What comes with technological convenience? Exploring the behaviors and performances of learning with computer-mediated dictionaries

Tzu-Chien Liu a,*, Po-Han Lin b,1

a Graduate Institute of Learning and Instruction and Center for Teacher Education, National Central University, Jhongli, Taiwan
b Graduate Institute of Learning and Instruction, National Central University, Jhongli, Taiwan

1. Introduction

Reading authentic texts has long been recognized as an essential part of accumulating knowledge and developing language ability (Gilmore, 2007; Krashen, 1989), and the dictionary is the most common type of reading aid. If an individual reads text and encounters an unfamiliar word, it is likely that he or she will refer to a dictionary if it is necessary to understand the word and its definition cannot be inferred from the text. While the primary goal is for the reader to understand the meaning of the text, many words can be unintentionally "pick(ed) up" (Hulstijn, Hollander, & Greidanus, 1996; Krashen, 1982, 1989; Laufer & Hulstijn, 2001). This "pick up" is usually referred to as incidental vocabulary learning and is considered a popular theory of language acquisition as a result of comprehensible input (Krashen, 1989). Therefore, the use of a dictionary while reading has the potential to assist learners not only in reading comprehension but also in incidental vocabulary learning.

As web-based technology rapidly develops, the availability of authentic materials on the Internet has expanded exponentially. This, in turn, has made online reading an increasingly popular way of learning (Abraham, 2008). Computer-mediated aids have become particularly important, as they are now often built in or are accessible using digitalized-article reading devices. Therefore, the development of a computer-mediated dictionary has received considerable interest due to its promising potential and increasingly prevalent and broad applications (Kramsch & Anderson, 1999).

Type-in and pop-up dictionaries are two major computer-mediated aids used today. They enable word definition consultation at an individual's fingertips. Type-in dictionary aids are generally in the form of computer software or online dictionary websites, which allow the desired word to be typed into an entry window while the original material is still visible on the screen. Pop-up dictionaries allow readers to double-click on any given word in a text to bring up a definition adjacent to or above the chosen word. Although the two computer-mediated dictionary forms serve the same purpose as a conventional book dictionary, the method of searching for information is quite different.

Without a doubt, in comparison with conventional paper-based dictionaries, type-in and pop-up forms of computer-mediated dictionaries provide learners with a more convenient way to search for words. This convenience may be thought of in terms of the reduction in the search effort for the target words. However, "convenience", as a consequence, could potentially eliminate any positive
learning process altogether. For example, when using a conventional book dictionary, the user must temporarily utilize his or her working memory (WM) by mentally holding the word while actively searching through the pages. This continuous rehearsal could potentially be helpful for incidental vocabulary learning. It is generally agreed upon that before a memory can become durable and recallable, additional processes (e.g., deep processing, elaboration, or cognitive effort) are usually involved (Anderson, 1995; Baddeley, 1998; Craik & Lockhart, 1972; Hulstijn, 2001). However, according to cognitive load theory (CLT), such occupation of the WM could leave fewer resources available for meta-cognitive processes, which could potentially impede reading comprehension (Sweller, van Merriënboer, & Paas, 1998). Therefore, it is possible that the convenience afforded by computer-mediated aids could potentially free up more cognitive resources for comprehension. The other apparent advantage of the advanced aids is the learning opportunities associated with their use. Due to the convenience, readers might be more willing to use computer-mediated dictionaries, thereby exposing themselves to more words.

However, little is known about the consequences of technological convenience. The exact nature of this information immediacy and how it affects the subsequent usage behavior, cognitive processes and learning performance remain unknown, despite the fact that these are important and interesting research issues. Although some studies on e-dictionaries have attempted to compare the effects of dictionary type on subsequent learning (e.g., Knight, 1994; Leffa, 1992), the lack of experimental control and data collection on usage behaviors limits the information necessary to answer the questions that we are interested in investigating. By avoiding the problems mentioned above, our study aims at understanding how technological convenience changes the long-standing pedagogical practice of dictionary usage and learning.

In this study, we compared type-in and pop-up dictionaries to conventional paper-based dictionaries and a situation in which no aid was available. The research presented herein was guided and motivated by three specific questions: (1) What cognitive processes are associated with different types of aids? (2) Do differences in accessing various types of dictionaries lead to differential reading comprehension and incidental vocabulary learning? (3) What are the relationships between cognitive processes and learning performance?

2. Theoretical background

A review of previous research and cognitive load theory (CLT) is presented below. This section summarizes foundational research and integrates CLT with the predictions of this study.

2.1. Effects of computer-mediated aids on reading and vocabulary learning

Although use of computer-mediated aids (e.g., dictionaries, glossaries, annotations, etc.) has the rational appeal of enhancing reading comprehension and incidental vocabulary learning, results regarding their effectiveness are mixed. Some researchers have examined the availability of annotations on reading and found improvements in learning (e.g., Chun & Plass, 1996; Jones & Plass, 2002), whereas others did not observe improvements (e.g., Brünken, Plass, & Leutner, 2004; Plass, Chun, Mayer, & Leutner, 2003). For instance, Leffa’s (1992) examined the effectiveness of electronic dictionaries compared with conventional dictionaries and found better comprehension performance and faster translation writing speed using the electronic dictionary. Knight (1994) also found the occurrence of incidental vocabulary learning and improved comprehension when participants used an electronic dictionary during a reading task (compared with not using an aid). However, other studies on computer-mediated aids and conventional book dictionaries failed to find any differential effect on comprehension or vocabulary learning. For example, Bensoussan, Sim, and Weiss (1984) questioned the benefit of dictionary usage on reading comprehension when comparing advanced students’ English reading performance with and without aids. Aust, Kelley, and Roby (1993) showed that participants who used an electronic dictionary looked up twice as many words as those who used a conventional dictionary, although no differences were found in comprehension.

