Language Processing”, “wireless network”, and so on. In the second step, the discipline name and its keywords as query expansion were both automatically queried from web to retrieve abstracts that contains queried words. For instance, the discipline name “Computer Science” and its keyword “Natural Language Processing” were both queried from web to retrieve abstracts that contains these two words; as a result, the retrieved abstracts were accumulated as corpus. In the third step, suggestions were produced from the corpus for users.

Figure 1. Main procedure of the method



Regarding the type of the suggestions for users, the third step could be divided into five parts. To begin with, words of high frequencies in the corpus were extracted as a dictionary, which was to provide function of word completion. Second, n-grams of each abstracts were extracted as phrase suggestions; third, collocations were also extracted from the abstracts as suggestions; fourth, words of the abstracts were tagged with structural moves as suggestions for frequently seen words among different moves; fifth, frequent transition phrases were extracted as transition suggestions. In addition to the corpus, instructions of writing reference were also added to the system as suggestions.

To implement the method, each step mentioned in the method will be specified in detail as follows. For first step, discipline names were queried from Wikipedia to retrieve discipline-specific keywords automatically. We adopted Wikipedia as query sources for the retrieval of discipline keywords. The reason why Wikipedia is adopted as query source is that it is the current largest knowledge internet resource in the world, which involves over 150 thousand users editing over 284 million knowledge entries, and that Weaver (2006) proved a high validity and usability of the contents of Wikipedia. In our method, we retrieved discipline-specific keywords by extracting

words and phrases with hyperlink in the entry name of its discipline in Wikipedia from twenty-seven different disciplines shown in Table 1.

Table 1. *Number of keywords in corresponding to their disciplines*

Discipline	Number of keywords	Discipline	Number of keywords
Electrical Engineering	157	History	249
Computer Science	123	Psychology	2715
Biology	772	Music	364
Chemistry	974	Sports	246
Physical Science	220	Health Science	125
Communication	166	Linguistics	337
Mathematics	209	Marketing	213
Management	67	Nutrition	409
Finance	667	Mechanical Engineering	244
Economics	468	Geography	92
Sociology	742	Architecture	790
Education	715	Philosophy	558
Law	38	Literature	623
Agriculture	216	Total Keywords	12499

In step 2, each of the discipline-specific keyword was regarded as a query expansion along with its discipline name as query. Google Scholar was assigned as a specialized search engine for retrieving data since most graduate students in Taiwan have experience consulting it (Sun, 2007). The goal of retrieving data is to retrieve PDF files of RA, especially abstracts in journal articles. The query and its query expansion then retrieved from Google Scholar a list of HTTP addresses which especially specified in PDF format. If the list contained many addresses that came from the same site, the site would be considered as the major source of that discipline; otherwise, the lists of PDF files would be automatically downloaded from internet. After observing all the of the retrieved list, an interesting empirical evidence was found that, eleven lists of disciplines contained most addresses of the same site called Springerlink, which is the second largest publisher in the world. Since we focus on building corpora of abstracts and Springerlink site automatically provides pure abstracts of these eleven disciplines, it is taken for granted that we turned to build corpora by retrieving abstracts from Springerlink. As for the remaining lists that can not apply for this case, we still provide ways to extract abstracts from retrieved PDF files, which will not be discussed here. The following method will focus on the case of retrieving purely abstracts.

In step 3, there are three steps of preparations to produce suggestions from a corpus. To begin with, each abstract extracted from Springerlink was sent into Tsujii English tagger (Tsuruoka and Tsujii, 2005), which aimed at providing each word with information of accurate POS (part of speech). Second, all of the words of the abstracts were classified by a language tool called *mkcls* developed by Och (1999), and were then filtered out by BNC (British National Corpus) to extract frequent words, which aimed at providing highlighting words for suggestions. Third, Lin (2009) developed *MoveTagger* to tag moves in the abstracts via Maximum Entropy with the average accuracy of 81%, which was used to analyze the extracted abstracts to produce the frequent content words (nouns, verbs, adjectives, adverbs) of each moves.

After the preparation work, the system automatically produced five specific types of suggestions from a corpus. As to the first type, the system counted all of the words in a corpus to provide function of word completion. Concerning the second type, the system applied n-gram model to divide sentences of abstracts into segments with the phrasal length of trigrams and four-grams, and counted their frequencies. Regarding the third type, the system calculated VN (verb-noun), NN (noun-noun), AN (adjective-noun), IN (preposition-noun), RV (adverb-verb) and NV (noun-verb) collocations and their frequencies with the help of POS-tagged abstracts. Respecting the fourth type, the system calculated verbs, nouns, adjectives and adverbs with frequencies from moves of background, purpose, method, result and conclusions. In concerned with the fifth type, the system listed all the transition phrases by judging a sentence that the number of word in a sentence is less than four and a sentence begins with an adverb or preposition. With a bonus, the reference materials from the academic writing class were also added to the system.

Figure 2 shows the graphic interface of the WriteAhead system. The upper-right window displays the function of word completion; the middle-right window displays n-gram phrases, collocations and moves, and the bottom-right window displays transition phrases along with an extra hyperlink linked to a new web page containing reference materials. The interface of the system is designed with JavaScript, Ajax, and the database was processed with Python and Sqlite.

To use the system in practice, when a user types an incomplete word and pause more than 0.5 sec, the system will automatically show Top 15 word that are most close to the user's input. When a user types more than two words and hit a space key, the middle-right window will be initiated, and the system will regard the last two words as input and do the matching work in the database. If there are no matching results in the database, the system will then use the last word as input instead. If a user wants to look up a collocation, by typing a wildcard '*' and a word, the system will show its precedent word. For example, when a user types '* method' and hits a space, the system will display 'proposed the method', 'a new method', 'the method' and so on. Furthermore, when a user wants to look up a specific collocation, by typing a wildcard '*' and a specific part of speech, by hitting a space the system will show its corresponding collocations. For instance, if a user types '*A method' and hits a space, the system will display 'available method', 'similar method', 'previous method' and so on. What is more, when a user wants to look up the frequent word in a certain move, by typing a wildcard '*' and a specific move, by hitting a space the system will show Top 15 frequent words in a certain move. For instance, if a user types '*B Verb' and hits a space, the system will display 'propose', 'present', 'describe' and so on.

Figure 2. The WriteAhead system interface.



CURRENT PROGRESS

Current progress is now focusing on an experiment involved with 58 graduate students and doctoral students from two classes of academic writing, where the majority is from CS background. The number of one class is 38 and the other is 20. Students will be divided into two groups by random selection, and each group is to complete a paraphrasing and translating task on two simplified CS-related abstracts in the form of entries of moves. On one hand, students are only allowed to write with the system only; on the other hand, students are allowed to write with word processors. We have first completed the first 38-student class on writing with the system and continued to conduct other three experiments later. This presentation shows a rough picture of the abstract writing system. Introduction, related work, method and current progress were shown. In the coming future, surveys and interviews will be conducted for the assessment. Findings and implications will be presented in this conference to reveal the effectiveness of the system on different aspects of student writing.

REFERENCES

- Cobb, T., Greaves, C. and Horst, M. (2001). 'Can the rate of lexical acquisition from reading be increased? An experiment in reading French with a suite of online resources' in P. Raymond and C. Cornaire (eds.) *Regards sur la didactique des langues secondes*. Montréal: Éditions Logique.
- Curado Fuentes, A. (2001). Lexical Behaviour in Academic and Technical Corpora: Implications for ESP Development, *Language Learning & Technology*, Vol.5, 106-129.
- Dong, Y. (1998). Non-native graduate students' thesis/dissertation writing in science: Self-reports by students and their advisors from two US institutions. *English for Specific Purposes*, 17(4), 369-390.
- Genoves L., Feltrim V.D., Dayrell C., Alusio S. (2007). Automatically detecting

