Do Pocket Electronic Dictionaries Influence Word Retention and Reading Comprehension? Their Effects and Mediating Factors

Chiho Kobayashi

<Abstract>
This study explores the effectiveness of pocket electronic dictionaries (EDs) compared with printed dictionaries (PDs) on searching behavior, word retention, and reading comprehension. It also investigates how factors such as test formats, learners' proficiency level, and text difficulty are related to the effects of EDs. Thirty-six Japanese university students participated in the study. They read an English text and answered comprehension questions, while circling the words they looked up. They performed this task under two conditions: ED and PD. A week later, the students were given four types of vocabulary tests to assess various degrees of word retention. The findings show that the ED facilitated dictionary consultation. Moreover, the increased frequency of consultation led to improvement in reading comprehension. In particular, the low-proficiency students benefited greatly from the use of the ED. However, dictionary type did not influence word retention significantly. Regardless of test formats, retention rates were low in both the ED and PD conditions. A major factor mediating the effects of the ED on reading comprehension appeared to be the text difficulty, in particular the proportion of unknown words.

<keywords>
electronic dictionaries, second language learning, vocabulary acquisition, reading comprehension

Introduction

Although pocket electronic dictionaries (EDs) have grown in popularity among second language (L2) learners, some educators are still concerned about their possibly negative pedagogical effects. They are worried that EDs' interface design may encourage students to look at only the meaning that appears first in the dictionary. Furthermore, due to the ease of searching, EDs may prompt students to turn to their dictionary without considering textual context, resulting in a negative influence on reading comprehension and word retention. The goal of this study is to elucidate the impact of pocket EDs on searching behavior, word retention, and reading comprehension while Japanese university students are reading an English text. It also attempts to discern how factors such as test formats and learners' proficiency level are related to these effects.
Effects of EDs on L2 Learning

In recent years, quite a few studies have investigated the effectiveness of EDs in different formats, including desktop computer-based EDs and pocket EDs. Most of these studies have focused on the effects of EDs on L2 learning while reading. A body of studies examined the effects of EDs on incidental vocabulary learning while reading. Incidental vocabulary learning refers to the learning of words as a by-product of another activity, such as reading or communication, without the learner's conscious decision to learn them (Laufer & Hill, 2000; Laufer & Hulstijn, 2001). Others investigated the effects on dictionary use or textual comprehension while reading. Some studies examined various combinations of these factors.

Considerable research has been devoted to examining the effects of desktop computer-based EDs on L2 learning (Aust, Kelly, & Roby, 1993; Knight, 1994; Koga, 1995; Laufer & Hill, 2000; Leffa, 1992, Peters, 2007; Peters, Hulstijn, Sercu, & Lutjeharms, 2009). For example, Leffa (1992) compared the effectiveness of an electronic glossary relative to a printed dictionary (PD) for comprehending authentic texts. The findings show that the electronic glossary enabled the students to understand 38% more of the passage, using 50% less time. Similarly, Aust, Kelly, and Roby (1993) found that the ED group consulted over twice as many words as the PD group in nearly 20% less time. However, the two groups' text comprehension did not differ significantly.

Some studies have looked into pocket EDs (Iso & Osaki, 2003; Kobayashi, 2007; Koyama & Takauchi, 2003, 2004, 2007; Osaki, Ochiai, Iso, & Aizawa, 2003). Koyama and Takauchi (2003) investigated the relative effects of an ED versus a PD on learners' lookup behavior and word retention. There were no significant differences between the ED and PD groups in the number of words looked up or the search time. Moreover, no significant differences were found between the ED and PD groups in word retention as assessed by two types of tests: recall and recognition. Likewise, Koyama and Takeuchi (2004) found no significant differences between the ED and PD groups in search time and the quantity of retrieved information. Although the two groups showed no significant differences in word recall, the PD group outperformed the ED group in word recognition.

Kobayashi (2007) conducted a mix-method study in which both qualitative and quantitative data were collected from 22 Japanese university students. Retrospective think-aloud protocols were elicited in a session where participants read a 517 word passage using either an ED or a PD, answered comprehension questions, and then reported on lexical processing strategies (LPSs; ignore, consult, infer) used to deal with unknown words. A week after the reading session, the participants took two types of vocabulary tests to assess word retention: multiple-choice and open-ended. The use of EDs tended to result in an increase in the frequency of dictionary consultation, accompanied by varying degrees of decrease in the frequency of inferring. However, EDs did not appear to significantly influence word retention or reading comprehension.