One potential explanation for the variations in findings could be methodological limitations. First, many of the studies noted above did not control for the amount of information presented in the dictionaries. In other words, whether the differences observed in performance were due to differences in the amount of information presented or were attributable to differences in learning processes cannot be distinguished. For example, in Leffa’s (1992) experiment, the types of conventional dictionaries (e.g., desk or pocket dictionaries) students bought and used in the experiment were not restricted. The amount of information (e.g., definitions and examples) presented could influence the search effort put forth to find the definition fitting the context, an elaborative process that others argue is related to reading comprehension (Hulstijn, 2001). Thus, to investigate the effects of learning processes on performance, we controlled for the information contained in various aids and attempted to distinguish the effects mediated by differences in learning processes rather than differences in content.

Second, a detail and thorough log file on dictionary usage behaviors is usually absent in previous studies. Few studies did provide additional information on the frequency of usage and overall reading time (Aust et al., 1993; Chun, 2001), but other important information is still missing (e.g., which words the participants looked up, vocabulary search time, definition reading time, and text reading time). These are some of the important insights that should be taken into account when considering the usage behaviors involved in using a dictionary. We believe that the learning process of using dictionary may be subdivided to several behavioral processes that we categorized. More rigorous research methodology on tracking learners’ “look-up” behaviors could provide clues to the cognitive processes involved in using various types of dictionary and how these processes may contribute to subsequent learning.

Moreover, both comprehension and vocabulary-learning performance should be considered when examining the effects of aids. A recent meta-analysis conducted by Abraham (2008) included 11 reports on computer-mediated glossaries in second language reading and found an overall medium effect size for comprehension and a large effect size for incidental vocabulary learning, suggesting variations in comprehension and vocabulary performance. We believe that an examination of the effect of aids should not be based merely on performance in one area. Complex relationships may exist between reading comprehension and vocabulary learning with different types of aids and, therefore, measuring both types of performance could provide an overarching picture of the effects of dictionary aids.

By controlling dictionary content across conditions and having a rigorous collection of detailed information on cognitive processes, comprehension, and vocabulary learning, this study can ultimately help to explain differences in reported by previous studies. In addition, CLT was incorporated in this study to provide guidance and a framework for our hypotheses and predictions.

2.2. Cognitive load theory

CLT suggests that learning performance is based on the interaction between the task, learners’ prior knowledge, and learners’
cognitive architecture constraints, namely the WM limitation (Paas, Renkl, & Sweller, 2004; Sweller et al., 1998). CLT has proved useful in terms of providing instructional design principles, guidance for multimedia design, and educational training (Mayer & Moreno, 1999, 2003; Sweller et al., 1998). According to CLT, learning performance is influenced by three kinds of cognitive loads: extraneous, intrinsic, and germane (Paas et al., 2004; Sweller et al., 1998). Intrinsic cognitive load refers to the basic memory requirement for holding the task elements in WM for simultaneous processes. Extrinsic cognitive load occurs when irrelevant information occupies and reduces WM resource available for learning. Finally, germane cognitive loads are efforts exerted by the learner that ultimately facilitate learning performance rather than impair it.

According to CLT, intrinsic, extraneous, and germane cognitive loads are additive (Sweller, 2010; Sweller et al., 1998). A decrease in extraneous cognitive load releases more WM resource available for handling intrinsic and germane cognitive loads (Paas, van Gog, & Sweller, 2010; Sweller et al., 1998). Based on CLT, Yeung, Jin, and Sweller (1997) conducted a series of experiments manipulating the physical distance between the explanatory notes and reading passages to examine the effect of presentation formats on reading comprehension and vocabulary learning for individuals with varied English proficiency levels. They argue that integrating vocabulary definitions within a passage (integrated format) helps to enhance comprehension because of the reduction in extraneous cognitive load, whereas separate vocabulary lists (separated format) are better suited for vocabulary learning due to reduced distraction from the reading material. This is exactly what they found in less proficient readers. For these participants (e.g., fifth or eighth graders with less proficiency), integrated format enhanced comprehension but did not improve vocabulary learning when compared to the separated format. The effect was reversed, such that the integrated format enhanced vocabulary learning but reduced comprehension, when participants were adults or eighth graders with higher proficiency. They argue that the same presentation format may either facilitate or interfere with performance, depending on the various cognitive loads imposed by the aid, as well as individual differences (Yeung et al., 1997).

Similarly, recent CLT researches focus on the goal of optimizing the learning environment by reducing the extraneous cognitive load to increase germane cognitive load. For instance, recent research on computer-assisted reading has investigated the effectiveness of types of supplementary information provided (e.g., annotations and multimedia) and the influence of presentation format (e.g., multiple representations or modality) that is beneficial to comprehension or vocabulary learning (e.g., Acha, 2009; Chun & Plass, 1996; Kablan & Erden, 2008; Mayer, Hegarty, Mayer, & Campbell, 2005; Peters, 2007; Yoshii, 2006). For example, Yoshii (2006) examined the effectiveness of different types of glosses (addition of pictorial cues in first language and second language glosses) on incidental vocabulary learning and found a beneficial effect of gloss, regardless of type. Acha (2009) also examined the effect of different annotation formats (visual, verbal, or both) on vocabulary learning in elementary students and found that children who received verbal annotations (written translation) learned better on both immediate post-test and delay post-test than children who received visual annotations (picture representing the words) or both.

This study also applies CLT and examines the effects of three different dictionaries (i.e., a conventional book dictionary, a type-in dictionary, and a pop-up dictionary) that varied in the effort required to access the relevant information. In this study, we are interested in determining whether information immediacy could potentially reduce the extraneous cognitive load by comparing pop-up and type-in dictionaries for which presentation formats are the same but the means of accessing the information differs. In addition, we explore how the reduction in search effort associated with these two types of dictionaries influences learning, and how this compares to a book dictionary. We believe, not only that the physical distance between the dictionary and the text matter (Yeung et al., 1997), but also that the effort associated with the dictionary must be considered when evaluating the effectiveness of these tools. Our study first explores the types of processes associated with different dictionaries; we then attempt to determine the nature of these processes by examining their influence on performance. We are also interested in whether the characteristics of these tools influence their usage frequency and how such differences may be related to subsequent learning.