- schematic structure components of English abstracts: building a high accuracy classifier for the task. *Proceedings of Recent Advances in Natural Language Processing (RANLP2007): International Workshop on Natural Language Processing for Educational Resources*, Borovets, Bulgaria, 27-29.
- Gledhill, C. (2000). The Discourse Function of Collocation in Research Article Introductions. *English for Specific Purposes*, 19, 115-135.
- Howarth, P. (1996). *Phraseology in English Academic Writing*. Tübingen: Max Niemeyer Verlag.
- Kuo, C. H. (1999). Certainty, probability, and capability: Modal verbs in scientific journal articles. *Proceedings of the 1999 English for Specific Purposes Conference*, 218-231. Taoyuan, Taiwan: Ming Chuan University.
- Kuo, C. H. (2002). Phraseology in scientific research articles. *Proceedings of the Eleventh International Symposium on English Teaching*. 405-411. Taipei: Crane.
- Lin, Y. H., Chen, M. H., Wu, J. C. and Chang, J. S. (2009) Automatically Identify Sections in Academic Abstracts. *Proceedings of 2009 International Conference on Applied Linguistics & Language Teaching (ALLT)*. Taipei, Taiwan: National Taiwan University of Science and Technology, 357-368.
- Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.
- Och, F. J. (1999). An Efficient Method for Determining Bilingual Word Classes. In *Ninth Conference of the European Chapter of the Association for Computational Linguistics (EACL'99)*, 71-76. .
- Scott, M., (1996), *Wordsmith Tools*, Oxford: Oxford University Press. ISBN 0-19-458984-6.
- Sun, Y. C. (2007). Learner perceptions of a concordancing tool for academic writing. *Computer-Assisted Language Learning*, 20(4), 323-343.
- Swales, J. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.
- Teufel, S., Moens, M. (2002). Summarizing scientific articles —experiments with relevance and rhetorical status. *Computational Linguistics* 28, 409-446.
- Weaver, G., Strickland, B., and Crane, G. (2006). Quantifying the accuracy of relational statements in Wikipedia: a methodology. In *Proceedings of ACM/IEEE Joint Conference on Digital Libraries*, 358–358.
- Tsuruoka, Y. and Tsujii, J. (2005). Bidirectional Inference with the Easiest-First Strategy for Tagging Sequence Data, In *Proceedings of HLT/EMNLP 2005*, 467-474.
- 楊永林 (Yang, 2000)。EAME：一種基於問句驅動、鏡像模擬基礎之上的文本生成系統。《現代外語》3，296-305。

Effects of Chinese Transliteration Systems on English Literacy Learning

Tzu-yin Yang (楊茨茵), I-ru Su (蘇怡如)

National Tsing Hua University

u1501012@nccu.edu.tw

irusu@mx.nthu.edu.tw

The present research followed the methodology of Holm and Dodd (1996)'s study in order to examine whether Taiwan learners of English differ from China and Hong Kong learners in phonological awareness and English literacy skills. Ten Taiwan EFL learners at the college level participated in this study. All of the participants took a battery of phonological awareness tasks, a non-word reading task, and a non-word spelling task. Results of the study showed that in the phonological awareness tasks at the phonemic level, Taiwan learners performed worse than the China learners but better than the Hong Kong learners; however, in the phonological awareness task at the onset-rhyme level, no significant difference was found between Taiwan and China groups, and both of them outperformed the Hong Kong group. In the English literacy tasks, it was found that in the reading task of non-words with high-frequency rhymes, no significant difference was found between Taiwan and China groups, and both of them outperformed the Hong Kong group, while in the reading task of non-words with low-frequency rhymes and non-word spelling task, Taiwan learners performed significantly worse than the China group but better than the Hong Kong group. The results of the study provide evidence that Taiwan EFL learners' experience with *Zhuyin Fuhao* does help them develop phonological awareness at the sub-syllabic level, though the phonological awareness is enhanced more at the onset-rhyme level than at the phonemic level, and that *Zhuyin Fuhao* does facilitate English literacy learning, though not as well as *Pinyin* does.

INTRODUCTION

In the literature on the acquisition of English literacy among Chinese learners of English, it has been consistently found that different transliteration experiences in learning Chinese literacy would influence how Chinese learners of English read or spell English words. For example, an instructional strategy like *Pinyin* that emphasizes the phonemic segmentation in Chinese may lead learners to develop a phonological awareness at the phonemic level (Cheung, Chen, Lai, Wong, & Hill, 2001; Holm & Dodd, 1996; Read, Zhang, Nie, & Ding, 1986), and this skill can be positively transferred to learners' English literacy acquisition (Holm & Dodd, 1996; Wang, Perfetti, & Liu, 2005). Conversely, an instructional strategy like "Look and Say" method that focuses on the relationships between entire Chinese word forms and their corresponding pronunciations may limit learners to a whole-word, visual strategy when processing English words due to their underdevelopment of phonological awareness (Holm & Dodd, 1996; Leong, Tan, Cheng, & Hau, 2005).

Zhuyin Fuhao, a phonetic system used in Taiwan, also helps learners develop phonological awareness at the sub-syllabic level (Bertelson, Chen, & de Gelder, 1997; Huang & Hanley, 1995, 1997). However, it differs from *Pinyin* in that not every

Zhuyin symbol represents a single phoneme; instead, the *Zhuyin* symbols representing rhymes may comprise more than one sound. Due to the difference between the two phonetic systems, it remains unclear whether the early experience with *Zhuyin Fuhao* will help Taiwan EFL learners develop phonological awareness that is required in learning English literacy as well as the phoneme-based *Pinyin* does. Hence, the purpose of the present study is to investigate the effects of Taiwan EFL learners' experience with *Zhuyin Fuhao* on the English literacy learning in terms of reading and writing words.

LITERATURE REVIEW

Different cognitive processes in reading English and Chinese

In world's orthographies, English and Chinese are inherently different in the writing system. In an alphabetic language like English, each graphic unit carries segmental phonological information and thus a linear assembly of the segments is vital in phonological decoding. Learning English therefore facilitates children's phonological awareness at the phonemic level. In contrast, the Chinese writing system does not reflect the segmental structure that is fundamental to alphabetic writing systems. Instead, the phonetic information in Chinese is defined at the syllable level rather than at the phonemic level, which suggests that learning to read Chinese requires mapping of spoken syllables onto written characters, rather than onto a phoneme (Perfetti, Liu, & Tan, 2000). In the literature, converging evidence has shown that compared with ESL or EFL learners with alphabetic L1 backgrounds, learners with non-alphabetic L1 backgrounds may have much more difficulty learning English, because with lack of experience with the alphabet, their phonological awareness is underdeveloped, which in turn, may impede the learning process in reading and writing English words. (Akamatsu, 1999; Hamada & Koda, 2008; Wang, Koda, & Perfetti, 2003). However, it should be noted that predictions about how native readers of Chinese read or write English cannot be made without taking different Chinese literacy instructions into consideration, because the instructional methods used in different Chinese-speaking regions may influence whether a reader of Chinese truly has a non-alphabetic literacy experience.

Different approaches in learning Chinese literacy

“Look and Say” Method in Hong Kong. Most Cantonese-speaking children learning to read Chinese characters in Hong Kong are not exposed to any phonetic system, but begin directly with characters. The approach to teaching reading in Hong Kong is the “Look and Say” method, with the whole character as the basic unit. Children are taught character-to-pronunciation mappings without the mediation of any phonetic system. A new character is pronounced and explained at the beginning of a lesson and then related to its context (Ho & Bryant, 1997).

***Pinyin* in China.** In 1958, China introduced an alphabet-based phonetic system employing Latin symbols, called *Pinyin*, as part of a general language reform movement. The *Pinyin* system uses letters of the western alphabet to produce a phonemic representation of Chinese syllables. Pre-literacy children are first introduced to *Pinyin* as a transitional alphabet for learning to read Chinese characters; therefore, we can say that children in China establish their alphabetic literacy during the process of learning to read logographic Chinese (Shu & Anderson, 1997).

***Zhuyin Fuhao* in Taiwan.** The initial reading instruction to children in Taiwan is based on a phonetic system called *Zhuyin Fuhao*, which represents the sounds of Mandarin by 37 symbols. Unlike the *Pinyin* system where each phonetic symbol represents a single phoneme, each *Zhuyin* symbol represents either an initial, a

medial or a rhyme. As in *Pinyin*, the initials (onsets) are represented by distinct symbols, and those symbols constitute 21 of *Zhuyin*'s 37 characters. However, different from *Pinyin*, in *Zhuyin Fuhao*, each rhyme also has a distinct symbol, representing a single vowel (see Table 1), a diphthong (see Table 2), or a vowel with nasal coda (see Table 3). For example, *luan* is written ㄌㄨㄢˊ (l-u-an), where the last symbol ㄢˊ represents the rhyme -an. These rhymes constitute the other 16 characters of *Zhuyin Fuhao* (Bertelson et al., 1997; Kuo, 2006). The fact that in *Zhuyin Fuhao*, not every symbol represents a single phoneme makes it different from the phoneme-based *Pinyin*.