The above studies on EDs have yielded mixed results as to their impact on L2 learning. Overall, the results of the studies seem to indicate that the use of EDs increases the frequency of dictionary consultation due to the speed and ease of searching. However, this increase in frequency does not necessarily improve word retention or reading comprehension.
Do Pocket Electronic Dictionaries Influence Word Retention and... — 3 —

Mediating Factors

This raises the question of why the changes in searching behavior do not have clear impact on word retention or reading comprehension. According to the depth of processing theory (Craik & Lockhart, 1972; Craik & Tulving, 1975; Hulstijn, 1992; Hulstijn & Laufer, 2001; Laufer & Hulstijn, 2001), processing information more elaborately will lead to higher retention. If so, shouldn’t the retention of words retrieved laboriously from PDs be greater than that of words retrieved with less effort from EDs?

Furthermore, it is also not convincing that the use of EDs does not influence reading comprehension greatly. When language learners engage in a reading task, they encounter too many unknown words to guess from context. However, if they frequently stop to look up a word in the dictionary, it will interfere with reading processes (Haynes, 1993). The use of EDs should alleviate this dilemma by allowing a learner to consult a dictionary quickly, which should influence comprehension positively.

One of the possible reasons for the mixed results obtained from the previous studies is that they were conducted using different groups of participants engaged in different types of tasks, suggesting that the effectiveness of EDs may vary depending on the conditions. More studies are needed to clarify the factors mediating their effects.

Accordingly, this study attempts to investigate the impact of pocket EDs on searching behavior, word retention and reading comprehension while Japanese university students are reading an English text. Moreover, in the discussion section, the results of the present study are compared with those of the researcher’s previous one (Kobayashi, 2007). The specific research questions addressed in the present study are as follows:

1. Are there any differences between ED and PD conditions in the frequency of dictionary consultation?
2. Are there any differences between ED and PD conditions in reading comprehension?
3. Are there any differences between ED and PD conditions in word retention?
4. What are some of the factors that influence the effects of EDs?

Method

Participants

Participants in the study were 36 Japanese students attending a mid-level private university in western Japan. They were all English majors whose years of study varied. They included 23 second-year students, 12 third-year students, and 1 fourth-year student. Among them, 15 were male, and 21 were female. The group’s level of proficiency varied. Their average Test of English for International Communication (TOEIC) score was 346.39 (SD=86.29), ranging from 220 to 550.

They were classified into high- and low-proficiency groups, according to their TOEIC scores, using the median split. The average score of the higher-proficiency group was 413.89 (SD=69.57), while that of the lower-proficiency group was 278.89 (SD=28.98).

Instruments

Reading text. Two articles taken from the Shukan ST were used. One was titled SOS (Appendix A), and the other Nonviolence for a Better World (Nonviolence; Appendix B). The two articles were similar in length and readability. The length of SOS was 380 words, and its...
grade level was 8.8. Nonviolence was 387 words, with a grade level of 8.3. Both articles were originally longer, but were edited by a native speaker to shorten them. Also, some of the words appearing in the texts were changed to less common ones to ensure that target words would be unfamiliar to all the students.

Thirteen target words were selected for each text. Prior to the selection, two students with slightly higher levels of English proficiency than the study’s participants circled all the unknown words in the two texts. The target words were selected from those that both students indicated to be unknown.


**Reading comprehension test.** A comprehension test was created for each text. It consisted of four open-ended questions. These questions focused on the paragraphs containing target words, making it difficult to answer them without knowing the meanings of these words. The questions had to be answered in Japanese.

**Vocabulary tests.** Vocabulary knowledge develops in an incremental manner (Schmitt, 2000). After one encounter, students may not learn the target words fully and only be able to recognize them. A possible reason that the researcher’s previous study (Kobayashi, 2007) did not detect any differences between the ED and PD groups is that the measures used were not sensitive enough to capture a small degree of word learning. Therefore, four types of vocabulary tests were designed for each text, in order to assess various degrees of word learning. The first was a recognition test that required indicating whether the students remembered seeing the word in the text (noticing test 1). It contained 13 target words with 10 distracters. The second was another recognition test that required indicating whether the students remembered looking up the word in the dictionary (noticing test 2). The third was a production test that required supplying word meaning by writing an L1 translation or L2 synonym (definition-supply test). The fourth was a recognition test that required matching the 13 target words with their Japanese translations, choosing from 18 alternatives (matching test).