2.3. Hypotheses

Regarding the cognitive processes associated with various types of dictionary aids used while reading, dictionary usage and reading behaviors are categorized into four different process indicators: (1) frequency of usage; (2) vocabulary searching time; (3) vocabulary reading time; and (4) text reading time. Although each indicator is used to represent the possible cognitive process, the nature of each cognitive process is still unknown and is therefore the main interest of this study.

Due to the physical differences in dictionary forms, searching through book pages to find a word would naturally require more effort than simply double-clicking a mouse to bring up a pop-up definition window or typing in the unknown word. Therefore, we state in Hypothesis 1 that readers will exert the most effort finding the target word in a book dictionary, followed by the type-in dictionary and then the pop-up aid (H1). Vocabulary search time is used to indicate the search effort.

Another indication thought to be related to the search process is the effort required to read the text. According to CLT, when more effort is associated with vocabulary searching and learning, fewer resources can be devoted to text comprehension. Because technology helps to reduce the effort required for vocabulary searching, more resource could potentially be devoted to reading comprehension. Thus, we postulate in Hypothesis 2 that readers in the pop-up condition will exert the most effort reading the text, followed by the type-in and then the book condition (H2). Text reading time is used to indicate text reading effort. Similarly, Hypothesis 3 states that willingness will differ by dictionary type due to differences in the search effort associated with each aid. Participants will be more willing to use the pop-up aid than other tools and will be least willing to use the conventional dictionary (H3). Willingness of usage is represented by usage frequency.

With respect to the effort exerted on vocabulary searching, it is not intuitively clear which type of loading is induced by the search process. From the CLT perspective, vocabulary search could be considered a germane cognitive load when the continuous rehearsal of the word form makes a more durable representation before the mapping of the semantic information. In such cases, the search effort would be considered a deeper or elaborating process, which generally leads to improved learning (Anderson, 1995; Baddeley, 1998; Craik & Lockhart, 1972; Hulstijn, 2001). If this is accurate, we would expect the pop-up condition to produce the worst vocabulary learning due to the elimination of search effort. However, effortless search could potentially free up more WM resource and enhance comprehension performance in the pop-up condition. Therefore, in Hypothesis 4, we postulate that if the effort required to find a word induces a germane cognitive load, higher vocabulary learning but lower comprehension performance is expected in the book condition (compared to the other two aids). Vocabulary scores in the pop-up condition will be lower than those in the book condition, yet better comprehension performance is expected (H4).
Conversely, the effort exerted to find a given word could induce an extraneous cognitive load. In addition, when definitions are presented on a physically separate location, readers must shift their attention from the main reading task, which might interfere with the comprehension process due to the split-attention effect (Cohen, 1990; Kalyuga, Chandler, & Sweller, 1999; Nation, 1990; Pressley, 2002; Sweller et al., 1998). In this case, pop-up and type-in dictionaries will lead to improvement in performance on both comprehension and vocabulary learning due to the minimization of the extraneous cognitive load of finding a word. Because the exact nature of the effort required to find a word (i.e., the vocabulary search process) remains unexplored, the goal of this study is to examine the cognitive process involved in this search. Thus, Hypothesis 5 states that if the effort required to search for a word induces an extraneous load, vocabulary and comprehension scores for the pop-up dictionary will be higher than those for the other two aids. Both vocabulary and comprehension performance should be lowest in the book condition (H5).

Regarding the effort spent on text and vocabulary reading (which are indicated by text reading time and vocabulary reading time, respectively), the nature of both should be clearer than that of the search effort. Like text reading time, vocabulary reading time is used to represent the effort spent on vocabulary reading. We hypothesize that the efforts exerted on text and vocabulary reading will increase the germane cognitive load and result in a positive effect on subsequent learning. It is expected that more effort spent on vocabulary and text reading will result in better performance. Therefore, according to CLT, Hypothesis 6 states that vocabulary reading time should be positively correlated with vocabulary learning (H6). Similarly, in Hypothesis 7, we expect that text reading time should be positively correlated with comprehension (H7).

Finally, a control condition (in which participants receive no dictionary aid) is also included in the experiment to examine the effect of using a dictionary while reading. Any differences observed among dictionaries could potentially be determined by an increase/decrease in overall performance with respect to performance in the control condition. We hypothesize that having a dictionary will be beneficial for both reading comprehension and vocabulary learning. Thus, Hypothesis 8 states that comprehension performance will be higher in all dictionary conditions than in the control condition (H8). Also, in Hypothesis 9, we expect that vocabulary learning performance will be enhanced in all dictionary conditions when compared to the control (H9).

3. Methods

3.1. Participants

Eighty college students enrolled in university in Tao-Yuan in Taiwan participated in the study. All participants were between the ages of 18 and 20 (M = 18.3 years, SD = 0.52 years; 33 females and 47 males, 41% and 59%) whose native language was Mandarin Chinese and learned English as their second language. To ensure similar English proficiency, participants were selected from a group of first-year college students enrolled in mid-level English classes; all had scored in the 50–69 range from a possible 100 points on the English proficiency classification exam administered at the university.

3.2. Design

The treatment (independent variable) used in the experiment is referred to as the "dictionary form." It includes the following four conditions: (1) pop-up dictionary; (2) type-in dictionary; (3) conventional book dictionary; and (4) no aid (control).