Table 1. *Rhyme without a coda.*

	High vowel			Mid & low vowel				
<i>Zhuyin</i>	一	ㄨ	ㄩ	ㄚ	ㄛ	ㄜ	ㄝ	ㄞ
<i>Pinyin</i>	i	u	ü	a	o	e	e	er
IPA	i	u	y	ɛ	o	ɤ	ɛ	ɛ̃

Table 2. *Diphthong: rhyme with a glide coda*

Glide coda	y		w	
<i>Zhuyin</i>	ㄞ	ㄟ	ㄞ	ㄟ
<i>Pinyin</i>	ai	ei	au	ou
IPA	ay	ey	ɔw	ow

Table 3. *Nasal-rhyme: rhyme with a nasal coda*

Nasal coda	n		ŋ	
<i>Zhuyin</i>	ㄢ	ㄣ	ㄤ	ㄥ
<i>Pinyin</i>	an	en	ang	eng
IPA	an	ən	aŋ	ɛŋ

The effects of Chinese L1 transliteration experience on phonological awareness

Given that learning an alphabetic system facilitates children's phonological awareness at a lower linguistic level, it is not surprising to find that much research has shown that different approaches in learning Chinese literacy would affect reader's phonological processing skills differently. In fact, converging studies have shown that Chinese children perform much better on the phonological awareness tasks after the introduction of *Pinyin* and *Zhuyin Fuhao* (Bertelson et al., 1997; Cheung et al., 2001; Huang & Hanley, 1997; Holm & Dodd, 1996; Read et al., 1986). However, for Hong Kong children who have learned Chinese without the assistance of any phonetic system, their performance on phonological awareness is worst compared with children from Taiwan or China (Bertelson et al., 1997; Cheung et al., 2001; Holm & Dodd, 1996).

In a pioneering study conducted by Read, Zhang, Nie, and Ding (1986), they compared a group of adult subjects who had learned to read Chinese characters through *Pinyin* with a group who had not learned *Pinyin*. The findings indicated that China adults who were literate only in Chinese characters found it difficult to add or delete individual consonants in spoken Chinese words. In contrast, China readers who were experienced in using an alphabet-based phonetic system in learning to read Chinese were more successful in manipulating speech sounds than those who were literate only in Chinese characters.

To further examine the effect of different Chinese instructional methods on children's phonological awareness development, Cheung, Chen, Lai, Wong, & Hills (2001) examined the effects of different transliteration experiences on phonological awareness by comparing young pre-readers and older literate children from Guangzhou and Hong Kong. Their results showed that Hong Kong and Guangzhou pre-readers performed very similarly at all levels of Chinese phonological awareness tasks; however, after the instruction of *Pinyin* to Guangzhou children, Guangzhou literate children outperformed their Hong Kong counterparts on the onset and coda analyses. The findings of their study provide further evidence that the early literacy experience with the alphabetic phonetic system, *Pinyin*, would help children enhance their phonological awareness.

With these findings showing that *Pinyin* and "Look and Say" method affect Chinese speakers' phonological awareness differently, one interesting question is raised. That is, to what extent does *Zhuyin Fuhao* influence Taiwan children's phonological awareness? To examine the effect of learning *Zhuyin Fuhao* on children's subsequent phonological awareness skills, Huang and Hanley (1997) administered Chinese phonological awareness tasks before and after the instruction of *Zhuyin Fuhao* to Taiwan children. The results revealed that Taiwan children performed much better on both onset/rhyme task and phoneme deletion task after the introduction of *Zhuyin Fuhao*, which is reminiscent of Read et al. (1986)'s finding showing that people who have learned *Pinyin* had no difficulty in adding or deleting individual consonants in spoken Chinese words because of the development of phonemic awareness. Huang and Hanley's findings seem to suggest that just like *Pinyin*, the phonetic script *Zhuyin Fuhao* can help children develop phonological awareness not only at the onset-rhyme level but also at the phonemic level. However, although it was found that the experience with *Zhuyin Fuhao* can improve learners' phonological awareness ability in Chinese, seldom did any research directly examined whether Taiwan learners' experience with *Zhuyin Fuhao* can be positively transferred to the development of phonological awareness that is needed in learning English literacy. The relationship between *Zhuyin Fuhao* and phonological awareness that is needed in learning English literacy therefore requires further clarification.

The effects of Chinese L1 transliteration experience on English literacy skills

In the previous review, we have learned that for Chinese learners of English, different Chinese transliteration experiences in learning Chinese literacy would influence their phonological awareness differently. With a substantial body of evidence showing that phonological awareness is significantly related to the acquisition of literacy skills (Bialystok, Majumder, & Martin, 2003; Bradley & Bryant, 1991; Goswami & Bryant, 1990), it seems reasonable to suggest that such effect of different transliteration experiences is very likely to pose a possible difference in the use of phonological information in reading processes among Chinese readers from different regions.

In Chen and Yuen (1991)'s study, for example, they compared second-grade children from China, Hong Kong, and Taiwan in the use of phonological information in pronouncing Chinese pseudo-characters. Results revealed that children from China and Taiwan utilized more phonetic cues from phonetic radicals to pronounce pseudo-characters than did children from Hong Kong, which suggests that the experience with a phonetic script, *Pinyin* or *Zhuyin Fuhao*, helps children become more aware of phonological information contained in characters during the Chinese reading process.

Given that the enhanced phonological awareness that is engendered by the learning of *Zhuyin Fuhao* or *Pinyin* might promote readers to make more use of the phonetic cues in reading Chinese, it is not surprising to find that a number of studies have indicated that Chinese readers from different regions differ in English word literacy skills due to the differences in exposure to phonetic symbols in learning to read their L1 (Holm & Dodd, 1996; Leong, Cheng, and Tan, 2005).

In a study conducted by Holm and Dodd (1996), they compared the performance of college students from Hong Kong and China on a set of English phonological awareness, reading, and spelling tasks. The results showed that students from Hong Kong read and spelled English real words about as well as the students from China. However, they were dramatically less competent than their China counterparts on a set of phonological awareness tasks and English non-word reading and spelling tasks despite the fact that Hong Kong students had the longer history of English learning. Holm and Dodd interpreted this difference as the result of the Hong Kong students' lack of exposure to a phonetic system in learning their first language. Thus, when the phonological awareness required in reading and writing English had not been developed in their L1, Hong Kong students were limited to a whole-word, visual strategy in their English literacy acquisition. In contrast to Hong Kong learners, China learners' experience with *Pinyin* helped them develop phonological awareness at the phonemic level, and therefore they were able to transfer these skills to English in reading and spelling English non-words.

This effect of *Pinyin* on English literacy skills is directly relevant to the present study because Taiwan EFL learners concerned in this study have a transliteration learning history similar to the learners in China. *Zhuyin Fuhao*, similar to *Pinyin*, also promotes Taiwan children's phonological awareness at the sub-syllabic level (Bertelson et al., 1997; Huang & Hanley, 1997). However, it still remains unknown whether the phonological awareness that Taiwan EFL learners have developed through their experience with *Zhuyin Fuhao* in acquiring Chinese literacy will help them acquire literacy skills in English. Further research is needed before we can have a more conclusive answer to this question.

PURPOSE OF THE STUDY AND RESEARCH QUESTIONS

The purpose of the present study aims to investigate the effects of Taiwan EFL learners' experience with *Zhuyin Fuhao* on their phonological awareness and English literacy skills. To guide the study, two research questions are proposed:

1. Does Taiwan EFL learners' experience with *Zhuyin Fuhao* in acquiring Chinese literacy help them develop the phonological awareness that is needed in learning English literacy? Do Taiwan EFL learners differ from learners from Hong Kong and China in English phonological awareness skills?
2. Does the phonological awareness that Taiwan EFL learners have developed through their experience with *Zhuyin Fuhao* in acquiring Chinese literacy help them acquire literacy skills in English? Do Taiwan EFL learners differ from learners from Hong Kong and China in English literacy skills?

In order to address the two questions, the study would follow the methodology of Holm and Dodd (1996)'s study. The results of Taiwan EFL learners obtained in this study would be compared to those of Hong Kong and China learners in Holm and Dodd (1996)'s study to examine whether Taiwan EFL learners would exhibit any difference from Hong Kong and China learners in phonological awareness skills and English literacy skills of reading and spelling English non-words.

METHOD

Participants

Since in Holm and Dodd (1996)'s study, each of the China and Hong Kong subjects had achieved a score of 6.5 on IELTS or 573 on TOEFL (paper version), in the present study, ten Taiwan EFL learners (five males and five females) at the college level with an internet-based TOEFL score of 89 (comparable to a paper-version TOEFL score of 573) were recruited. The demographic information of Hong Kong learners, China learners and Taiwan learners are presented in Table 4. In each group, there were ten participants. Hong Kong subjects had the longest English learning experience, while China subjects the least. It should be noted that in the present research, the data of Hong Kong and China subjects were drawn from Holm and Dodd (1996)'s study.

Table 4. *Subject characteristics for each group*

Subjects	Mean number of years	SD
<i>Hong Kong (n=10)</i>		
Age	23.2	3.71
English literacy exposure	15.0	3.19
<i>China (n=10)</i>		
Age	31.5	3.37
English literacy exposure	10.4	3.37
<i>Taiwan (n=10)</i>		
Age	22.0	2.40
English literacy exposure	12.2	3.33

Instruments

In order to compare performance of Taiwan EFL learners with that of learners from Hong Kong and China, the present research followed the methodology of Holm and Dodd (1996)'s study and adopted the instruments and the procedure used in their study.