**Procedure**

A week before the reading session, the students took a pre-test consisting of the target words plus dummy words in order to validate that all of the target words were indeed unfamiliar to them. They were asked to mark the words they thought they knew. Before the reading session, there was a 10-minute training period in which the students received instruction on how to use the ED and practiced using it. Then, students were randomly divided into two groups, and each group completed a reading task under two conditions: ED and PD. In order to ensure a counterbalanced design, the groups read different texts using different types of dictionaries. Group 1 read Text A with the PD and Text B with the ED, while Group 2 read Text A with the ED and Text B with the PD. For each reading task, the students were required to read a text, circle the words they looked up, and answer comprehension questions. Thirty minutes were allocated to complete each reading task.

A week later, unexpectedly, the students were given two sets of the four types of vocabulary tests to assess word retention. They took the vocabulary tests in this order: the noticing test 1, the noticing test 2, the definition-supply test, and the matching test. This ensured that the pre-
ceeding tests would not provide additional hints for those that followed. Lastly, in order to figure out the proportions of unknown words, the students were asked to circle all the words in the texts that they had not known prior to the reading session. They were allowed to spend as much time as they needed for each test. Most students finished all the tasks within 30 minutes.

Analysis

The reading comprehension test was scored using a three-point scale (2 = correct, 1 = partially correct, 0 = incorrect). The researcher and the second rater scored the students’ answers separately. Initial inter-rater reliability was 87.2% for SOS and 91.2% for Nonviolence. Any disagreements were resolved through subsequent discussion to achieve 100% consensus.

The noticing test 1 and the noticing test 2 were scored as either remembered or not remembered (1 = remembered, 0 = not remembered). Similarly, the matching test was scored as either correct or incorrect (1 = correct, 0 = incorrect).

The definition-supply vocabulary test was scored using a three-point scale (correct = 1, partially correct = 0.5, incorrect = 0). The researcher and the second rater scored the students’ answers separately. The initial inter-rater reliability was 99.8% for both the test based on SOS and the one based on Nonviolence. Any disagreements were resolved through discussion to achieve 100% agreement.

The vocabulary and reading test scores were compared between the ED and PD conditions, using paired t-tests. In addition to comparing the ED and PD conditions for the participants as a whole, separate paired t-tests were performed to compare the ED and PD conditions within each proficiency-group. Prior to performing these paired t-tests, the nature of the data was carefully checked. If the data was not normally distributed, the Wilcoxon signed-ranks test was performed instead. For all the analyses, the significance level was set at .05.

Results

Lookup Behavior

The ED and PD conditions were compared in the frequency of consultation in two ways: the total number of words looked up and the number of target words looked up. As can be seen in Table 1, the students looked up more total words as well as more target words in the ED condition than in the PD condition. The results of the paired t-tests showed that the differences between the ED and PD conditions were statistically significant for both the total words, \( t(35) = 4.627, p < .001 \), and target words, \( t(35) = 3.69, p = .001 \). This suggests that the ED facilitated dictionary consultation.

However, the students in both the ED and PD conditions looked up less than half of the target words (6.33 words for the ED condition and 4.53 words for the PD condition). These results were rather unexpected, given that answering comprehension questions required knowledge of the target words.

Table 1. Frequency of Dictionary Consultation by Dictionary Type

<table>
<thead>
<tr>
<th></th>
<th>ED Condition</th>
<th>PD Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( N )</td>
<td>( M )</td>
</tr>
<tr>
<td>Total Words</td>
<td>36</td>
<td>17.78</td>
</tr>
<tr>
<td>Target Words</td>
<td>36</td>
<td>6.33</td>
</tr>
</tbody>
</table>

Note. Total Words, Max = 28; Target Words, Max = 13.
**Reading Comprehension**

As can be seen in Table 2 below, the students understood the texts better in the ED condition than in the PD condition. The results of the paired $t$-test revealed that the difference was statistically significant, $t(35) = 3.14, p = .003$. This indicates that the ED helped the students with text comprehension. Taken together with the results on lookup behavior, it seems that the ED allowed the students to understand a larger number of unknown words, which led to better text comprehension.

Table 2. Reading Comprehension by Dictionary Type

<table>
<thead>
<tr>
<th></th>
<th>ED</th>
<th>PD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>36</td>
<td>36</td>
</tr>
<tr>
<td>$M$</td>
<td>4.81</td>
<td>3.78</td>
</tr>
<tr>
<td>$SD$</td>
<td>2.00</td>
<td>2.13</td>
</tr>
</tbody>
</table>

*Note. Max = 8.*

**Word Retention**

Table 3 shows the comparison of mean scores between the ED and PD conditions for four vocabulary tests. The students scored higher in the ED condition than the PD condition on the noticing test 1 and scored the same on the noticing test 2. In contrast, they scored higher in the PD condition on the matching and definition-supply tests. However, none of these differences were statistically significant, $t(35) = -0.93, p = .359; t(35) = 0.1, p = 1; t(35) = -1.99, p = .054; z = -1.605, p = .108$, respectively. The difference found in the matching test approached the significance level ($p = .054$). However, this seems to be due to an outlier. When another $t$-test that eliminated it was performed, the $p$-value fell away from the significance level, $t(34) = -1.719, p = .095$. This seems to show that dictionary type did not influence word retention greatly.