In the pop-up condition, students were instructed to double-click the mouse to bring forward a word definition window (Fig. 1a). In the type-in condition, students double-clicked the mouse to produce an entry window into which they typed the unknown word to retrieve the definition (Fig. 1b). Because copy-and-paste is restricted by the program, students had to strictly follow the aforementioned procedure to retrieve the definitions. Students in the conventional aid condition were provided with a book dictionary while reading (Fig. 1c), and no aids were given to the control group (Fig. 1d). All words in the reading text could be found in all three dictionaries, and participants were free to check any words they wished.

In both the type-in and pop-up conditions, the rule for which the definition window appeared on screen relative to the chosen (searched) word was the same between conditions (see Fig. 1c and d for comparison). This resulted in an identical learning environment that differed only in terms of the effort required to retrieve the definition. In both conditions, because the definition window covered the subsequent reading passage and remained visible until participants actively closed the window, this helps to distinguish the time for reading the dictionary content from time to reading the text (definitions regarding to times measured in the study will be provided in the variables and time measurement sections later). In addition, to keep the contents of the three dictionaries the same across conditions, a specific printed version of a book dictionary was created for this study. Words that did not appear in the text are omitted in all dictionaries to reduce the possible distractions from irrelevant words. Therefore, the total possible words that can be found in the dictionaries are the same across conditions. The size of the printed version of the dictionary is about half of an A4 paper and contain 344 pages.

English proficiency, reading materials, dictionary content and reading environment were either controlled or standardized in all conditions. To ensure that they possessed similar English proficiency, students were first randomly assigned to each condition. Analysis of variance (ANOVA) of students’ English proficiency classification scores indicated that no differences were found between the four groups (F < 1), suggesting that students were similar in their English abilities before the treatment and across conditions. The reading material selected for the study was a science topic on the importance of ocean and water preservation (Barnett, 1954). The section of the article was selected (contained 517 words) and revised. It contained four paragraphs (~ 130 words each) and was used in all four conditions. Low-frequency words were used to ensure a consistent level of difficulty. Furthermore, the information contained in each dictionary was the same across conditions to eliminate any difference in vocabulary-learning material. Finally, the reading apparatus was a PC with a 17-in. monitor, which simulated an online reading environment. The computer was programmed to record the number and duration of "look-ups." Students used the Safari browser to read the text and gain access to the type-in or pop-up dictionary aids, depending on their assigned group.

ANOVA was used to detect any significant differences between groups for all measurements. Post hoc comparisons using the Scheffe test were conducted after a significant effect was found with ANOVA. Correlational analysis was also used to investigate the relationship between cognitive processes and performance. Statistical significance for all tests was set at p < .05.

3.3. Instruments

Two tests were developed for the evaluation of students’ vocabulary learning and reading comprehension.
3.3.1. Vocabulary matching test

A pilot study was conducted to develop the vocabulary matching test. Six additional college students were selected from the same university and were asked to look over the reading material and highlight any unknown words. Items that were indentified by at least four students were selected for the vocabulary matching test, yielding a total of 24 items. Because it has been shown that the relevance of the words to the task affects the chance of dictionary consultation (Hulstijn, 1993), two teachers from the university's English Department were asked to select 15 of the 24 words as the vocabulary test set. Their selections were based on the words' relevance to the text to increase the chances that they would be looked up. This procedure also verified and increased the expert validity of the test. In the vocabulary matching test, students were asked to match the 15 pre-selected items to their definitions; all words and definitions were presented at once. For any given test item, participants had to select the correct definition from the 16 choices presented (an additional none-of-above choice was included to increase the level of difficulty). Participants could not merely rely on their recognition memory because all 15 definitions were associated with words that had previously appeared in the reading. Total score for the vocabulary test was 15 but later converted into a 100% scale. The vocabulary test covered about 5.38% (15 out 279) of non-repeated words that appeared in the text. The inter-reliability (KR20) of the vocabulary test was .713.

3.3.2. Comprehension test

To develop the reading comprehension test, we first developed 16 multiple-choice questions, which were revised by two English teachers to establish expert validity. The six students who took the vocabulary test also answered multiple-choice questions to determine the suitable items (averaging accuracy ratings of 60–80%). Items that had too high or too low scores were eliminated. Finally, 10 items were selected for the reading comprehension test set. Students received 10 points for each correct answer and 0 points for each incorrect answer. Hence, the maximum total score was 100. The inter-reliability (KR20) of the comprehension test was .741.

3.4. Variables

For research purposes, nine variables related to cognitive processes and learning performances were proposed in this study. The following sections will introduce each variable and its corresponding indicator.

The following four variables concerned students’ cognitive processes:

- Degree of willingness to use a dictionary was represented by consultation frequency while reading.
- Effort exerted to find a word in the dictionary was represented by the average vocabulary search time that student spent on locating the individual unknown words in the dictionary.
- Effort exerted to read a word in the dictionary was represented by the average vocabulary reading time that students spent on reading the definition of the unknown word.
- Effort exerted to read text was represented by text reading time that students spent on reading the text.

The following five variables concern students’ learning performance:

- Overall vocabulary performance was represented by students’ scores on the vocabulary matching test.
- Performance on checked items was represented by the number of correct items in the checked items. Here, “Checked item” refers to the consulted words that were also assessed on the vocabulary test.
• Accuracy of vocabulary learning was represented by performance on checked items divided by the total number of checked items.
• Learning efficiency for vocabulary was represented by the accuracy of vocabulary learning divided by the average dictionary using time (average vocabulary searching plus reading time) (Rasch & Schnitz, 2009).
• Reading comprehension performance was represented by students’ scores on the reading comprehension test.

3.5. Time measurement

This section provides details on how time measurements were recorded in the current experiment. Vocabulary search time in the pop-up condition was obtained using the time required to double-click the mouse. Vocabulary reading time in the pop-up condition was measured from the time the definition window appeared until the participant closed it. Vocabulary search time for the type-in condition was measured from the initial double-clicking (to open a type-in window) until participants had typed in the desired word and hit the enter key (keyboard or on screen button) to retrieve the definition window. Vocabulary reading time for the type-in condition started from the enter command and ended with the close of the definition window. Vocabulary search time for the book conditions started from the moment when participants flipped the page and ended when they found and pointed to the target word. Participants were instructed to point to the target word immediately after finding it and to keep their finger on the appropriate line and page while reading. Vocabulary reading time was determined as the moment that participants’ finger touched the page until the finger was removed from the page. Such instruction was necessary in order to determine which word the participant was searching for and how long he or she spent reading the definition.

