The Role of Social Presence in Learner-centered Communicative Language Learning Using Synchronous Computer-mediated Communication: Experimental Study
Masanori Yamada, The University of Tokyo mark@mark-lab.net

Abstract
This study aimed to clarify the relationship between media, learners’ perception of social presence, and output in communicative learning using synchronous computer-mediated communication (SCMC). In this study, we developed four types of SCMC: videoconferencing (image and voice), audioconferencing (voice but no image), text chat with image (image but no voice), and plain text chat (no image and no voice). Each system allows learners to be conscious of and utter a target formulaic expression. I investigated the effect of each system on psychological perception and productive output as well as the relationship between perception and output. The results show that image and voice promote consciousness of natural communication and relief, while a text-mediated system enhances confidence in grammatical accuracy. In order to clarify the relationship between media, affective side, and output, path analysis was conducted using SPSS Amos 7.0. The results indicated that voice communication strongly affects both learners’ affective side and output. The existence of a partner’s image enhances the consciousness of natural communication, which leads to a number of self-corrections, an aspect of learning performance. However, voice communication has a negative effect on confidence in grammatical accuracy.

Keywords: Synchronous computer-mediated communication, Computer assisted language learning, Social presence, Communicative approach

1. Introduction

1.1 Language learning with CMC use

With advancements in information technology, there has been a growing interest in the use of computer networks for second language acquisition (SLA). In recent times, network technologies such as computer-mediated communication (CMC) have enabled teachers to offer Internet-based collaborative learning in SLA. CMC has two types: Synchronous CMC (SCMC) and Asynchronous CMC (ACMC).

Hilz & Goldman (2005) show the advantages of ACMC use in educational settings, such as allowing learners to study anytime and anywhere, allowing learners to study at
their own pace, promoting reflective learning, and giving learners time to collect their thoughts before posting. However, Weller & Mason (2000) pointed out that the advantages mentioned above can be disadvantages in online learning. For example, since learners are not required to study at a specific time, other demands on their time can take precedence. While the recording of posts in ACMC allows learners to learn through reflection, some learners hesitate to post when they lack confidence in the validity of their ideas or opinions. And it seems to take considerable time to comprehend the overall flow of communication in ACMC due to complex interchanges and threads of posts. Levy & Stockwell (2006) pointed out that ACMC lacks social cues, which tends to demotivate learners from involvement in communicative language learning. With respect to language learning, an advantage of ACMC is the ability for reflection, which promotes self-correction and allows the learner to consider their idea with consciousness of grammatical accuracy (White, 2004; Lamy & Hampel, 2007).

On the other hand, it has been suggested that SCMC is effective in instruction concerning communication skills, because, in SLA, its forms such as text chatting offer an environment similar to that in face-to-face communication (Blake, 2000). SCMC enables learners to communicate in an environment that has real time settings, which motivates learners to communicate with each other in the second language. Therefore, CMC and other interactive media can be used to promote learning (Furstenberg, 1997; Warschauer, 1997). Moreover, previous research has indicated a positive effect of the use of SCMC (e.g., Hampel, 2003; Rosell-Aguilar, 2006; Wang, 2004). SCMC promotes more equal participation than face-to-face communication does in discussions in the second language (Chun, 1994, 1998; Warschauer, 1996). Language learners who have studied using SCMC outperform those who have studied using ACMC and without CMC with respect to the amount of speech in face-to-face discussion (Abrams, 2003). In addition, compared to the case in regular classrooms, SCMC enhances task-based communication such as discussions, due to the combination of the rapid nature of communication exchange; linguistic effects such as the amount of speech; cognitive effects, which promote continuous communication through a communication strategy (Tarone, 1981); and emotional effects such as shy learners’ increased participation in discussions (Beauvois, 1998a, 1998b). Language learners use communication devices in SCMC as well as in face-to-face communication (Lee, 2002, 2004; Smith, 2003). In particular, SCMC use in task-based communication is effective in promoting the use of communication strategies such as negotiation of meaning (Smith, 2002) and feedback for repairing lexical and syntactic errors (Morris, 2005). CMC enables learners to speak with
reduced anxiety (Kelm, 1992). In a similar study, learners who were trained using SCMC performed better on an oral test than those who were trained using regular classroom instruction (Beauvois, 1994). Beauvois (1994) suggested that a positive attitude to technology, a low-stress environment, and anonymity, which allows learners to keep personal information such as race, gender, and timidity confidential, leads to a positive effect in oral performance in SLA. These features facilitate the acquisition of communication skills in a second language. These positive effects promote interaction between learners, which many researchers regard as one of the most important skills in communication (e.g., Long, 1981, 1989; Gass et al., 1989).

SCMC also has some disadvantages in communicative language learning (Lamy & Hampel, 2007). For example, they suggest that high-proficiency learners utter more than low-proficiency learners; learners tend to utter short sentences, which lead to inadequate output for language acquisition; and learners feel considerable pressure to utter rapidly. Moreover, SCMC is not flexible in terms of time (Levy & Stockwell, 2006); since learners engage in “live” communication with their partners, they must schedule a specific time for study.

Recent advances in technology have created a new type of SCMC using videoconferencing, enabling interlocutors to feel others’ presence to a much greater degree than in text-based communication. Several studies have explained the effects of such types of CMC in language learning. Videoconferencing helps learners to eliminate physical barriers, motivates them to speak in the second language (McAndrew et al., 1996), and enables them to use communication devices such as eye gazing and gestures for understanding other people (Bruce, 1996). In task-based language learning, videoconferencing can improve performance in collaborative learning (Zähner et al., 2000). However, it was suggested that the practical use of IT-enhanced CMC in SLA has not yet been considered (Wang, 2004). In particular, instructional design should consider the features of IT-enhanced CMC, including face-to-face instruction, for effective learning.

1.2 Theoretical background of SLA

With respect to SLA, as mentioned above, previous work suggests that CMC is effective in communicative language learning, because it promotes social interaction such as negotiation of meaning between learners, and facilitates comprehensive input and output. Interaction and comprehensive input and output seem to play an important role in
language learning. The importance of these factors in classroom-based communicative instruction has been verified by many previous studies.

Comprehensive input implies written or spoken information in the target language that the learner can comprehend (e.g., Krashen, 1985; Gass et al., 1998). Interaction is based on comprehensive input. In SLA, communication skills, in particular, are acquired through communication between participants, for example, between learners and teachers (e.g., Long, 1981, 1989).