Phoneme segmentation task. This task required subjects to analyze the internal composition of words at the phonemic level by identifying the number of phonemes. There were twenty-four stimuli in total and they were grouped into three lists. The three lists included real words with a one-to-one phoneme to grapheme correspondence (e.g., stamp); real words with a one-to-many phoneme to grapheme correspondence (e.g., whistle); and pseudo-words with a one-to-one phoneme to grapheme correspondence (e.g., stelp). Three practice items, one for each type, were given, and it was stressed that sounds not letters were to be counted.

Phoneme manipulation task. This task tested subject's ability to segment and transpose word initial consonants of word pairs. It assessed subject's ability to manipulate complex phonemic information. Subjects were required to make spoonerisms of ten word pairs by transposing the initial phonemes of each pair (e.g., big dog → dig bog). The ten pairs of words included: two pairs of words where both words began with a single letter/ sound (e.g., poor teddy); four pairs of words where a digraph occurred in the first or second element of the spoonerism (e.g., soft cheese); two pairs of words where both initial elements contained digraphs (e.g., chip shop); and two pairs of words which began with clusters (e.g., crowd play). Five practice items, two for the cluster type and three for the other three types, were provided to ensure the subjects understand instruction of the task. To be counted as correct the

onsets that were transposed both had to be correct.

Rhyme judgment task. In this task, subjects were required to decide whether a word pair rhymed. Four different types of word pairs were presented: orthographically similar rhyming words (e.g., rang/ sang); orthographically dissimilar rhyming words (e.g., through/ new); orthographically similar non-rhyming words (e.g., said/ paid); and orthographically dissimilar non-rhyming words (e.g., wait/ wet). Test items were presented at random order. Twenty pairs of words were presented, five for each type. Four practice trials were given, one for each type.

Non-word reading task. In the reading task, subjects were asked to read two lists of non-words and they were asked to pronounce each one however they thought best. The stimuli included 48 non-words, 24 words for each list. The graphemes of one list were arranged so that the rhyme units of the items were more frequent in real words than those items in the other list. The theory behind the task is that if experience with real words like *main* and *rain* helped the subjects pronounce non-words like *tain*, there will be a difference in the accuracy of reading the two lists. If a straightforward grapheme-phoneme conversion occurs, the accuracy of reading the two lists should be similar. The total score of each subject was calculated based on the percentage of non-words that were pronounced correctly.

Non-word spelling task. In the non-word spelling task, twenty-four non-words were presented. It was emphasized that there were no correct spellings for the non-words, and subjects were encouraged to spell the words according to how those words sound. One-third of selected stimulus items were composed of segments analogical to real words (e.g., complefation), one-third had morphemic unit endings (e.g., kleppering), and one-third were made up of discrete morphemes and they were not easily spelled by an overt analogy strategy with derivational endings (e.g., shepelartoe). Subjects were given enough time to finish one item before the next was presented. Venezky (1970) grapheme-phoneme correspondence rules were used to identify the spelling counted as correct. In scoring non-word accuracy, any spelling that is plausible according to grapheme-phoneme correspondence rules is accepted as correct, even if it was not identical to the stimuli. The total score of each subject was calculated based on the percentage of non-words that were spelled correctly.

Procedure of the study

All the stimuli were recorded by a native English speaker. Subjects were tested individually in one 60-minute session in a quiet room. All subjects received the tasks in the same order. All the test sessions were recorded.

RESULTS AND DISCUSSION

A series of independent two-sample *t* test was carried out to examine whether Taiwan EFL learners would exhibit any difference from learners from Hong Kong and China in English phonological awareness skills and English literacy skills of reading and spelling English non-words. First, the means and standard deviations of each task were calculated for Taiwan EFL learners. With the mean and standard deviation, we were able to use *t* test to compare Taiwan subjects with the other two groups. The results of descriptive statistics and *t* test are shown in Table 5.

Table 5. Comparison of performance in the experimental tasks among learners from Hong Kong, China and Taiwan

	Mean	SD	<i>t</i> value
<i>Phoneme segmentation task</i>			
Hong Kong	36.9%	7.42	-4.463***
China	89.9%	4.58	2.665*
Taiwan	82.9%	6.93	
<i>Phoneme manipulation task</i>			
Hong Kong	60.5%	8.64	-4.191***
China	97.5%	5.40	6.633***
Taiwan	76.0%	8.43	
<i>Rhyme judgment task</i>			
Hong Kong	65.0%	4.08	-5.879***
China	86.0%	3.94	0.622
Taiwan	84.0%	9.37	
<i>Reading task of non-words with high frequency rhymes</i>			
Hong Kong	56.3%	4.6	-6.790***
China	71.7%	2.5	-1.716
Taiwan	76.3%	8.1	
<i>Reading task of non-words with low frequency rhymes</i>			
Hong Kong	36.3%	2.4	-7.995***
China	65.4%	3.5	2.660*
Taiwan	57.9%	8.2	
<i>Non-word spelling task</i>			
Hong Kong	23.0%	9.85	-3.897**
China	60.2%	5.13	5.840***
Taiwan	40.0%	9.66	

Notes: * $p < .05$ ** $p < .01$ *** $p < .001$

Phonological awareness tasks

In the phoneme segmentation tasks, *t* test showed that the China group performed significantly better than the Taiwan group, $t(18) = 2.665$, $p < .05$, and that the Taiwan subjects' performance was significantly better than that of the Hong Kong group, $t(18) = -4.463$, $p < .001$. In the phoneme manipulation task, in line with the result of the phoneme segmentation task, the China subjects outperformed the Taiwan subjects, $t(18) = 6.633$, $p < .001$, and the Taiwan subjects outperformed the Hong Kong subjects, $t(18) = -4.191$, $p < .001$. Next we move on to the phonological awareness tasks at the onset-rhyme level. In the rhyme judgment task, the Welch *t* test showed that Hong Kong subjects' performance was again significantly inferior to Taiwan subjects', $t(12) = -5.879$, $p < .001$. However, no significant difference was found between the Taiwan group and the China groups in the rhyme judgment task, $t(12) = 0.622$, $p > .05$.

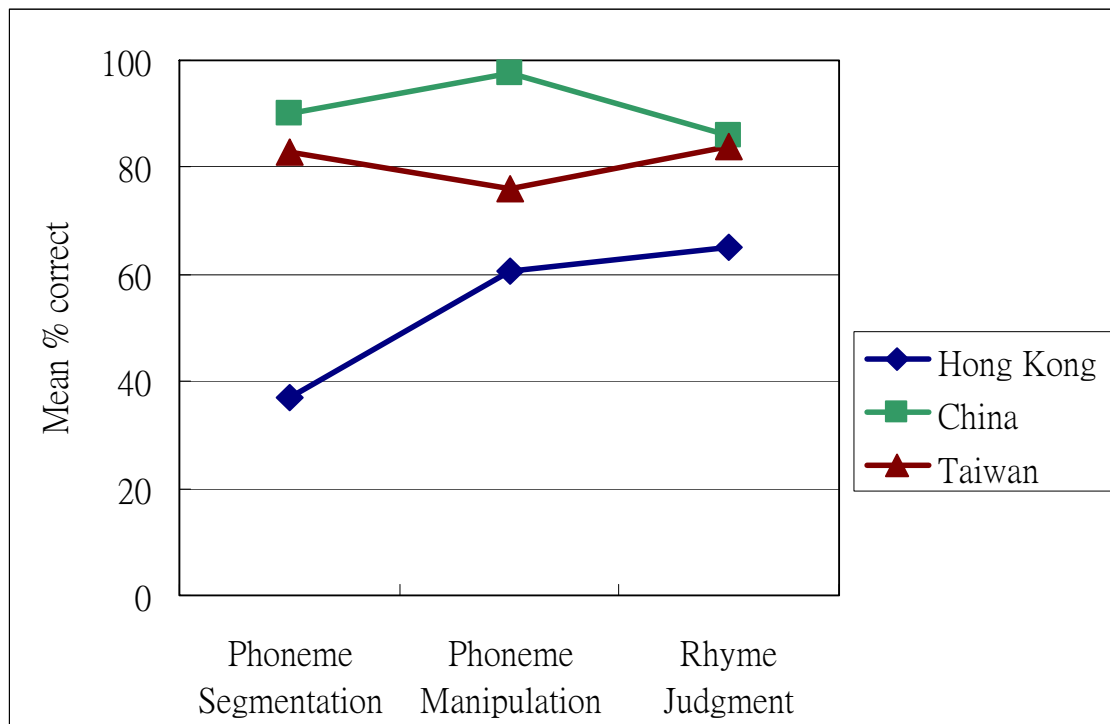


Figure 1 Mean percentage of correct responses in the phonological awareness tasks.