3.6. Procedure

The experiment included three phases: instruction, treatment, and testing. In the instruction phase, participants were first asked to complete a consent form (for participation and video recording) and a background survey (age, gender, major, etc.). Participants were randomly assigned and escorted to each experimental condition. Students completed a practice session on how to use a particular dictionary (if provided), which also provided them with time to familiarize themselves with the tool and the environment. No other special instructions were given, except students in all four conditions were told to focus on comprehending the text passage and were encouraged to use the dictionary (if provided). They were also informed that there was no time limit on reading and that there would be 10 multiple-choice comprehension questions. However, students were NOT told about the surprise vocabulary test for the purpose of incidental learning. The comprehension test was administered before the vocabulary test. It also served as an intervening task for prolonging the time between the treatment and vocabulary test to ensure that the students’ vocabulary knowledge was not based on immediate recognition.

All students were video recorded during the treatment phase. Data on searched words, search time, and reading time were extracted from the video for the book dictionary, whereas the detailed data for the other two aids were automatically logged into the computer upon each consultation. Students were compensated after finishing the experiment.

4. Analysis results

4.1. Cognitive processes

In this section, variables related to dictionary usage (including average vocabulary search time, vocabulary reading time, and frequency of usage) are analyzed and compared across the three dictionary conditions; the text reading time is analyzed and compared across all four conditions. Descriptive statistical analyses are summarized in Table 1.

4.1.1. Average vocabulary searching time

As an indication of how much effort students spent finding individual unknown words in the dictionary, ANOVA of average search times revealed a significant difference across conditions: F(2, 57) = 124.69, MSE = 3562.84, and p < .001. Post hoc tests indicated that all conditions were significantly different from each other. The pop-up condition had the shortest average search time, which was significantly shorter than those of the book (p < .001) and type-in (p < .05) dictionaries. Search time for the type-in dictionary was significantly shorter than that of the book condition (p < .001).

4.1.2. Average vocabulary reading time

ANOVA also revealed a significant time difference in average vocabulary reading time: F(2, 57) = 12.06, MSE = 193.95, and p < .001. More specifically, the pop-up condition had the shortest average vocabulary reading time when compared to the other two conditions (p < .001). The type-in and book conditions did not differ significantly from one another.

4.1.3. Usage frequency (consultation frequency)

The ANOVA also indicated that usage frequency varied across conditions: F(2, 57) = 13.46, MSE = 3827.27, and p < .001. Students had the highest consultation frequency when using the pop-up dictionary, followed by the type-in and then the book dictionaries. Usage frequency of the pop-up dictionary was significantly higher than usage frequency of the other two aids (p < .05), although consultation did not differ between the type-in and book dictionary conditions.

4.1.4. Text reading time

ANOVA indicated that reading time differed significantly across all four conditions: F(3, 76) = 4.66, MSE = 293077.99, and p < .01.

Table 1

<table>
<thead>
<tr>
<th></th>
<th>Pop-up</th>
<th>Type-in</th>
<th>Book</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average vocabulary search time</td>
<td>0.73 (0.03)</td>
<td>5.62 (1.30)</td>
<td>25.91 (9.14)</td>
<td></td>
</tr>
<tr>
<td>Average vocabulary reading time</td>
<td>4.13 (1.05)</td>
<td>7.94 (3.47)</td>
<td>10.2 (5.93)</td>
<td></td>
</tr>
<tr>
<td>Frequency of usage</td>
<td>43 (25.30)</td>
<td>22.1 (13.15)</td>
<td>16.85 (6.36)</td>
<td></td>
</tr>
<tr>
<td>Text reading time</td>
<td>526.84 (213.98)</td>
<td>589.11 (231.00)</td>
<td>447.56 (269.59)</td>
<td>733.76 (281.81)</td>
</tr>
<tr>
<td>Treatment time</td>
<td>736.19 (260.56)</td>
<td>907.27 (442.98)</td>
<td>1029.42 (277.65)</td>
<td>733.76 (281.81)</td>
</tr>
</tbody>
</table>

Note: Each cell contains the mean and the standard deviation (in brackets).
Time variable units are in seconds.
For the frequency of usage, the unit is the number of dictionary uses.
More specifically, reading time for the control group \((p < .05)\) was significantly longer than the book dictionary group, although it did not differ from the type-in or pop-up conditions. No other differences were found among the three dictionary conditions.

### 4.2. Learning results (performance)

In this section, variables related to students’ learning performance are analyzed. Descriptive statistical analyses are summarized in Table 2.

#### 4.2.1. Reading comprehension

Concerning text comprehension (comprehension scores), ANOVA indicated that there was not a significant main effect of using a dictionary across all four conditions: \(F(3, 76) = 1.53, \text{MSE} = 646.09, \text{and } p = .21\).

#### 4.2.2. Overall vocabulary performance

Concerning incidental vocabulary learning, as indicated by the surprise vocabulary test, revealed a significant main effect of using a dictionary: \(F(3, 76) = 2.98, \text{MSE} = 802.57, \text{and } p < .05\). Post hoc tests indicated that students who used any type of dictionary scored better on the vocabulary test than students without aids (controls). All three types of dictionaries enhanced vocabulary learning, but no differences were found between them.