Interaction refers to meaningful communication that enables the understanding of and stimulates the production of comprehensive input. When people face a problem such as misunderstanding the other person, they prefer to resolve the problem before continuing communication (Clark, 1994). For example, when learners cannot understand their interlocutor’s utterance, their interlocutor may modify or paraphrase it to facilitate the learners’ understanding, or the learners may ask their interlocutor to repeat it. Learners learn communication skills through the production of comprehensive input in interaction.

Output refers to the learning activities performed in language education. Learners need to perform learning activities such as uttering, repeating, or writing, because they produce comprehensive input through interaction (Swain, 1985, 1995). Swain (1995) claimed that output has three functions: noticing the gap between what the learner can and cannot express, hypothesis testing such as the trial-and-error method, and metalinguistic functions such as reflective learning.

A communicative approach is effective in fostering communication skills by combining the three functions mentioned above. However, in an interactive classroom setting, it is difficult to consciously make learners aware of the learning objectives. In general, learning objectives are not described clearly in communicative task-based instruction because evaluation criteria are concerned with task accomplishment and the outcome of communication rather than the fluency and accurate form of learners’ utterances (Ellis, 2003). In SLA research, one common issue is how to raise learners’ consciousness of target language forms in communication tasks. Previous studies have suggested the effectiveness of tasks that increase the consciousness of grammar communication (Fotos & Ellis, 1991; Fotos, 1994). The significance of these studies was based on the importance of learners’ awareness of communicative instruction. Some studies report that in second language learning, learners cannot be aware of learning objectives without being conscious of them (e.g., Schmidt, 1990, 2001). The next section explains this point in detail.
1.3 Learning consciousness and motivation

The recent advances in information technology have provided people involved in the field of education with the opportunity to use a variety of educational methods such as e-learning, which is a useful educational method in both online and face-to-face communication and is used across the world. However, e-learning clearly seems to have two major drawbacks: the absence of the instructor and the lack of motivation, both of which seem to cause drop-out in e-learning. The absence of the instructor can cause two problems. One is to reduce learners’ consciousness of the necessity of study; since learners don’t need to go to a face-to-face classroom for their study and can complete their courses by networked computer, they can give their personal issues priority, thus failing to study (Bersin, 2004). The other is concerned with feedback; the absence of an instructor can lead to the decline of appropriate and immediate feedback (Lou et al, 2003; Butler & Winne, 1995). Therefore, a system which raises learning consciousness is needed in an instructorless learning environment.

Motivation is a central issue for learning (e.g., Bersin, 2004; Schunk & Zimmerman, 2007, Van Lier, 1996). Highly-motivated learners can be high performers, because they recognize their learning purposes and they can proceed in their studies by themselves (e.g., Pintrich et al, 1993; Zimmerman & Martinez-Ponz, 1986). Learners must be self-regulated able to recognize their ability, make learning plans, and use suitable learning strategies for high performance (Matsuda & Harada, 2007). In language learning, various strategies for motivating learners have been suggested, such as collaborative learning in learner-centered communication and giving authentic learning materials from the viewpoints of socio-culture and cognitive psychology (e.g., Dörnle, 2001; O’Malley & Chamot, 1990; Oxford, 1993).

In order to solve the problems above, blended learning with CMC and face-to-face lectures can be practical and effective for learning (Bersin, 2004; Jochems et al, 2003; Matsuda & Harada, 2007). In such blended learning, online learner-centered study should motivate learners to study but should also be learner-directed. It has been suggested that learner-centered instruction may promote the negotiation of meaning and increase motivation in language learning (e.g., Pica & Doughty, 1985; Fernandez-Garcia & Martinez-Arbelaez, 2002). On the other hand, the discourse between teacher and learner is more grammatical than that in learner-centered study (Pica & Doughty, 1985). Therefore, it seems that learners do not study accurate speech in learner-centered communication because they are not conscious of the learning objectives; the absence of
a teacher prevents learners from understanding what they have to learn and do in learner-centered instruction. Some studies report that in second language learning, consciousness of the learning objective can be necessary for learners to learn a second language (e.g., Schmidt, 1990, 1993, 2001). Many previous studies (e.g., Abrams, 2003; McAndrew et al, 1996) have compared the features of learner-centered communication performed using various communication media; however, the effectiveness of a particular medium with respect to the consciousness of learning objectives and the requirements for effective communicative language learning have not been discussed. Although making communication as realistic and natural as possible is a major challenge in computer-assisted language learning (CALL) (Bax, 2003), it also has practical value in terms of making learners conscious of learning objectives even in learner-centered communication. It is important to design instruction that raises the consciousness of learning objectives and maintains a high level of motivation without the teacher necessarily being in front of the learner (Wang, 2004). The relationship between system design, learners’ affective side and output will be helpful for researchers and practitioners to design, develop, and use SCMC in their research and classroom.

1.4 Social presence and language learning

The effect of CMC in language learning should be considered from the viewpoint of the psychological factor. Saraberry (2000) pointed out that the new paradigm of CMC use in language learning is unclear due to the lack of evaluative study from the viewpoints of psychological factors and theories. Through her practical research, Hampel (2003) also insisted on not only the viewpoint of language learning but also the significance of the psychological view in the evaluation based on learning theory.

In light of the spread of e-learning, one of the useful viewpoints of evaluation in CMC use in learning is social presence. Learners’ perception of presence is affected by social presence, which Short, Williams, and Christie (1976) described as the “degree of salience of the other person in the interaction and the consequent salience of the interpersonal relationship,” that is, the perceived proximity to real time communication in face-to-face settings. Short, Williams, and Christie (1976) suggested that the two factors which promote social presence are “immediacy,” which is the psychological proximity of the interlocutors and “intimacy,” which is the perceived familiarity caused by social behavior such as eye gazing, nodding, and smiling. White (2004) suggested that interactive learning is more effective than one-way (noninteractive) learning; however, a
relevant concern is the connection between learning and interaction. Social presence is a significant concept for considering the method of connecting interaction to learning. Social presence is an important factor for promoting learning in distance learning (McIsaac & Gunawardena, 1996) and is considered to be emotionally effective. It enhances learners’ satisfaction with learning (Gunawardena & Zittle, 1997).

In asynchronous CMC (e-mail), social presence motivated learners and promoted interaction such as requests for help (Leh, 2001). Social cues such as nodding, smiling, and gestures facilitated effective learning in interactive television settings (Hackman & Walker, 1990).

In traditional text-based CMC, which lacks social cues, learners tend to increase their social presence in continuous communication by expressing their emotions in ways such as using emoticons. Derks et al. (2008) suggests that nonverbal devices have social meaning, for example, feelings, and transmit it from person to person. Social presence plays an important role in this transmittance. An invisible situation has a special influence on realizing and understanding an interlocutor’s feelings, due to the absence of social cues in receiving messages from the interlocutor (Derks et al., 2008). In some types of SCMC, such as audioconferencing and text chat, learners cannot use social cues such as eye gazing and nodding; as a result, they “are not aware when one person starts to type a message and may continue with a topic, or else may change the direction of the discussion while a potential contributor to the discussion types his or her message” (Levy & Stockwell, 2006).