The main finding of the phonological awareness experiment was that in the rhyme judgment task, Taiwan learners of English performed as well as China learners and both of them outperformed the Hong Kong group; however, in the phoneme segmentation and phoneme manipulation tasks, it was found that Taiwan subjects' performance was inferior to China subjects' but superior to Hong Kong subjects' (see Figure 1). The results of the phonological awareness tasks seem to suggest that Taiwan EFL learners' experience with *Zhuyin Fuhao* helps them develop phonological awareness more at the onset-rhyme level than at the phonemic level. To interpret these results, we need to scrutinize the different phonological levels that are represented between *Pinyin* and *Zhuyin Fuhao*. Despite the fact that *Pinyin* and *Zhuyin Fuhao* adopted in China and Taiwan respectively are both the phonetic scripts used to transcribe Mandarin speech, the letters of *Pinyin* represent individual phonemes, and this makes the system truly alphabetic. *Pinyin* learning therefore boosts phonemic awareness and the experience with the *Pinyin* transliteration system would sensitize China learners to segmental features of spoken English words. In contrast to the phoneme-based *Pinyin*, in *Zhuyin Fuhao*, while initials and medials are represented by single symbols, the nucleus and coda components within the rhyme are not separately coded. Therefore, the superior performance on the part of China group in the phonological tasks at the phonemic level may suggest that *Zhuyin Fuhao* does not help Taiwan students develop phonemic awareness as well as *Pinyin* does. Nonetheless, the Taiwan group still outperformed the Hong Kong group in both the phoneme segmentation task and phoneme manipulation task, suggesting that the experience with *Zhuyin Fuhao* does foster Taiwan EFL learners' phonemic awareness, though to a lesser degree compared with the phonological awareness at the onset/rhyme level.

English non-word reading and writing tasks

In the H non-word (non-words with high-frequency rhymes) reading task, the Welch *t* test revealed no significant difference between the Taiwan group and the

China group, $t(10) = -1.716$, $p > .05$. However, for Taiwan and Hong Kong groups, it was found that the Taiwan subjects performed significantly better than the Hong Kong subjects, $t(18) = -6.790$, $p < .001$. In the L non-word (non-words with low-frequency rhymes) reading task, the Welch t test showed that the Taiwan group was significantly inferior to the China group, $t(10) = 2.660$, $p < .05$, but superior to Hong Kong subjects, $t(10) = -7.995$, $p < .001$. In the non-word spelling task, the t test showed that the China group significantly outperformed the Taiwan group, $t(18) = 5.840$, $p < .001$, and Taiwan subjects significantly outperformed the Hong Kong subjects, $t(18) = -3.897$, $p < .01$.

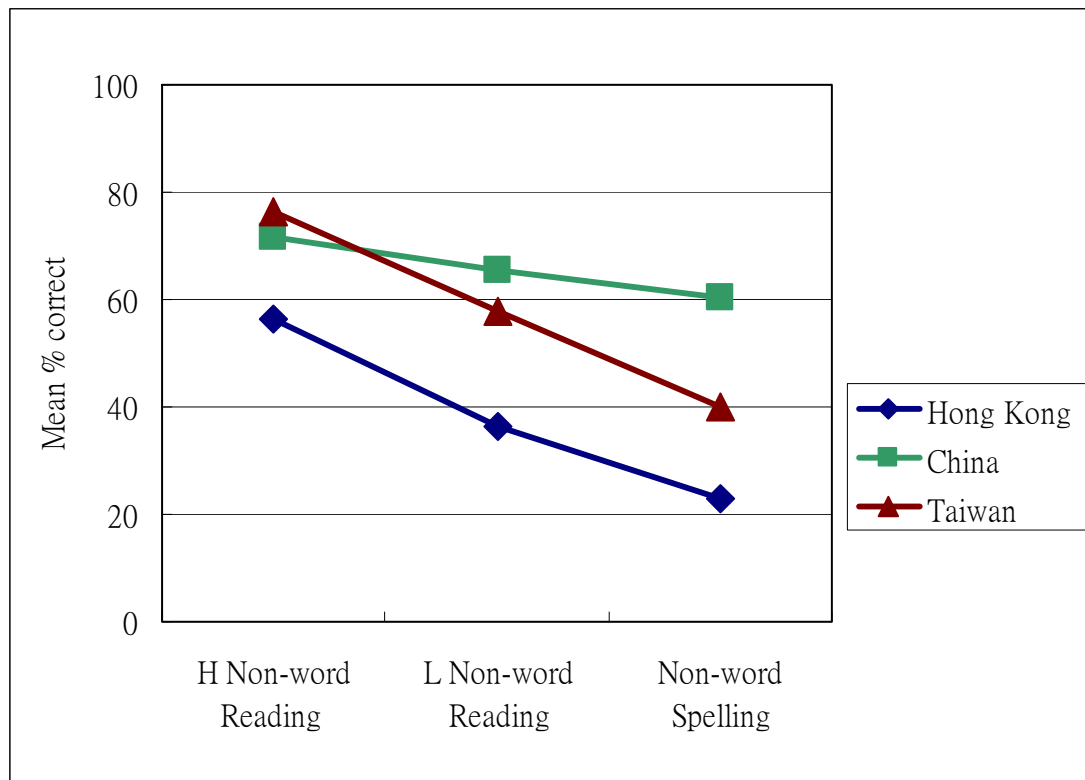


Figure 2 Mean percentage of correct responses in the literacy skill tasks.

The main finding of the literacy tasks is that although Taiwan subjects read H non-words as well as China learners did and both groups outperformed the Hong Kong group, in the L non-word reading task and non-words spelling task which required more sophisticated phonological awareness skills, Taiwan EFL learners were found to be inferior to the China group but superior to the Hong Kong group (see Figure 2). It is worth noting that in the non-word reading task, the correlation between GPC (grapheme-phoneme correspondence) knowledge and L non-word reading was significantly higher than the correlation between GPC knowledge and H non-word reading (Treiman, Gowswami, & Bruck, 1990), because when trying to decode L non-words like *haip*, subjects are forced to use GPC rules due to the fact that few words ends with *aip*. In contrast, when decoding H non-word like *tain*, subjects' phonemic awareness is less important because they are more likely to rely on the easier strategy of making analogies to the rhyme unit of familiar words. Given that the onset-rhyme awareness is closely related to readers' ability to use orthographic rhymes within clue words to read unfamiliar words (Goswami & Mead, 1992), in the present study, it is very likely that Taiwan EFL learners' enhanced

onset-rhyme awareness developed through *Zhuyin Fuhao* helps them to make analogies to the rhymes of known words in assembling the pronunciations of H non-words, and therefore their performance were not found to be inferior to the China group in this task. However, in the L non-word reading task and non-word spelling task, Taiwan subjects had no choice but to use the GPC rules because the stimuli used in the two tasks could not be easily read or spelled through the analogy strategy. Therefore, in this research, Taiwan EFL learners' performance in L non-word reading and non-word spelling tasks were found to be poorer than the China learners' but better than the Hong Kong learners' as a result of the reflection of Taiwan EFL learners' phonemic awareness developed through *Zhuyin Fuhao*. In sum, the results of the study provide further empirical evidence that *Zhuyin Fuhao* does not facilitate English literacy learning as well as *Pinyin* does.

CONCLUSION

In conclusion, the present study yielded valuable information about the effect of *Zhuyin Fuhao* on Taiwan EFL learners' English literacy learning. The results of the phonological awareness tasks and the English literacy tasks provided two major findings. First, Taiwan EFL learners' experience with *Zhuyin Fuhao* did help them develop the phonological awareness that is needed in learning English literacy, though the phonological awareness was enhanced more at the onset-rhyme level than at the phonemic level. Second, the phonological awareness that Taiwan EFL learners had developed through the experience with *Zhuyin Fuhao* in acquiring Chinese literacy did facilitate their English literacy learning, though not as well as *Pinyin* did.