#### 4.2.3. Performance on checked items

An analysis of the number of words participants answered correctly on the checked items revealed a significant main effect of dictionaries: \(F(2, 57) = 5.45, \text{MSE} = 52.27, \text{and } p < .01\), and the Scheffe test indicated that students who used the pop-up dictionary had a significantly greater number of correctly answered checked items than those who used the type-in or book dictionaries. The other two types of dictionary did not differ from each other.

#### 4.2.4. Accuracy of vocabulary learning

Accuracy was obtained by dividing the number of correct items from the checked item list by the total number of checked items. However, when considering accuracy, no difference was found across dictionaries: \(F(2, 57) = 0.162, \text{MSE} = 95.43, \text{and } p = .85\).

#### 4.2.5. Vocabulary-learning efficiency

Learning efficiency for each type of dictionary was obtained by dividing the accuracy on the checked items by the sum of the average vocabulary search and reading times. This equation provided an indication of how well the student learned the vocabulary in a given amount of time. When analyzed using ANOVA, a significant difference was found: \(F(2, 57) = 69.65, \text{MSE} = 975.53, \text{and } p < .001\), and the Scheffe test indicated that students who used the pop-up dictionary learned words significantly more quickly than those who used the type-in or book dictionaries. Learning efficiency in the type-in condition was also significantly better than that of the book condition.

#### 4.2.6. Correlational analysis

The correlational analysis of vocabulary reading time and overall vocabulary performance was not significant \((r = .06, \text{p } = .65)\). Correlational analysis did indicate a positive correlation between the text reading time of the three dictionary conditions and the respective comprehension scores \((r = .29, \text{p } < .05)\).

### 5. Discussion

The goal of this study was first to analyze how comprehension and vocabulary learning are mediated by different types of dictionaries. Different processes associated with various types of aids were then identified and examined in conjunction with hypotheses regarding how these processes affect learning. The nature of the process was explored according to the subsequent performance. The following sections will first discuss the cognitive processes and then learning performance. Hypotheses will be revisited and the relationships between cognitive processes and performance will also be discussed.

#### 5.1. Cognitive processes

Observations regarding behavioral patterns usually provide important clues as to the underlying cognitive processes involved during learning. Possible relationships between the cognitive processes and subsequent performance can be inferred from differences observed in behavioral patterns. In this section, dictionary usage behaviors (e.g., vocabulary search time, reading time, and frequency of usage) and text reading time will be discussed in detail. The natures of these behaviors are discussed in relation to the cognitive processes they represented in the next section. General behavioral patterns are depicted in Fig. 2.

As illustrated in Fig. 2, behavior patterns differed across the four conditions. This figure shows how different processes are distributed within the treatment, as well as the proportion of each process. Because no aid was provided, students in the control group spent all of their time to reading the text. Students in the pop-up group spent the least amount of time on vocabulary searching and reading the vocabulary. In contrast, students in the book group spent the longest time on vocabulary searching and reading, but the least amount of time on text comprehension. Frequency information is also depicted in the figure by the number of segments. Each line within a process represents a consultation; this also shows how each process is segmented by frequency.

Concerning the dictionary usage behaviors, we first examined the vocabulary searching time, vocabulary reading time, and frequency. As predicted in H3, students using the pop-up dictionary consulted it twice as many times (43 times) as students using the other two types of aids (22 and 17 times for type-in and book dictionaries, respectively). Such differences in consultation fre-

### Table 2

<table>
<thead>
<tr>
<th></th>
<th>Pop-up</th>
<th>Type-in</th>
<th>Book</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading comprehension score</td>
<td>49.44 (24.84)</td>
<td>50.00 (19.25)</td>
<td>47.22 (17.24)</td>
<td>37.78 (20.20)</td>
</tr>
<tr>
<td>Overall vocabulary performance</td>
<td>79.06 (18.50)</td>
<td>78.13 (18.08)</td>
<td>77.50 (12.13)</td>
<td>65.63 (15.64)</td>
</tr>
<tr>
<td>Performance on checked items</td>
<td>10.15 (3.18)</td>
<td>7.25 (2.75)</td>
<td>7.10 (2.92)</td>
<td>7.10 (2.92)</td>
</tr>
<tr>
<td>Accuracy of vocabulary learning</td>
<td>79.10 (25.6)</td>
<td>79.58 (26.76)</td>
<td>75.59 (20.00)</td>
<td>75.59 (20.00)</td>
</tr>
<tr>
<td>Vocabulary-learning efficiency</td>
<td>15.89 (5.55)</td>
<td>6.33 (3.20)</td>
<td>2.29 (1.00)</td>
<td>2.29 (1.00)</td>
</tr>
</tbody>
</table>

Note: Each cell contains the mean and the standard deviation (in brackets).

Reading comprehension score, overall vocabulary performance, and accuracy of vocabulary learning are on a 100% scale. Vocabulary-learning efficiency is measured as learning rate (accuracy divided by time spent).
The split-attention effect (i.e., the constant switching between searching and reading processes and between reading media from computer screen and paper material). This effect may have caused the interruption and dislocation of the previous reading position and thus increased the vocabulary reading time (Cohen, 1990; Nation, 1990). Extra wasted time was added to the length of the vocabulary reading time without benefitting performance. This assumption was again supported an exploratory correlational analysis, which revealed a significant positive correlation between the average vocabulary search time and reading time ($r = .35, p < .001$). This indicates that longer searches tend to increase reading time. However, the reasons why such correlations exist deserve greater experimental attention and direct empirical evidence. Future research might consider investigate why such relationships exist, as well as why differential vocabulary reading time was found despite the fact that the contents provided in the dictionaries were identical.

Reading comprehension and text reading time were also examined. Results indicate that the book group had the shortest text reading time when excluding vocabulary searching and reading time (see H2). A positive correlation was found between the text reading time and comprehension score, suggesting that longer reading times are associated with better comprehension (as predicted by H7). An interesting relationship between the processes is revealed when looking at the proportional distribution of each process within the treatments (see Fig. 2): when students spent more time searching and reading the vocabulary, they devoted less time to reading comprehension. This trend is observed when looking at all three dictionary conditions. For example, although the treatment time was longest in the book group, the text reading time was shortest due to the long vocabulary search and reading time. As previously explained, longer periods of disruption could be consequences of the split-attention effect.