Learners can also increase their social presence in text-based CMC through the community created by learners during communication or with the teacher’s effort (Gunawardena, 1995). Moreover, social presence has a greater effect on an active community with increased frequency of interaction between learners; as a result, it promotes learners’ engagement in communicative learning (Polhemus et al., 2000). Lomicka and Lord (2007) also pointed out that social presence affected an active community; in particular, social presence concerning interaction was the most frequent. It led to better performance than that in individual learning in terms of received responsibilities and precise description skills. Lomicka and Lord (2007) suggested that social presence enhances the interaction between learners, which, in turn, affects learning performance.

Broadband network technology is capable of offering a new type of SCMC involving the use of multimedia, audioconferencing, and videoconferencing; it appears that such media might be more effective for learning due to the availability of social cues (e.g.,
Hackman & Walker, 1990; McIsaac & Gunawardena, 1996; Gunawardena & Zittle, 1997; Zähner et al., 2000). Social cues such as smiling and gestures can support effective learning in interactive TV situations (Hackman & Walker, 1990). This study aimed to illustrate the effect of videoconferencing software as an interactive language learning tool on the learners’ affective side and output from the viewpoint of social presence, which is one of important psychological factor in using CMC in language learning (e.g., Garrison et al, 2003; Lamy & Hampel, 2007).

1.5 Research Objectives

In this study, we aimed to investigate the issues related to the application of videoconferencing software in SLA. For this purpose, we analyzed the following points:
1: the effect of the communication medium on learners’ affective side
2: the effect of the communication medium on learning performance
3: the relationship among the affective aspect of learning, communication media, and learning performance

For this study, we used four types of web-based software (videoconferencing software, audioconferencing software, text chat with images, and plain text chat) that enable learners to be conscious of their SLA objectives through learning tasks in learner-centered instruction. Using the software, we evaluated and discussed the media’s influence on affective learning, output and the relationship between media, learners’ affective side and output.

2 Experiment

2.1 System

For this study, four SCMC software systems were re-developed, based on videoconferencing software developed in a previous study (Yamada et al, 2008): videoconferencing, audioconferencing, text chat with image, and text chat without image. This system allows learners to be aware of and utter target formulaic speech during second language communication. Using this software, we evaluate the effect of each system from three viewpoints above. Figs. 1, 2, 3, and 4 display the interface of each system.
2.2 Participants

The participants in this study were 40 university students (females: 9, males: 31, age range: 18–35, mean age: 22.1). The participants did not know each other prior to the experiment. All the participants were non-native speakers of English. The participants’ proficiency in English varied from the highly proficient level (students who had participated in international conferences and had interacted with international students in their daily life) to the low level (students who needed help to understand others’ utterances); however, all the participants had reached at least the high school standard level in grammar and vocabulary, and had TOEIC (Test of English for International Communication) C level scores (C level: score range 470 – 730; learners’ average: 582.5).

A learner at the TOEIC C level “has sufficient knowledge for daily activities and conducting business within certain limits,” and “has acquired fundamental knowledge of grammar and structure and has the vocabulary to communicate essential information, even when lacking expressive power” (Educational Testing Service, 2008). Computer literacy among all the students was high, and all of them used computers daily for e-mail, text chatting, surfing, writing reports, etc.

2.3 Procedure

The participants were randomly divided into four groups: the videoconferencing group, the audioconferencing group, the text chat with images group, and the plain text chat group. The groups were selected by lottery, with each participant drawing a piece of paper with a media name written on it from a sack. After listening to the instructions on using the system and tasks prior to the experiment, each participant moved to his/her room. A laptop with a webcam and headset was set in each room, and all the computers were connected to a LAN. The participants were divided into pairs randomly by lottery, and each pair engaged in a learner-centered discussion for 15 minutes; all the pairs were given the same topic—choosing a new schoolteacher from among four candidates, taking into consideration certain given conditions (a decision-making task). Information about and the background of the school and candidates were provided in the learning material displayed in each system (see the explanation of the system interface above). Each pair tries to make a decision through communication; in voice-based communication, learners exchange their opinions and argue by voice; in text chat, they discuss this theme by text-based communication. As mentioned above, each pair consisted of participants who had not met before because familiarity between participants may have an influence on
communication (e.g., the participants slipping back into their native language) and evaluation (e.g., participants may be affected by their friends’ thoughts when they evaluate the system). From the educational viewpoint, although familiarity is an important factor in educational performance, we controlled this condition by eliminating this factor, because this study aimed to investigate the effect of systems that enhance social presence on second language learning issues. Finally, the participants were asked to answer a questionnaire.

2.4 Data Collection

This study aimed to investigate the contribution of subjective evaluation to learning performance. Data was collected in two ways. The first was by means of a questionnaire. All the participants were required to answer a questionnaire after the experiment. The questionnaire required all participants to rate (on a 6-point rating scale) their perceived presence of the interlocutor, the perceived ease of communication, and the perceived consciousness of SLA while communicating using each type of SCMC. The questions asked to the participants are displayed in Table 1.

The second method of data collection was video-recording. To ensure that the research was objective, the entire communication was video-recorded, and the items listed in Table 2 were calculated for each participant.

2.5 Path Analysis

In order to see the relationship between communication media, psychological perception, and learning performance, path analysis was conducted using each perception as an independent variable and language learning performance metrics as dependent variables. In order to do path analysis, the 15 questionnaire items were categorized into three groups following the results of factor analysis in a previous study which used the same questionnaire items (Yamada & Akahori, 2007). Yamada & Akahori (2007) extracted four factors as the result of factor analysis. However, this study employs three factors, because factor four in their research has only one questionnaire item and explains only five percent of the variance, which indicates that this factor does not have enough influence to explain learners’ perceptions.