REFERENCES

- Akamatsu, N. (1999). The effects of first language orthographic features on word recognition processing in English as a second language. *Reading and Writing: an Interdisciplinary Journal*, 11, 381-403.
- Bertelson, P., Chen, H. C., & De Gelder, B. (1997). Explicit speech analysis and orthographic experience in Chinese readers. In Chen, H.C. (Eds.) *Cognitive processing of Chinese and related Asian language* (pp.27-46). Hong Kong: Chinese University Press.
- Bialystok, E., Majumder, S., & Martin, M. M. (2003). Developing phonological awareness: Is there a bilingual advantage? *Applied Psycholinguistics*, 24, 27-44.
- Bradley, L., & Bryant, P. E. (1991). Phonological skills before and after learning to read. In Brady, S.A., & Shankweiler, D.P. (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Lieberman* (pp. 37-45). Hillsdale, NJ: Erlbaum.
- Chen, J., & Yuen, C. (1991). Effects of pinyin and script type on verbal processing: Comparisons of China, Taiwan, and Hong Kong experience. *International Journal of Behavioral Development*, 14, 429-448.
- Cheung, H., Chen, H. C., Lai, C. Y., Wong, O. C., & Hill, M. (2001). The development of phonological awareness: Effects of spoken language experience and orthography. *Cognition*, 81, 227-241.
- Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read*. London: Erlbaum.
- Goswami, U., & Mead, F. (1992). Onset and rime awareness and analogies in reading. *Reading Research Quarterly*, 27, 152-162.
- Hamada, M., & Koda, K. (2008). Influence of first language orthographic experience on second language decoding and word learning. *Language Learning*, 58(1), 1-31.
- Holm, A., & Dodd, B. (1996). The effect of first written language on the acquisition

- of English literacy. *Cognition*, 59, 119-147.
- Huang, H. S., & Hanley, J. R. (1995). Phonological awareness and visual skills in learning to read Chinese and English. *Cognition*, 54, 73-98.
- Huang, H. S., & Hanley, J. R. (1997). A longitudinal study of phonological awareness, visual skills, and Chinese reading acquisition among first-graders in Taiwan. *International Journal of Behavioral Development*, 20, 249-268.
- Kuo, C. C. (2006). Phonetic and phonological background of Chinese spoken languages. In C.H. Lee, H. Li, L.S Lee, R.H. Wang, & Q. Huo (Eds.) *Advances in Chinese spoken language* (pp. 33-55). World Scientific Publishing Company
- Leong, C. K., Cheng, P. W., & Tan, L. H. (2005). The role of sensitivity to rhymes, phonemes, and tones in reading English and pseudowords. *Reading and Writing*, 18, 1-26.
- Leong, C. K., Tan, L. H., Cheng, P. W., & Hau, K. T. (2005). Learning to Read and Spell English Words by Chinese Students. *Scientific Studies of Reading*, 9(1).
- Perfetti, C. A., Liu, Y., & Tan, L. H. (2002). How the mind can meet the brain in reading: A comparative writing systems approach. In H. S. R. Kao, C-K. Leong, & D-G. Gao (Eds.), *Cognitive neuroscience studies of the Chinese language* (pp. 36-60). Hong Kong University Press.
- Read, C., Zhang, Y. F., Nie, H. Y., & Ding, B. Q. (1986). The ability to manipulate speech sounds depends on knowing alphabetic writing. *Cognition*, 24, 31-44.
- Shu, H., & Anderson, R. (1997). Role of radical awareness in the character and word acquisition of Chinese children. *Reading Research Quarterly*, 32, 78-89.
- Treiman, R., Goswami, U., & Bruck, M. (1990). Not all nonword are alike: Implications for reading development and theory. *Memory & Cognition*, 18, 559-567.
- Venezky, R. L. (1970). *The structure of English orthography*. The Hague: Mouton.
- Wang, M., Koda, K., & Perfetti, C. A. (2003). Alphabetic and nonalphabetic L2 effects in English word identification: A comparison of Korean and Chinese English L2 learners. *Cognition*, 87, 129-149
- Wang, M., Perfetti, C. A., & Liu, Y. (2005). Chinese-English biliteracy acquisition: Cross language and writing system transfer. *Cognition*, 97, 67-88.

APPENDIX

I. Phoneme Segmentation

l: l	Many :l	Non-words
it	out	og
on	itch	ap
pet	white	zeg
big	cake	lek
swim	ocean	klon
frog	plate	vist
stamp	friend	stelp
robin	whistle	oskad

II. Phoneme Manipulation

Single Letters	One Digraph	Two Digraphs	Clusters
big dog poor teddy	dark ship soft cheese chilly seats short date	chip shop sharp chain	crowd play clown prince

III. Rhyme Judgment

OSR	ODR	OSNR	ODNR
rang / sang	laugh / staff	beast / breast	wait / wet
bush / push	fern/yearn	said/paid	cart / kit
jar / tar	sheep / heap	bear / rear	beg / bag
fate / mate	fought/port	were/where	shop/ship
fine / mine	through / new	wreath / death	fit / fat

OSR: orthographically similar rhyming words; ODR: orthographically dissimilar rhyming words; OSNR: orthographically similar non-rhyming words; ODNR: orthographically dissimilar non-rhyming words

IV. Non-word Reading

More frequent non-word	Less frequent non-word
tain /ten/	goan /gon/
goach /gotʃ/	taich /tetʃ/
joal /dʒol /	soag /sog/
sug /sʌg/	jul /dʒʌl /
vess /vɛs/	fiss /fɪs/
fip /fɪp/	vɛp /vɛp/
chail /tʃel/	paig /peg/
pog /pag/	chol / tʃal /
juck /dʒʌk/	leck /lɛk /
lef /lɛf/	juf /dʒʌf/
foud /fʌud/	choub / tʃʌub /

chob	/tʃʌb/	fod	/fʌd/
vag	/væg/	paf	/pæf/
peef	/pif/	veeg	/vig/
foop	/fup/	haip	/hep/
hain	/hen/	foon	/fun/
jub	/dʒʌb/	vud	/vʌd/
veed	/vid/	jeeb	/dʒib/
cheed	/tʃid/	leem	/lim/
lum	/lʌm/	chud	/tʃʌd/
meep	/mip/	fep	/fɛp/
fesh	/fɛʃ/	meesh	/misʃ/
losh	/lɔʃ/	yol	/yɔl/
yoal	/yol/	loash	/losʃ/

V. Non-word Spelling

dorf	lont
sheve	wump
suts	craid
strecker	pitfair
truckow	baytle
klondly	delsom
tismuppered	strapobees
beartelfoam	soptipal
elinam	kleppering
churtlesneedee	complefation
shepelartoe	diskerpillshine
dramplehoffer	strimperdiction

Concordance-enhanced Vocabulary Instruction on EFL Elementary School Students

Li-tang Yu (余立棠), Hsien-chin Liou (劉顯親)

National Tsing Hua University

cdpsttarng@gmail.com

hcliu@mx.nthu.edu.tw

From the report of a nation-wide English achievement assessment held by National Academy for Educational Research in 2005 in Taiwan, it was revealed that beginning learners of English at primary schools had difficulty in manipulating words and phrases to make sentences. The report also suggested that students need to cultivate the competence of organizing grammatical patterns in a sentence. It is thus crucial to teach the beginning learners the grammatical functions of a word. However, few empirical studies have been conducted to investigate this issue. The purpose of this study is to investigate the effects of concordance-enhanced vocabulary instruction on the beginning learners' productive word use.

Seven fifth graders with varying proficient levels were involved in the study. In order to build their baseline of productive vocabulary knowledge, a time-series design was adopted. A pretest on 63 verbs from their English textbooks was administered to assess the word meaning, usage, and sentence production. Individual unknown 12 verbs out of the pretests were re-assessed to make sure they were unfamiliar to learners. For the treatment period, the participants explored the meanings and grammatical patterns of 12 verbs, which were unknown to them, through consulting a self-designed bilingual concordancer with corpora from English textbooks over four weeks. Afterwards, they received an interview concerning their attitudes toward concordance and took two posttests and one delayed posttest four weeks later after the second posttest completed. It is expected that the findings will help us understand how concordance-enhanced vocabulary teaching facilitates the beginning English learners' productive word use.

INTRODUCTION

Vocabulary knowledge has strong, positive relationships with four language skills and literacy. Particularly at the primary school level in the context of English as a foreign language (EFL), vocabulary, which is considered as a cornerstone of acquiring a new language, is the foundation of English learning. Vocabulary instruction at primary education stage is extremely important.

In 2005, Ministry of Education in Taiwan assessed primary school students' English proficiency level through a nation-wide test on more than fifteen thousand randomly selected sixth graders. The results revealed that Taiwanese beginning learners of English at primary schools had difficulty in organizing words and phrases to make sentences (Research, n.d.). It was also suggested that students need to cultivate the competence of organizing grammatical patterns in a sentence. Accordingly, it is crucial to teach the grammatical functions of a word to the beginning learners.

LITERATURE REVIEW

Web-based concordancers on vocabulary learning and teaching

The effectiveness of integrating technologies in vocabulary instruction on children has been demonstrated by several studies (e.g. Herselman, 1999; Tsou, Wang, & Li, 2002; Tsou, Wang, & Tzeng, 2006). However, the majority of the empirical studies focused primarily on receptive vocabulary knowledge and were limited within a short span of research duration. Few were concerned with productive vocabulary knowledge (e.g. Kang, 1995; Segers & Verhoeven, 2002; Sun & Dong, 2004; Tsou et al., 2002; Shamir et al., 2008). Besides, it is unclear that how long the effects of applying computer technologies on children's vocabulary development can last.

Among various technologies, web-based concordancing has been extensively used for writing and vocabulary learning. Known for its characteristics of data-driven learning (DDL), the online concordancer could provide a context for the words by presenting plentiful examples for learners to inductively discover grammatical patterns of language use from corpus data (Chambers, 2005; Johns, 1994). Instead of playing passive roles in the process of vocabulary learning, learners are expected to play a more active role and assigned to control their learning pace, thus becoming active, autonomous learners like language researchers (Bernardini, 1996).