One might question why the split-attention effect did not interfere with text reading, but did interfere with vocabulary reading. We believe the same negative effect also applied to the text reading time, but the magnitude of the effect differed in these two situations because of the time differences between them. For example, in the pop-up condition, the average text reading time (total text reading time divided by usage frequency) was 27.62 s and 4.13 s for the average vocabulary reading time. The negative effect only added a fraction of time to the text reading time, and thus the beneficial germane effect was still observable in the correlational findings. On the other hand, the addition of as little as 1 s of extraneous cognitive load time occupied a large proportion of the vocabulary reading time and had a substantial impact on vocabulary performance.

5.2. Learning performance

Performance is discussed in terms of comprehension and vocabulary learning. According to CLT, limited resources are shared by vocabulary reading and meta-cognitive process, such as reading comprehension. An increase of resource in one process can be at the expense of another. Therefore, we investigated whether such interactions exist when using various types of dictionaries. The following section will link the cognitive processes described above to both comprehension and vocabulary learning.

5.2.1. Performance on incidental vocabulary learning

For incidental vocabulary learning, a significant enhancement effect for the dictionary group was observed. Interestingly, students in all three types of dictionary conditions demonstrated better vocabulary learning than those in the no dictionary condition (as predicted in H9), although no noticeable differences were detected among them. Does this finding indicate that all dictionaries are equally effective learning tools? We do not believe so. Indeed, different patterns of results emerged when learning efficiency, performance on checked items, accuracy of the check items, and consultation frequency were taken into account. Learning efficiency, as operationally defined, is the accuracy of the checked items (see Section 3.4) divided by the sum of vocabulary search and reading times. When analyzing learning in terms of efficiency, significant differences were found across dictionary types. In other words, given the same amount of time spent reading the words, students using the pop-up dictionary were most efficient at learning the vocabulary.

Although learning efficiency is one indication of learning, one might question the validity of such an indicator given that the pop-up dictionary group had the shortest vocabulary search and reading times, and that accuracies on the checked items were similar across conditions. Thus, the efficiency effect could be attributed to differences in the denominator (time) instead of a true reflection of differences in learning. However, we still considered the pop-up dictionary to be a more efficient tool when attention was shifted to closely examine the accuracy of the checked items. This argument may seem unreasonable and ironic at first, especially given the fact that no differences were found for accuracy. However, as explained below, inference logic is quite clear and straightforward, as it is based on simple mathematical concepts.

The accuracy of the checked items was defined as performance on the checked items divided by the total number of checked words. To obtain equal accuracy on checked items between condi-
tions, students in the pop-up group would need to get more items correct on the checked items because the denominator is increased. Our results supported this fact by showing that students in the pop-up condition had more correct items on the checked items. In other words, students' performance when using the pop-up dictionary was not impaired, even though they had encountered more words. From a qualitative perspective, without sacrificing performance, students using the pop-up dictionary learned as well as those using the book and type-in dictionaries. However, from a learning quantity perspective, we argue that opportunities for which students were exposed to words was greatly increased when using the pop-up dictionary. This “exposing opportunities” argument is in accordance with frequency data (most exposures), providing converging evidence in support of superior learning in the pop-up group.

One possible explanation for better vocabulary learning in the pop-up group is the re-exposure to the checked words. Students were more likely to check the same word two or more times using the pop-up dictionary. Evidence supporting such incremental learning (multiple exposures) is incidental vocabulary learning is supported by Hulstijn et al. (1996), who found better learning on words that reoccurred in a passage multiple times. However, such learning is unlikely to account for the results obtained in this study. First, the occurrence of multiple checks for the same word happened only rarely in all conditions. Out of the 1476 total consultations (pooled across all conditions), only 25 were repeated checks. This limited amount of data was not sufficient for drawing any meaningful conclusions. Second, the accuracy attribution of repeatedly checked words was eliminated because the accuracy scores were collected for the first checked word only (i.e., the scores analyzed did not include the accuracy of rechecked words). This rules out the possible contamination of the data.

Finally, by evaluating vocabulary performance, two alternative hypotheses (H4 and H5) were tested against one another (e.g., the vocabulary search process is associated more with extraneous or germane cognitive load). If a search is germane in nature, we would expect higher performance in the book group than in the other two aid groups. If a search adds to the extraneous load, we would expect reduced vocabulary learning. The overall vocabulary test scores did not support either of theses hypotheses. However, as previously mentioned, the pop-up dictionary was more efficient than the other two aids when then data were closely examined. Therefore, we suggest that although the search process could potentially involve the germane process, it is more likely to be associated with extraneous cognitive load in this study. In the pop-up group, the time that was saved during the search process could be used to learn vocabulary. Thus, students using the pop-up dictionary would have increased exposure to the words without sacrificing or jeopardizing learning performance.

5.2.2. Performance on comprehension

Based on the vocabulary results, we suspect that vocabulary searching may be associated more with extraneous cognitive load and thus may negatively affect reading comprehension. For example, look-ups in the conventional book dictionary would be the worst case scenario due to the constant attention-switching, head-turning, and task-switching between media (monitor and paper-based). Because of information immediacy, the pop-up condition is expected to have the highest comprehension scores, followed by the type-in and then the book conditions (H5). However, such a differential effect was not found. Results are instead in accordance with other studies that did not find evidence of comprehension enhancement (Akbolut, 2007; Aust et al., 1993; Bensoussan et al., 1984), as overall reading comprehension did not differ across conditions (H8). Comprehension findings differed from the CLT prediction that there would be a differential effect of dictionaries on comprehension. Possible explanations for the lack of effects obtained by this and other studies are described below.