In Yamada & Akahori (2007), factor 1, named “perceived ease of communication in English due to the interlocutor’s presence” contains questions #3 “Rate your perceived ease of understanding your partner’s utterance”, #5 “Rate your perceived consciousness of your partner’s presence”, #6 “Rate your perceived relief of communicating using
SCMC”, #7 “Rate your perceived ease of communicating in English”, and #8 “Rate your perceived feeling of the similarity between face-to-face and SCMC”. Factor 2 was named “consciousness of natural communication”. It includes questions #1 “Rate the frequency of your utterances while using SCMC”, #2 “Rate your perceived ease of initiating communication”, #4 “Rate your perceived ease of saying what you wanted to say”, #10 “Rate your perceived consciousness of responding as soon as possible”, and #13 “Rate your perceived consciousness of the grammatical accuracy of your partner’s utterances”. Factor 3, “confidence in grammatical accuracy”, contains questions #14 “Rate your perceived consciousness of the response speed of your partner”, #9 “Rate your perceived consciousness of accuracy in English communication”, and #11 “Rate your perceived consciousness of communicating the desired meaning in English, even if you made a grammatical mistake”. Following these results, the responses to each item were added up by factor to calculate each factor’s score.

3. Results

3.1 Determining Factors

Thirteen of the fifteen items were used for determining three factors following Yamada & Akahori (2007). In order to judge the validity of each factor, cronbach alpha for each factor was calculated (factor 1: 0.885, factor 2: 0.727, factor 3: 0.622). Factor 3 is somewhat low to explain the cohesion between questionnaire items in this factor. However, Yamada & Akahori (2007) demonstrated the validity of factor 3 by showing cronbach alpha as 0.715. Therefore, this factor seems to be valid, and the questionnaire items cohesive, to some extent.

3.2 Difference Between the Three Factors in Subjective Data

A two-way analysis of variance (ANOVA) revealed that the main effect of the partner’s image was statistically significant in factor 1 (factor 1: $F(3, 36) = 4.598, p < .05$). Further, an interaction effect between image and use of voice versus text was confirmed for factors 1 and 2 (factor 1: $F(3, 36) = 7.185, p < .05$; factor 2: $F(3, 36) = 6.031, p < .05$). With respect to factor 3, only the main effect of the communication medium was confirmed ($F(3, 36) = 24.326, p < 0.001$). Table 3 shows the average score, main effects, and interaction effect for each item.
3.3 Difference Between Communication Media in Language Performance

The results of the ANOVA revealed a statistically significant effect on all items in the use of voice versus text. Further, the main effect of the partner’s image emerged in 2-1, “mean number of turns,” and 2-3, “mean number of self-corrections.” However, the interaction effect between the use of voice, text, and image also emerged in 2-1. Table 4 shows the results in detail.

3.4 Path Analysis

This study aimed to investigate the relationship between affective evaluation, image, and language learning performance for the effective design of CALL using SCMC. In order to determine the relationship, path analysis was conducted between the three factors’ score and three language learning performance metrics as dependent variables. Additionally, dummy variables were used to differentiate the medium used. The variable “partner’s image” was set to 1 when a learner’s own image was displayed for communication and to 0 when it was not. Another variable, “voice communication,” was set to 1 when voice communication was available and to 0 when it was not (text chat). The significant relationships are displayed in Fig 5.

Fig 5 shows the relationship between factors, medium, and overall performance. In this relation, partners’ images have a direct effect on the perceived ease of communication in English due to interlocutor’s image (factor 1), number of turns, and number of self-corrections as performance. Factor 1, “perceived ease of communication in English due to the interlocutor’s presence,” positively affects factor 2 “consciousness of natural communication” and the number of turns. Factor 2 has no influence on other factors or performance metrics. The communication medium (1: voice, 0: text chat) has a strong influence on most performance metrics Voice communication has a strong positive effect on all performance metrics; in particular, the number of turns and use of target expression are affected strongly by voice communication. However, learners using text chat tend to be extremely confident in grammatical accuracy; the effect of text chat on the number of self-corrections was confirmed by the path analysis.

3.5 Qualitative analysis

One of the significant differences between voice-based communication and text-based communication is the use of filler. Previous studies suggested that filler is a predictor of problems in voice communication (Smith & Clark, 1993; Clark, 1994). Filler was rarely
used in text-based communication, because learners can avoid communication problem by taking extra time to type accurate English words.

However, filler was often used in voice communication. It seems that the possible function of the filler used as a unique feature of voice-based SCMC. For example, the filler “ah” seemed to be used for constructing sentences, finding suitable words after consideration of opinion, or correcting errors in previous utterances (Smith & Clark, 1993). Some tendencies after fillers in SCMC were found in this study, which are similar to those of face-to-face communication. Examples 1, 2, and 3 display particular features of learning in dialogue through voice-based CMC.

Example 1 (Audioconferencing)
17. Subject 1 (S1): He is clever, but, the other candidate is more, so I shouldn't
18. Subject 2 (S2): uhnnn
19. S1: correct the teacher.
20. S2: (4 sec) Why do you think so?
21. S1: Ahhh he is smarter than uhnn . . . (3 sec) candidate 2.
22. S2: Okay . . . ahhh so but the good thing for him is . . . he . . . is that ahhh he . . .
23:   he . . . he has
(...)
44. S2: Probably we should ha . . . have . . . ahhh hire the candidate 4.
45. S1: Yes.
46. S2: And then ah ahh wu . . . wu . . . we need the . . . the . . . ahh negotiate annnd (and) he can
47.   work longer or we should look for another person during summer and after summer.

In example 1, the usage of each filler can be seen. The filler “ah” was used as a bridge between words; after “ah”, subject uttered the next word without intervening silence. Lines 22 and 44 display another function of “ah”: after the utterance of “ah”, subject 2 repeated and modified his words. For example, in line 44, subject 2 intended to convey the meaning “employ” but first uttered “have” which is an incorrect word choice. However, the subject then modified the word “have” to “hire” after the filler “ah”.

Example 2 (Videoconferencing)
32. Subject 4 (S4): But, it is . . .
33. Subject 3 (S3): Uhnnn . . .(putting his elbow on table)
34. S4: How do you think?
35. S3: But we don’t need experience in teaching so . . . ah I . . . I think . . . uhnnn . . .
36. number 2 is okay.

In example 2, filler was used with social cues such as putting one’s elbow on the table as a sign of disagreement in communication. In line 33, subject 3 seemed not to agree with his partner’s opinion, and uttered “uhnnn”. Subject 4 guessed that subject 3 had an opinion different from his own, and asked “How do you think?”. Subject 3 then expressed his opinion.

Example 3 (Videoconferencing)
26. Subject 6 (S6): Okay . . . right. But it **takes some cost** to teach him how to  
27. teach.(looking at the ground)  
28. Subject 5 (S5): **Uh? Uh?** (cocking head)  
29. S6: (2 seconds) **Ahhh** but, it costs us in teaching him how to teach, right?  
30. S5: **Uhhnn . . .**

Example 3 indicates social cues helping the learner to understand the interlocutor’s situation. In line 26, Subject 6 made a grammatical error. Subject 6 did not seem to have a confidence in grammatical accuracy, and spoke in a soft voice. Subject 5 asked subject 6 to repeat his utterance by saying “uh? uh?” because subject 5 could not hear subject 6. Subject 6 translated subject 5’s response and gesture as a sign of difficulty in comprehension. Therefore, subject 6 tried to modify his previous utterance. Filler seemed to be used along with social cues as an indirect trigger for feedback.