In Cobb's studies (1997, 1999), it was revealed that college students could improve their amount of academic words and gain lasting and transferable word knowledge in a short period of time via direct corpus-based vocabulary learning. Learners could become more interested in word deep processing via concordancing (Horst, Cobb, & Nicolae, 2005). Moreover, in terms of Sun and Wang's study (2003) and Chan and Liou's study (2005), learners can also benefit from consulting the concordancers by improving collocation knowledge. Kaur and Hegelheimer (2005) also found learning vocabulary with an online concordancer and dictionary did help ESL learners transfer the word knowledge correctly in their writing task. Instructions for learners to realize the importance of words and how to use them productively are essential.

From the viewpoint of an in-service senior high school English teacher, the first researcher in Lee and Liou (2003) integrated web concordancing on receptive and productive vocabulary ability in a regular senior high English curriculum, aiming to investigate whether the senior high school students could enhance their receptive and productive vocabulary knowledge. It was evidenced that the students of lower proficient level could improve their vocabulary knowledge. The gap between the lower level students and higher level students were successfully minimized via concordancing consultation. Furthermore, the concordancing instruction had a better effect on the students who were inclined to inductive learning than on those inclined to deductive learning.

From the empirical studies on concordancer application on vocabulary learning it was supported that a concordancer is a useful tool for assisting learners' vocabulary size growth, receptive and productive lexical abilities, and collocation knowledge. However, little research has been done on how children, the beginning English learners, can benefit from consulting the concordancer on their vocabulary development.

The time-series design

According to Ortega and Ibarra-Shea (2005), one of the best strategies to investigate effects of instruction over a period of time is time-series designs. Mellow, Reeder, and Forster (1996) elucidated the underlying principles of the time-series designs. The time-series designs are known for its reliable measurement instances

used for each subject in a study. The repeated pretests, assessing preintervention behaviors, provide a baseline and the behaviors are able to be compared with postintervention behaviors. If learners' preintervention behaviors are stable and consistent but their postintervention behaviors improve, the researchers are confident in internal validity. Furthermore, a large number of measurement instances in pre- and post-intervention sessions allow researchers to confirm construct validity. The data from time-series design can be analyzed statistically. A small size of subjects is allowed with the time-series designs as well. Hence, detailed description of possible baseline, treatment conditions, characteristics of subjects, and measurement procedures is possible, allowing subsequent researchers to do replications to gain generalizability (Mellow, Reeder, & Forster, 1996).

In line with Mellow et al. (1996) research methodology of time-series designs, Ishida (2004) investigated the effect of intensive recasting on the aspectual acquisition, -te i-(ru), in Japanese as a foreign language. Isida closely documented learners' progress at each instructional session and explored the relationship between recasting instruction and their development. The results showed that the intensive recasting instruction had positive effects on the overall use of -te i-(ru). The number of recasts was positively related with the frequency in accurate use of -te i-(ru). From the outcome of the delayed posttest, the instructional effects were shown to last. More replications with a larger number of participants were needed for generalizability of the recasting effects and the prototypical aspect order (Ishida, 2004).

METHODOLOGY

To fill the void of little research on concordancing instruction to primary school students, the purpose of this study was to investigate the effects of concordance-enhanced vocabulary instruction on primary school students in the framework of the time-series design. After the students' consultation of concordancing, their development of verb-phrase productive knowledge, operationalized as sentence production with appropriate structural patterns of the verb phrases, was examined.

An English-Chinese concordancer specifically for Taiwanese primary school students was developed. The URL of the concordancer is <http://140.114.75.17>. There were around 46730 words in 10078 sentences as a corpus of the concordancer, modeled from primary schools English textbooks censored by Ministry of Education (MOE). The Chinese translations were provided so that the students could better comprehend the sentence meanings and further compare the structure difference between Chinese and English (Wang, 2001).

Participants

The participants in the study were seven grade five students in an intact class at a public, regular elementary school in the suburban zone of Hsinchu city, northern Taiwan. They were all native speakers of Mandarin-Chinese and learners of English as a foreign language.

The participants had at least 4-year experiences of learning English as a subject in the school. They were all qualified to the following criteria to take part in the study. First, they had acquired the rule of English sound-letter correspondence and had a certain amount of vocabulary. Then, due to characteristics of time-series design, the duration of the study was taken place over 6 months. The participants had to be willing to attend the whole research process. Lastly, the number of target words set in the study was 12. If the result of pretests showed that the number of words they did not understand was less than 12. They would be ruled out from the study.

Instruments

Several instruments were used in the study to collect data.

Background questionnaire

The participants were required to complete a background questionnaire at the beginning of the study. The background questionnaire was designed to obtain information concerning the participants' backgrounds and learning experience, such as their English learning conditions and the use of computers and the Internet.

Tests

There were four types of tests in the study: three pretests, twelve treatment tests, two posttests, and one delayed posttest. As shown in the following table, the format in these tests was the same, including three categories of word knowledge: word meaning, word usage, and sentence making that illustrates the pattern. The participants were required to fill in the categories of their own target words. In vocabulary learning, the mapping of lexical forms to meanings is one of the most integral tasks. It is essential to add this part, "word meaning," in the test. Moreover, through inductively discovering the grammatical patterns of the target words via the concordancing consultation, learners could record their findings in "word usage," and make a sentence with the target words with regard to the grammatical patterns they induced.

Table 1. *A sample form of a target word in the test task*

feed	Word meaning: 餵
	Word usage: 人 + feed + animals + 時間
	Sentence production: I feed the fish on Sunday.

Interview

After the concordancing treatment, the researcher would interview the seven participants in Chinese in terms of their viewpoints toward using concordancing on their vocabulary learning. Furthermore, it is interesting to explore whether they would subsume their improvement or decline in productive vocabulary knowledge performed in the posttests into the concordancer consultation.

Data collection procedure

The data collection procedure involved three phases: a pretest period, a treatment period, and a posttest period. The purpose of administering the pretest was to identify the participants' unfamiliar verb phrases before using concordancing in the treatment period. In the pretest period, the participants took three pretests to investigate their understanding on verb phrases over a period of 16 weeks. 51 verb phrases occurred in their grade 5 and 6 English textbooks and 12 distracters from junior high school English textbooks were taken as test words. Through the pretests, 7 individual 12-word lists were produced as individual participants' target words in this study. The intervention lasted for twelve sessions across four weeks during the treatment period. In each session, the participants firstly had hands-on experience of concordancing for the search of their target words of the day. They were required to make notes about the target word. When they thought they had learned the target word of the day, they turned off the computer screen, taking translation practices which were taken from the concordancing output. Once they finished the translation practices, they turned on the screen again to check the English. If time was enough, they could consult the previous target words to retrieve the understanding of the words. Afterwards, they took the treatment test to write down the meaning, the usage of the word and a sentence that

illustrates the pattern. In the posttest period, the purpose of the posttest was to examine whether the effect of concordance-enhanced vocabulary instruction could retain. Two posttests after the treatment period and one delayed posttest four weeks after the second posttest will be involved in the posttest period. The researchers will also interview the participants to explore their viewpoint of concordancing on vocabulary learning.

Data analysis

Both quantitative and qualitative analyses were administered to explore the effectiveness of the concordance-enhanced vocabulary instruction. The data collected include the participants' background questionnaire, three types of tests (the pretests, the treatment tests, and the posttests), and one interview. Quantitative analyses, a parametric t-test, will be conducted on data from the pretests, the treatment tests, the posttests, and the delayed posttest. The scoring method is displayed in Appendix A. To evaluate the immediate effects of instruction, the mean of scores of the target words in the pretest period and the treatment period was compared. To evaluate the retention of concordance-enhanced vocabulary instruction, the mean of the scores of verb phrase use in the treatment period and the posttest period will be compared. Data collected from their background questionnaire and the interview would be qualitatively analyzed to explore their learning background and perspectives into concordancing on vocabulary learning.

PRELIMINARY RESULTS

So far the progress of the study has been done at the tenth session in the treatment period. Thus, the findings are preliminary. From the response of the participants' questionnaire, it was revealed that their English learning background and proficiency levels varied. The only common characteristics were that all participants could have access to the Internet and they were all technical-skilled. After the three pretests, every participant obtained their individual 12-word list which was completely unknown to them, as displayed in Table 2. The first six words were common to all participants; the latter ones were different.