One possible reason for the lack of effect could be the differences in experimental procedures. Scores obtained in this study are lower than those of other studies (e.g., Knight, 1994; Leffa, 1992; Yeung et al., 1997). In fact, the students only achieved an average score of 48.2% (averaged over all four conditions). However, such low scores could be attributed to the test difficulty, as students were told in advance that referring back to the text during the comprehension test was prohibited. This restriction was intended to encourage students to take full advantage of the dictionary before they were confident in taking the exam. In other studies, the text was visible when taking the comprehension test (Knight, 1994; Leffa, 1992). Thus, these procedural differences might have minimized the effects of the dictionary. Evidence supporting this explanation is taken from a study by Akbulut (2007) in which no effect of multimedia annotations was found across groups when prohibiting participants from referring back to the text while taking the comprehension quiz.

In addition to the test difficulty, the English proficiency level of the participants should also be considered. Studies that obtained beneficial effects of dictionaries also recognized the influence of participants’ language proficiency on the results (Knight, 1994; Leffa, 1992; Yeung et al., 1997). When Knight (1994) divided the comprehension results according to proficiency level and analyzed the effects of the dictionary, no differences were found among the high verbal ability groups. Leffa (1992) compared electronic and traditional, paper-based dictionaries and also suggested that differences in comprehension become greater in less proficient participants. Furthermore, direct empirical evidence from comparisons between participants’ abilities and the presentation format of explanatory notes found an interaction effect indicating that the same explanatory notes may facilitate or interfere with performance, depending on the learners’ language expertise (Yeung et al., 1997). Thus, participants’ language proficiency had a great influence on the effects of the dictionary. Students who participated in this study had been continuously exposed to and studied in English during their formal education for at least 8 years, starting in fifth grade or earlier. Thus, differences in readers’ abilities may explain the absence of an effect of the dictionaries on comprehension (Knight, 1994).

In summary, we conclude that if recommendations must be made for which type of dictionary is the best to use in terms of facilitating comprehension and vocabulary learning, then the pop-up dictionary is the recommended aid. Although comprehension across the four conditions did differ, students were exposed to more words using the pop-up aid and had more opportunities to learn new words. The benefit of acquired words might be attenuated by the requirement of searching for the learned word, although this knowledge may be helpful in future readings.

6. Conclusions

In the process of reading comprehension, one of the fundamental requirements is the ability to decode or understand individual vocabulary items. The dictionary, which Thorndike (1991) described as “one of the most important instruments of instruction,” has developed into different forms, such as the computerized dictionary (e.g., online or a pre-installed program) and the portable electronic dictionary (e.g., pocket or cellular phone). Each type is associated with a different degree of effort, convenience, and information. As technology advances and the Internet becomes more popular, the combination of online reading and the use of com-
puter-mediated aids will inevitably become an important method of acquiring information.

In this study, we considered a very basic question: "How do changes in technology affect learning?" As technology advances, one must consider the effects of this technological convenience. Does effortless dictionary use enhance understanding of text and facilitate vocabulary learning, or does it have the opposite effect due to the reduced effort of the input process?

Types of dictionaries that varied in required search effort were examined based on their effects on comprehension and vocabulary learning. Regarding vocabulary learning, vocabulary searching was more likely to involve an extraneous cognitive load, and the dictionary consultations neither interfered with nor benefited the students’ understanding of the passage. This study also contributes to the evidence suggesting that the type of dictionary can enhance incidental vocabulary learning. Although the apparent beneficial effects appeared to be equal for all dictionary types, when examined closely, the pop-up dictionary revealed advantages in terms of frequency of use and learning efficiency.

This study differed from others because we controlled for the information provided by various dictionaries and focused on exploring the nature of its associated processes (e.g., vocabulary and text reading) in relationship to participants’ subsequent performance. More importantly, we stressed the important point of consideration of both comprehension and vocabulary learning when evaluating the effectiveness of the reading aids. As the data suggest, complex interactions exist between these two types of learning, such that enhancement in one may be at the cost of another.

Although the results are intriguing and the corrected data seem adequate for drawing certain conclusions, nevertheless, some issues could be improved in future research. First, the power to confidently draw a conclusion is always limited by the size of the sample, especially when there is a null effect on comprehension performance. There is always the concern that the lack of an effect is due to a small sample size. Second, when college students are used as participants, one must be cautious about the generalizability of the data because of the restricted age range. This is particularly true in present experiment, as a restricted age group was used to control English proficiency level. A third and similar concern relates to the limitations on individual differences and the use of a uniform population is the generalizability of the result. In a future follow-up investigation, a larger sampling of different populations would help to address the problems mentioned above, as well as to provide converging evidences of our results.

The following areas are additional recommendations for future research. First, individual differences, such as English proficiency and learning style, are important factors to consider when investigating the learning effects of online reading mediated by computer aids. As mentioned above, differences in language proficiency could have a great impact on comprehension and dictionary usage (Knight, 1994; Leffa, 1992; Yeung et al., 1997). A direct empirical study examining the interaction between the types of aids and proficiency levels could have important implications for adaptive learning in formal and adult education. Second, research investigating incremental learning could also examine the impact of repeated checked-word frequency on incidental vocabulary learning (Hulstijn et al., 1996). Third, as CLT suggests, a germane cognitive load could potentially be beneficial to later performance, whereas extraneous cognitive load could lead to impaired performance. The precise nature of the searching and reading processes of the unknown word require further empirical investigation. It is interesting that the conventional book dictionary did not lead to better or worse performance in vocabulary learning, despite the fact that students devoted more attention to holding the word form in memory.

The data suggest that the vocabulary search process may be related more to extraneous loading than to germane loading. Understanding the nature of this process could prove helpful in developing future dictionaries and their functions.

Acknowledgment

The authors acknowledge the support of the National Science Council, Taiwan under Grant NSC 98-2628-S-008-001-MY3.

References