The effects of the presence of the partner’s image in communication were revealed not only in perceived presence, but also in perceived consciousness of second language communication and in productive performance. The presence of the partner’s image seemed to motivate learners to communicate, and served as a reinforcement medium for voice communication. Examples 4 and 5 show the effect of this presence in videoconferencing.

Example 4 (Videoconferencing)
35. Subject 7 (S7): Yes. Horiiike is the best one for this work.  
36. Subject 8 (S8): Hai (“yes” in Japanese) (with laughing)  
37. S7: **His life and ability is very good.** And the…  
38. S8: (interrupting) Yeah (nodding)  
39. S7: And he can go to work **not take** long time.  
40. S8: Uhnmm (with nodding) (41 seconds), is our choice Miss Horiiike (cocking inclining head)? (S8 mistook of the sex of the character “Mr. Horiiike” in the task)  
41. S7: Yes.  
42. S8: Yeah. (with laughing)
This dialogue indicates that non-verbal behaviors seemed to help subjects to speak English without frustration due to grammatical and lexical errors. In this communication, we can find several errors in grammar (e.g., use of the singular “is” in line 37). However, behaviors such as nodding and laughing allowed them to relax and speak positively. Moreover, such behaviors facilitate the transfer of intention to their interlocutors. In line 40, we can see the concrete functions of non-verbal behavior. In order to confirm their decision, subject 8 asked uncertainly, “is our choice Miss Horiike?”, while giving a sign of negative intention. Subject 7 responds to this question in line 41. In lines 42 and 43, they relaxed and began laughing, having confirmed their agreement. As can be seen from this example, the partner’s image seemed to play an important role in facilitating communication; however, in text-chat with image, such behaviors were not confirmed. Immediacy, which is one of the features of social presence, seems to facilitate communication in both videoconferencing and text chat, but non-verbal behavior has a strong power of not only immediacy but also negative feedback which may lead to effective learning (Lou et al, 2003).

Example 5 (Videoconferencing)
23. Subject 9 (S9): So I . . . I think she has not ability to teach to them . . . so . . . I think number 3 is not good. So . . . I . . .
24. Subject 10 (S10): (interrupting) Yeah I think so.
25. S9: Uhn . . . let's think about 1, 2, 4.
26. S10: (8 seconds) Uhn . . . (inclining head)
27. S9: (interrupting) Ahh . . . ahh . . . which person do you have any opinion whether
28 she or he has no ability?

Example 5 shows a situation in which the utterance of filler as a learner works to express their desired meaning can promote the use of formulaic speech, depending on the learner’s proficiency. In lines 26 and 27, subject 9 looked at subject 10’s image and suspected that subject 10 had some problem or idea; subject 9 then interrupted with filler and asked a concrete question. However, subject 9 seemed to have difficulty uttering her question, said filler “Ahhh”. Therefore, she clicked the target expression buttons and uttered the expression “which person do you have…”. It is a notable possibility that formulaic speech may be utilized during the construction phase of the desired meaning in
communication.

On the other hand, text chat does not allow learners to use non-verbal behavior. Examples above reveal that non-verbal behavior has a central role in communicating and may lead to language learning, but text chat has its own unique communication features, and the limitations imposed on communication also provide a possibility to promote successful learning. Examples 6 and 7 show the notable features of text chat in communicative language learning.

*Example 6 (Plain text chat)*

9. Subject 11 (S11): (23 seconds) I think that we need to think about the ability of communication.
10. 
11. Subject 12 (S12): (18 seconds) How do you think?
12. S11: (40 seconds) Yes, that is the most important.
13. S11: (26 seconds) I think the last one is good about the ability of communication.
15. S12: (15 seconds) Why do you think so?
16. S11: (45 seconds) *I'm sorry. I am so slow to type words.* She can speak the language of that country.
17. 
19. S12: (18 seconds) Yes, I agree with you.

In example 6, learners took considerable time to type words in order to communicate in English. Compared with voice communication, it’s only natural that communicating with the interlocutor would take time, but learners felt sorry that they were slow to type English words. The text chat used in this experiment cannot show learners’ actions, such as typing. Learners sometimes had to wait for their interlocutor’s response, making learners worry about miscommunication. However, there is an outstanding feature of learning in this medium: Overall, learners in this medium rarely made grammatical mistakes. In this example, subjects 11 and 12 had only a few grammatical errors, though there are some unclear sentences (e.g. in lines 13 and 14). This suggests that learners have enough time to construct the desired meaning and accurate sentences before typing words.

*Example 7 (plain text chat)*
31. Subject 13 (S13): (17 seconds) Why?
32. S13: (37 seconds) She have the knowledge, but she have no experience.
33. Subject 14 (S14): (91 seconds) She have been study the skill, but her has not the experiences
34. S14: (12 seconds) I'm sorry. "She has"
35. S13: (14 seconds) All right. "She has"
36. S13: (73 seconds) But I think the communication skill is the most important thing
37. for this job.
38. S14: (33 seconds) Then I think the best election is the member 2.
39. S14: (7 seconds) Why?

This example displays the feature of modification in communicative language learning using text chat. Both subjects 13 and 14 had grammatical errors about the third person singular, writing “She have” in lines 32 and 33 (subject 14 also made a nominative error in line 33, “her has…”). However, the reflective feature of text chat seemed to raise the learners’ consciousness of grammatical accuracy, due to the display of their utterances. In fact, in line 34, subject 14 apologized about the grammatical error and modified her error. This unique feature of text chat supports Lee’s research (Lee, 2002).

3.6 Opinions and Suggestions From Participants

Some participants who used each type of SCMC commented on the SCMC implementation and other functions. Almost all those who used SCMC with their partner’s image reported a positive effect on communication. On the other hand, participants who used SCMC without their partner’s image tended to have a negative opinion.