Table 2 *Individual participants' target words*

Participants' code	A	B	C	D	E	F	G
1	exercise	exercise	exercise	exercise	exercise	exercise	exercise
2	meet	Meet	meet	meet	meet	meet	meet
3	remember	remember	remember	remember	remember	remember	remember
4	sound	sound	sound	sound	sound	sound	sound
5	surf	Surf	surf	surf	surf	surf	surf
6	taste	Taste	taste	taste	taste	taste	taste
7	buy	Cut	brush	buy	brush	feed	buy
8	forget	point	feed	call	call	forget	feed
9	row	Start	forget	feed	start	start	forget
10	smell	throw	paint	smell	sweep	sweep	start
11	think	Wait	row	sweep	walk	think	sweep
12	walk	Walk	wear	wear	wear	wear	think

The development of the participants' productive vocabulary knowledge was shown in Figure 1. The first three sessions were the pretests. All participants received no scores

of the target words in the pretests. Since the fourth session, the treatment started. The participants' performance in the treatment tests was not in chorus.

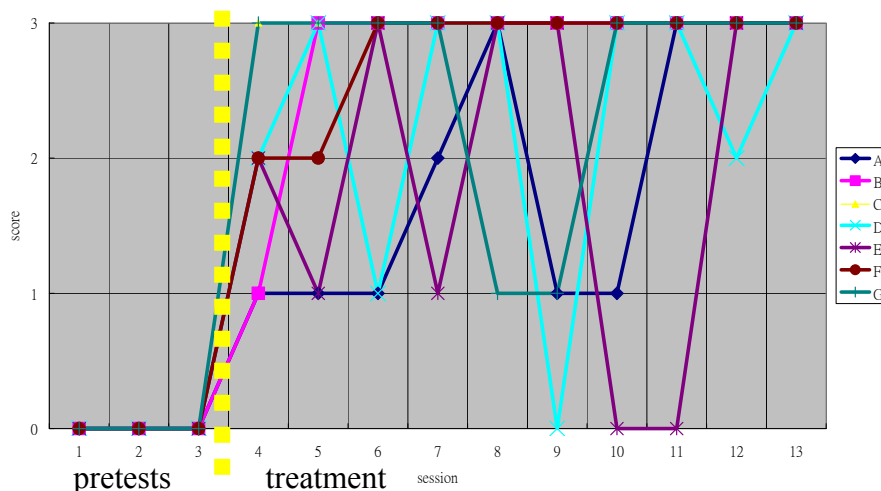


Figure 1 The development of the participants' productive vocabulary knowledge

As displayed in Table 3, the mean of the participants' productive vocabulary performance on the treatment tests was much better than on the pretests. The analysis of a parametric paired t-test on the mean of scores between the pretest period and the treatment period indicated a statistically significant effect on the participants' productive vocabulary knowledge ($\alpha=.05$, $t= -14.66$, $p=.00$).

Table 3 *Descriptive analysis of the participants' performance on the pretests and the treatment tests*

	average	
	Mean	SD
Pretest	.00	.00
treatment	2.47	.45

DISCUSSION AND CONCLUSION

Through the time-series design, a solid baseline of the participants' productive vocabulary knowledge was built by means of the three pretests. In the pretest, It was demonstrated that the participants had no understanding of their target words. After consulting the concordancer, they had immediate treatment tests to assess whether they learned something from the concordancing output. In Figure 1, it was clearly recorded that the students' vocabulary development was not linear. Given the 10-session treatment observation, the learners' development routes could be different from each other. Some were fluctuating drastically in the cases of Student D and Student E, but the route was also developed gradually stably upward like Student C and Student F. The average score of each target word for each participant was nearly 2.5 out of 3. It is expected that the participants could induce the structural patterns of the target words from hands-on concordancing experience and discover the differences between English and Chinese structures from the bilingual parallel corpus. Besides, from the links to the online dictionaries and the bilingual corpus, the participants could have a better understanding of word sense.

Overall, it seems that the students could benefit from consulting the concordancer to learn vocabulary meaning and pattern. In addition, they could advance their

knowledge into making sentence with the word they learned. Concordance-enhanced vocabulary instruction could be administered in primary school English education.

ACKNOWLEDGEMENTS

The paper is sponsored by projects, NSC95-2520-S-007-002-MY3 and NSC96-2411-H-007-033-MY3 under National Science Council, Taipei, Taiwan R. O. C.

REFERENCES

- Bernardini, S. (1996). Corpora in the classroom: an overview and some reflections on future developments. In J. M. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 15-36). Philadelphia: John Benjamins.
- Chambers, A. (2005). Integrating corpus consultation in language studies [Electronic Version]. *Language Learning & Technology*, 9, 111-125. Retrieved July 20, 2008, from <http://llt.msu.edu/vol9num2/chambers/default.html>
- Chan, T.-P., & Liou, H.-C. (2005). Effects of web-based concordancing instruction on EFL students' learning of verb-noun collocations. *Computer Assisted Language Learning*, 18(3), 231-250.
- Cobb, T. (1997). Is there any measurable learning from hands-on concordancing? *System*, 25(3), 301-315.
- Cobb, T. (1999). Breadth and depth of lexical acquisition with hands-on concordancing. *Computer Assisted Language Learning*, 12(4), 345-360.
- Herselman, M. E. (1999). South African Resource-deprived learners benefit from CALL through the medium of computer games. *Computer Assisted Language Learning*, 12(3), 197-218.
- Horst, M., Cobb, T., & Nicolae, I. (2005). Expanding academic vocabulary with an interactive on-line database [Electronic Version]. *Language Learning & Technology*, 9, 90-110. Retrieved July 31, 2008, from <http://llt.msu.edu/vol9num2/horst>
- Ishida, M. (2004). Effects of recasts on the acquisition of the aspectual form -te i-(ru) by learners of Japanese as a foreign language. *Language Learning*, 54(2), 331-394.
- Johns, T. (1994). From printout to handout: grammar and vocabulary teaching in the context of data-driven learning. In T. Odlin (Ed.), *Perspectives on pedagogical grammar*. Cambridge: Cambridge University Press.
- Kang, S.-H. (1995). The effects of a context-embedded approach to second-language vocabulary learning. *System*, 23(1), 43-55.
- Kaur, J., & Hegelheimer, V. (2005). ESL students' use of concordance in the transfer of academic word knowledge: an exploratory study. *Computer Assisted Language Learning*, 18(4), 287-310.
- Lee, C. Y., & Liou, H. C. (2003). A study of using web concordancing for English vocabulary learning in a Taiwanese high school context. *English Teaching & Learning*, 27(3), 35-56.
- Mellow, D., Reeder, K., & Forster, E. (1996). Using time-series research designs to investigate the effects of instruction on SLA. *Studies in Second Language Acquisition*, 18, 325-350.
- Ortega, L., & Ibarra-Shea, G. (2005). Longitudinal research in second language acquisition: recent trends and future directions. *Annual Review of Applied Linguistics*, 25, 26-45.
- Research, N. A. F. E. (n.d.). 2005 年小六學生國語、英語、數學評量結果報告

- [Electronic Version], 1-20. Retrieved October 13, 2008, from <http://tasa.naer.edu.tw/report.htm>
- Segers, E., & Verhoeven, L. (2002). Multimedia support of early literacy learning. *COMPUTER & EDUCATION*, 39, 207-221.
- Shamir, A., Korat, O., & Barbi, N. (2008). The effects of CD-ROM storybook reading on low SES kindergarteners' emergent literacy as a function of learning context. *Computer & Education*, 51, 354-367.
- Sun, Y., & Dong, Q. (2004). An experiment on supporting children's English vocabulary learning in multimedia context. *Computer Assisted Language Learning*, 17(2), 131-147.
- Sun, Y.-C., & Wang, L.-Y. (2003). Concordancers in the EFL classroom: cognitive approaches and collocation difficulty. *Computer Assisted Language Learning*, 16(1), 83-94.
- Tsou, W., Wang, W., & Li, H.-y. (2002). How computers facilitate English foreign language learners acquire English abstract words. *Computer & Education*, 39, 415-428.
- Tsou, W., Wang, W., & Tzeng, Y. (2006). Applying a multimedia storytelling website in foreign language learning. *Computer & Education*, 47, 17-28.
- Wang, L. (2001). Exploring parallel concordancing in English and Chinese [Electronic Version]. *Language Learning & Technology*, 5, 174-184. Retrieved July, 18, 2008, from <http://llt.msu.edu/vol5num3/wang>

APPENDIX: Scoring methods

Aspects	scores	description
meaning	0	Wrong, inappropriate definition
	1	Complete, appropriate, and accurate.
pattern	0	Wrong pattern Verbs are not used as the verb function
	1	Complete, appropriate, and accurate.
sentence	0	The produced sentence is not related to the pattern.
		The verb is not functioned as "verb."
		Incomplete sentences can not be recognized.
1	Complete sentence with clear meaning in correspondence with the pattern..	
* Word meaning is evaluated first, followed by pattern, and finally, sentence. Without receiving any scores in meaning, no score for pattern. If no score is given to pattern, no score is for sentence.		