Positive comments

Comment 1: I could communicate with my partner in videoconferencing software as in a face-to-face setting. It made me want to speak during the communication. (videoconferencing)

Comment 2: I could perceive my partner’s presence, so I could enjoy communicating in English. But a larger image size would have been given a greater impression of my partner’s presence. (videoconferencing)

Comment 3: I could consider what to say thanks to seeing my partner’s nonverbal response. It helped me to learn English. (videoconferencing)

Comment 4: When I made a grammatical error, my partner could understand my
desired meaning. So I felt relaxed, even when I made a mistake. *(text chat with image and plain text chat)*

Comment 5: I could communicate with my partner. I watched her image to see her behavior, so I could wait for her response. *(videoconferencing)*

Negative comments

Comment 1: I felt that it was difficult to speak with only voice. When I could not imagine suitable words, I wanted to transfer my desired meaning with gesture. *(audioconferencing)*

Comment 2: I could not find an opportunity to start communication, because I could not see my partner’s nonverbal response, which would have enabled us to understand the timing to start communication. *(audioconferencing)*

Comment 3: My partner had to wait for my response many times, because it took me a long time to find the right words. I was worried that he thought I wasn’t serious. *(plain text chat)*

Comment 4: I could not voice my thoughts and feelings without seeing my partner’s image. *(audioconferencing)*

With respect to the use of voice or text in English communication, most opinions were negative for both media. Comments about voice communication generally reflected self-evaluation of the learner's ability. On the other hand, comments about text-chat tended to focus on the sense of discomfort.

Comment 1: Text-chat made me consider what to say rapidly, because I felt compelled to respond rapidly. It caused me to use inaccurate English.

Comment 2: I could not find a suitable expression during English communication, and I felt bad for making my partner wait for my response.

Comment 3: I could not use gestures with voice explanation in text chat with images, in order to transfer my desired meaning. Gestures alone made it difficult for me to understand what my partner meant.

Comment 4: I felt uncomfortable when the text was mixed with the image. I prefer to speak. Voice communication enables us to reply rapidly.

Comment 5: I had to change my message depending on my partner’s response while I was typing.

4. Discussion

In this experiment, I conducted comparative research between four types of SCMC in
order to find the relationship between medium, perceived evaluation, and learning performance. I found two types of consciousness: consciousness of the proximity to face-to-face communication, which is concerned with factors 1 and 2, and consciousness of language learning, as reflected in factor 3. The results of this experiment reveal that the combination of interlocutor’s image and voice affects the perceived consciousness of natural communication in SCMC use. On the other hand, the communication tool used (voice versus text) influences the consciousness of learning such as grammatical accuracy. Moreover, the results of the ANOVA reveal an interaction effect between the image and communication tool for factor 1 ("perceived ease of communication in English due to interlocutor's image"). This suggests that the use of image and voice gives learners an enhanced perception of presence; that is, a perception of similarity to a face-to-face setting as displayed in the dialogue data, in which filler and social cues assist in solving communication problems.

Text chat can enhance learners’ consciousness of grammatical accuracy as compared to videoconferencing and audioconferencing. This is because displaying words enables learners to focus on grammatical form and encourages reflection on the accurate spelling of vocabulary. Learners can check the accuracy of their messages as they type; when they find an error in typing, they can revise it before sending the message to their partner. This reflective feature of text chat was pointed out by Lee (2002). In path analysis in this study, text chat led to high confidence of grammatical accuracy in English communication. Moreover, learners can use more time to type, reflect and send their message to their partner than they do during communication using videoconferencing and audioconferencing. It can be said that text chat has a positive influence on grammatical accuracy, allowing learners to type their message while feeling relaxed. This suggests that the Japanese learners can be more relaxed and satisfied with English communication in text chat, considering foreign language anxiety and learning belief as features of Japanese English learners (Kubo, 1997). However, these results also show that learners tended to enjoy natural communication rather than focusing on learning consciousness in SCMC. Proximity to face-to-face communication enhances affective support in communication-based learning, because it can be easy to transfer the desired meaning through the use of social cues, particularly emotion transfer with social cues such as smiling (e.g., McIssac & Gunawardena, 1996; Gunawardena & Zittle, 1997). The results of this research also suggest that media close to face-to-face (partner’s image display and voice-based communication) promotes natural communication similarly to an actual face-to-face setting. The ease of the desired meaning transfer and immediacy based on
social presence seems to be an importance factor on learner’s affective side in order to continue communicative language learning, while text chat has effective feature for communicative language learning on affective side such as reduction of foreign language anxiety.

Learning output was affected by both the presence of the partner's image and the use of voice. This result shows that the interaction effect of image and voice promotes an increased number of utterances and self-corrections in English communication: videoconferencing leads to more frequent utterances with grammatical reflection, considering the results displayed in table 5 and dialogue analysis, while the use of voice versus text influences other aspects of output. The results of this experiment indicate that the choice of communication tool influences English communication. This is related to learners' ability: comments from learners suggested that inadequate English proficiency prevented learners from continuing English communication. The target expression and self-correction features complement their lack of English proficiency. However, it seems that communication tools may have an effect on conscious raising of natural communication, judging from the results of ANOVA. Learners' comments indicated that the added presence of the interlocutor's image enhanced this aspect, while the partner’s image alone has a significantly strong effect on factor 1.

Moreover, the presence or absence of the partner's image affected learners’ self-correction, judging from the results of ANOVA and path analysis. Learners mentioned that they considered using gestures to determine their partner's level of comprehension and could modify their utterances on recognizing inadequate proficiency in transferring comprehensive meaning to their partner. Learners reveal all their thoughts explicitly, including the affective aspects, because plain chat does not allow learners to use social cues (Garrison et al, 2003; Levy et al). In addition to this aspect, the immediacy of SCMC seems to compel learners to produce language quickly (Levy, et al, 2006). Taking this point into consideration, learners seem to feel responsibility toward their partners while communicating in English with SCMC, and their partner’s image helps them judge when to speak and determine their partner’s level of comprehension. The case in which a subject using text chat apologized for their slow response because of slow typing seems to support this view.

In the path analysis, significant relationships between some psychological perceptions and interactive performance were confirmed. The aspect of presence caused more active communication due to the presence of the partner’s image, thus enhancing the perceived ease of communication in English. Learners tended to speak English actively even in
learner-centered communication, finding it easy to speak English due to the partner’s image (factor 1). The presence of the partner’s image has an indirect effect on factor 2, “consciousness of natural communication.” The presence of the partner’s image helps learners to understand their partner’s situation such as in the case of miscommunication. Sometimes learners are unable to comprehend each other; they notice this miscommunication and try to modify their former utterances.

Voice communication has a clear influence on all aspects of learning performance. It enables learners to speak naturally, as in face-to-face communication. Learners can use filler, which indicates communication problems or is a sign of sentence construction, in voice-based CMC much like they would in face-to-face communication. Moreover, judging from the dialogue data, learners can interrupt their partner’s utterances due to the ease of comprehension of their partners’ meaning with the help of social cues, thus promoting active English communication.

Voice communication also promotes self-correction. Nonverbal devices such as facial expressions enable learners to understand their partner’s comprehension. When a learner’s partner shakes their head, it is easy for the learner to understand that their partner is unable to comprehend what is being said. While using voice communication, in which learners’ social cues are absent, they attempt to speak English accurately. Therefore, they try to modify their previous expression or use another expression in order to transfer the desired meaning. Judging from the participants’ comments, learners reflect on their utterances, based on their partner’s level of understanding, and then modify the grammatical errors in the utterances.

As the results of this research reveal, interactive media such as videoconferencing which promotes social presence promote active communicative language learning from the viewpoint of the affective side. Moreover, the signs concerned with social presence such as social cues seem to be effective on reflective learning. In fact, learners tried to modify the grammatical errors in their utterances based on the facial expressions shown by their partners. Nonverbal devices play an important role in compensating for the lack of oral proficiency; however, it is difficult for learners to transfer their desired meaning only through verbal expressions. The same logic can be applied to the background of the effect of voice communication on the use of target expressions. This point suggests social presence leads to learning consciousness, enhancing learning performance as a result. According to previous studies (e.g., Gunawardena, 1995; Garrison & Anderson, 2003; Sato & Akahori, 2005), social presence affects the will to learn. This research demonstrates that social presence effectively promotes interaction in communicative
language learning, raising the consciousness of learning and leading to increased learning performance such as the frequency of utterances and grammatical modification, as suggested by previous SLA research (e.g., Long, 1991; Swain, 1995).

In this study, target expressions are concerned with communication strategies or theme-specific expressions, helping learners to continue English communication by solving communication problems. Based on the qualitative data shown in 3.5, the use of target expressions assists learners in transferring their desired meaning without social cues when they recognize their lack of English proficiency. This suggests that the similarity to face-to-face communication, that is, the use of social cues and filler can promote the use of target expressions, and may lead to successful communicative language learning. This point needs to be investigated in detail by subsequent research.

On the other hand, the negative relationship between voice communication and grammatical accuracy is significant; as confirmed above; that is, the lack of social cues encourages learners to be conscious of the accurate transfer of meaning during communication in text chat. Learners can use nonverbal devices in videoconferencing sessions in order to transfer their desired meaning (Garrison et al., 2003). The above results coincide with the results of the ANOVA in this study. The only relationship found for target expression usage was that with voice communication; psychological perception had no significant effect on expression usage. The use of voice makes learners feel burdened while communicating in English. In this situation, learners have the opportunity to use target expressions related to communication strategies and theme-specific expressions.

5. Conclusion

This study aimed to examine three points about the use of SCMC in second language learning: the effect of the communication medium on learners' affective side, the effect of the medium on output, and the relationship among the communication medium, learners’ affective side, and productive performance. Prior research about the effects of social presence on learning has focused on affective side during learning; this is why, in e-learning using text-based communication tools such as BBS, one of the common problems is how to satisfy online learners through the promotion of continued learning (Bersin, 2004). The viewpoint of social presence research in the area of e-learning has been mainly focused on the communication process. Previous research (e.g., Tu, 2002; Garrison & Anderson, 2003) suggested evaluation criteria for the enhancement of social presence in e-learning, which are increasingly being seen as important criteria for
successful e-learning.

It is not clear how social presence affects learning performance directly and indirectly. In language learning, SCMC has been used frequently throughout the world, but it has been suggested that the enhancement of social presence leads learners to speak in their native language (Yamada & Akahori, 2007). Setting aside the issue of whether the use of native language facilitates language learning, it seems that the use of SCMC does not always increase output in the target language, which plays an important role in second language acquisition (Swain, 1985; Swain, 1995).

An interlocutor’s image and voice communication enables learners to understand each other, since learners can use nonverbal cues such as smiling and nodding. Nonverbal cues promote comprehensive communication between learners in a second language. Learners can transfer their desired meaning and judge whether their partner understands what they want to say. Image and voice raises learners' consciousness of natural communication as it brings about a situation that is similar to a face-to-face one. In particular, social cues and filler seem to play an important role in both enhancing social presence and promoting effective communicative language learning, based on the results of qualitative and quantitative data. Communication via text chat, on the other hand, raises learners’ grammatical consciousness. The reflective feature of text chat allows learners to be conscious of grammatical accuracy and modify errors without feeling rushed. Finding errors and revising the message before sending it allow learners to communicate with each other using accurate grammar.

The results of path analysis suggest that affective perception such as the relief in and ease with regard to communication, along with the use of audio and image, directly and indirectly enhances aspects of learning performance such as number of turns and self-corrections. Learners frequently use target expressions as learning objectives in voice communication because target expressions concerning communication strategies assist learners in continuing communication during voice SCMC, during which learners might feel burdened. On the other hand, the consciousness of grammatical accuracy and the use of voice communication have a negative relationship. This point supports the results of the ANOVA.

Considering the implementation of each medium from an educational standpoint, it can be said that text chat encourages learners to be conscious of grammatical and lexical accuracy. In particular, novice learners tend to be conscious of continuous communication rather than the form of communication (Van Patten, 1990); on the other hand, Japanese English learners tend to prefer studying form (Kubo, 1997). From these
viewpoints, text chat is effective for studying both form and for communication skills. However, I cannot deny the influence of multimedia-based SCMC such as videoconferencing. Voice communication provides many opportunities to study by negotiating meaning (Japson, 2002). Voice communication with images offers the opportunity to foster practical skills in communication such as laughing and nodding (Kock, 2001), and it can enhance distance learning, thus increasing learners’ motivation and leading to continued participation in learning (Smyth, 2005). In this study, videoconferencing had a significant effect on all aspects of performance and affective evaluation.

However, it is essential to understand the features of each medium with respect to SLA and effectively use them, for example, blending text chat and videoconferencing, considering learners’ language proficiency, learning objectives, target language, computer literacy, and computer environment. For example, Toyoda et al. (2002) suggested the effectiveness of the use of CMC in interactive language learning, similarly to our research though using Japanese as the second language. However, unique language features such as filler, which is an important measure of communicative competence for successful communication in Japanese, should be considered in order to set learning objectives.

From the viewpoint of intercultural communication as a learning objective, highly interactive communication and the narration of personal experiences are essential factors (Kelm, 1992; O’Dowd, 2005). An interlocutor’s presence, which is the key factor for highly interactive communication, plays an important role in self-disclosure, which is the desire to narrate one’s personal experiences and ideas. The anonymity of forms of CMC such as text chat promotes self-disclosure (Joinson, 2001). Thus, SCMC in which the interlocutor/partner is not present, such as plain text chat and audio conferencing, are effective in highly interactive intercultural communication.

This study proposes the effects of SCMC in communicative second language learning, considering the relationship between consciousness during English communication and productive performance. In particular, the relationship between them can suggest a direction in learning design for developers and teachers. However, this study has certain limitations; for example, this study was designed within an experimental setting, and the results may not be directly applicable to practical environments. Long-term investigation in classroom settings with a larger number of participants is required to further examine the effectiveness of SCMC and the applicability of this model.
<table>
<thead>
<tr>
<th>#</th>
<th>Questions</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rate the frequency of your utterances while using SCMC</td>
<td>1: less than face-to-face–6: more than face-to-face</td>
</tr>
<tr>
<td>2</td>
<td>Rate your perceived ease of initiating communication</td>
<td>1: very low–6: very high</td>
</tr>
<tr>
<td>3</td>
<td>Rate your perceived ease of understanding your partner’s utterances</td>
<td>1: very low–6: very high</td>
</tr>
<tr>
<td>4</td>
<td>Rate your perceived ease of saying what you wanted to say</td>
<td>1: very low–6: very high</td>
</tr>
<tr>
<td>5</td>
<td>Rate your perceived consciousness of your partner’s presence</td>
<td>1: very low–6: very high</td>
</tr>
<tr>
<td>6</td>
<td>Rate the perceived relief in communication while using SCMC</td>
<td>1: very low–6: very high</td>
</tr>
<tr>
<td>7</td>
<td>Rate your perceived ease of communication in English</td>
<td>1: very low–6: very high</td>
</tr>
<tr>
<td>8</td>
<td>Rate your perceived feeling of the similarity between face-to-face and SCMC</td>
<td>1: very different–6: very similar</td>
</tr>
<tr>
<td>9</td>
<td>Rate your perceived consciousness of accuracy in English communication</td>
<td>1: very low–6: very high</td>
</tr>
<tr>
<td>10</td>
<td>Rate your perceived consciousness of responding as soon as possible</td>
<td>1: very low–6: very high</td>
</tr>
<tr>
<td>11</td>
<td>Rate your perceived consciousness of communicating the desired meaning in English, even if you made a grammatical mistake</td>
<td>1: very low–6: very high</td>
</tr>
<tr>
<td>12</td>
<td>Rate your perceived ease of reflecting on your utterances</td>
<td>1: very low–6: very high</td>
</tr>
<tr>
<td>13</td>
<td>Rate your perceived consciousness of the grammatical accuracy of your partner’s utterances</td>
<td>1: very low–6: very high</td>
</tr>
<tr>
<td>14</td>
<td>Rate your perceived consciousness of the response speed of your partner</td>
<td>1: very low–6: very high</td>
</tr>
<tr>
<td>15</td>
<td>Rate your perceived consciousness of the comprehension of your partner’s utterances</td>
<td>1: very low–6: very high</td>
</tr>
</tbody>
</table>
Table 2
Data collected through the analysis of video records

<table>
<thead>
<tr>
<th>#</th>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1</td>
<td>Total number of turns</td>
</tr>
<tr>
<td>2-2</td>
<td>Utterances of the target expression</td>
</tr>
<tr>
<td>2-3</td>
<td>Self-corrections</td>
</tr>
</tbody>
</table>

Table 3
Mean scores and effects for each factor

<table>
<thead>
<tr>
<th>#</th>
<th>Factor</th>
<th>Image</th>
<th>Voice/Text</th>
<th>Mean</th>
<th>Main effect of image</th>
<th>Main effect of voice/text</th>
<th>Interaction effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Perceived ease of communication in English due to the interlocutor’s presence</td>
<td>Yes</td>
<td>Voice</td>
<td>21.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>Voice</td>
<td>12.90</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yes</td>
<td>Text</td>
<td>16.80</td>
<td>*</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>Text</td>
<td>17.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fac1</td>
<td></td>
<td>Yes</td>
<td>Voice</td>
<td>20.10</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>Voice</td>
<td>15.90</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yes</td>
<td>Text</td>
<td>17.10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>Text</td>
<td>20.10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fac2</td>
<td>Consciousness of natural communication</td>
<td>Yes</td>
<td>Voice</td>
<td>7.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>Voice</td>
<td>8.10</td>
<td></td>
<td></td>
<td>***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yes</td>
<td>Text</td>
<td>10.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>Text</td>
<td>11.70</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

***: p < .001; *: p < .05
Table 4
Mean values and effects for language performance

<table>
<thead>
<tr>
<th>#</th>
<th>Metric</th>
<th>Image</th>
<th>Voice/Text</th>
<th>Mean</th>
<th>Main effect of image</th>
<th>Main effect of voice/text</th>
<th>Interaction effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1</td>
<td>Mean number of turns</td>
<td>Yes</td>
<td>Voice</td>
<td>65.30</td>
<td>***</td>
<td>***</td>
<td>**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>Voice</td>
<td>34.40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yes</td>
<td>Text</td>
<td>10.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>Text</td>
<td>9.10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-2</td>
<td>Mean number of utterances of the target expressions</td>
<td>Yes</td>
<td>Voice</td>
<td>1.30</td>
<td>**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>Voice</td>
<td>1.70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yes</td>
<td>Text</td>
<td>0.60</td>
<td></td>
<td></td>
<td>**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>Text</td>
<td>0.40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-3</td>
<td>Mean number of self-corrections</td>
<td>Yes</td>
<td>Voice</td>
<td>1.50</td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>Voice</td>
<td>0.50</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yes</td>
<td>Text</td>
<td>0.40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>No</td>
<td>Text</td>
<td>0.20</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

***: p < .001; **: p < .01; *: p < .05; +: p < .1
Fig. 1
Interface of video conferencing

Fig. 2
Interface of audio conferencing

Fig. 3
Interface of text chat with images

Fig. 4
Interface of plain text chat
Fig. 5 The relationship between media, psychological factors and performance metrics (dotted line means negative effect)

REFERENCES
Beauvois, M.H. (1998a) Write to speak: the effects of electronic communication on the oral achievement of fourth semester French students, In Muyskens, J.A (Eds), New ways of learning and teaching: focus on technology and foreign language education, 93-115, Heinle & Heinle, Boston, MA, USA.
Bersin, J. (2004), The Blended Learning Book - Best Practives, Proven Methodologies, and Lessons Learned, Pfeiffer, CA, USA
Educational Testing Service (2008), *ETS TOEIC Data and analysis 2007*


Matsuda, T., and Harada, M (2007), Mentoring in e-learning: for the practice of the learning support, Tokyo Denki University Press, Tokyo, JAPAN


O’Dowd, R. (2005). Negotiating sociocultural and institutional contexts: The case of