In both the ED and PD conditions, retention rates were very low. For example, even on the noticing test 1 that demanded the smallest degree of word learning (i.e., remembering seeing the word in the text), the students retained only 3.81 words in the ED condition and 3.47 words in the PD condition.

Table 3. Word Retention by Dictionary Type

<table>
<thead>
<tr>
<th></th>
<th>ED Condition</th>
<th>PD Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N$</td>
<td>$M$</td>
</tr>
<tr>
<td>Noticing 1</td>
<td>36</td>
<td>3.81</td>
</tr>
<tr>
<td>Noticing 2</td>
<td>36</td>
<td>0.75</td>
</tr>
<tr>
<td>Matching</td>
<td>36</td>
<td>0.97</td>
</tr>
<tr>
<td>Definition</td>
<td>36</td>
<td>0.17</td>
</tr>
</tbody>
</table>

*Note. Max = 13.*

Additional sets of analyses including only the words looked up were performed, due to the concern that the effectiveness of the PD for word retention might have been offset by the relative infrequency of consultation. In these analyses, the students scored higher on all the four tests in the PD condition than in the ED condition (see Appendix C). However, none of these differences were statistically significant.

**High Proficiency Group**

In order to discover whether students’ proficiency level influenced the effects of EDs, the two dictionary conditions were compared within each proficiency group. Table 4 presents the frequencies of dic-
Do Pocket Electronic Dictionaries Influence Word Retention and... — 7 —

tionary consultation of the high-proficiency group in both conditions. The students in the high-proficiency group not only consulted more total words but also more target words in the ED condition than in the PD condition. The results of the paired $t$-tests show that the differences were statistically significant, $t(17) = 3.492, p = .003; t(17) = 3.078, p = .007$, respectively.

Table 4. Frequency of Dictionary Consultation by Dictionary Type for High-Proficiency Group

<table>
<thead>
<tr>
<th></th>
<th>ED Condition</th>
<th>PD Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N$</td>
<td>$M$</td>
</tr>
<tr>
<td>Total Words</td>
<td>18</td>
<td>18.89</td>
</tr>
<tr>
<td>Target Words</td>
<td>18</td>
<td>7.56</td>
</tr>
</tbody>
</table>

Note. Total Words, Max=28; Target Words, Max=13.

As Table 5 shows, the high-proficiency students scored better on the reading comprehension test in the ED condition than in the PD condition. However, the difference was not statistically significant, $t(17) = 1.56, p = .137$.

Table 5. Reading Comprehension by Dictionary Type for High-Proficiency Group

<table>
<thead>
<tr>
<th></th>
<th>$N$</th>
<th>$M$</th>
<th>$SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ED</td>
<td>18</td>
<td>5.33</td>
<td>2.00</td>
</tr>
<tr>
<td>PD</td>
<td>18</td>
<td>4.61</td>
<td>1.82</td>
</tr>
</tbody>
</table>

Note. Max=8.

They scored better in the ED condition than in the PD condition on the noticing test 1, while they scored better in the PD condition on all the other vocabulary tests (see Appendix D). None of the differences were statistically significant, however.

**Low Proficiency Group**

As Table 6 shows, the low-proficiency group looked up more total words as well as more target words in the ED condition than in the PD condition. The results of the paired $t$-tests show that the difference between the ED and PD condition was significant in the total number of words looked up, $t(17) = 3.004, p = .008$, but not in that of target words looked up, $t(17) = 2.078, p = .053$, although the difference approached the significance level.

Table 6. Frequency of Dictionary Consultation by Dictionary Type for Low-Proficiency Group

<table>
<thead>
<tr>
<th></th>
<th>ED Condition</th>
<th>PD Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N$</td>
<td>$M$</td>
</tr>
<tr>
<td>Total Words</td>
<td>18</td>
<td>16.67</td>
</tr>
<tr>
<td>Target Words</td>
<td>18</td>
<td>5.11</td>
</tr>
</tbody>
</table>

Note. Total Words, Max=28; Target Words, Max=13.

As can be seen in Table 7, the students understood the texts better in the ED condition. The results of the paired $t$-tests show that the difference was significant, $t(17) = 2.871, p = .011$. This indicates that the

Table 7. Reading Comprehension by Dictionary Type for Low-Proficiency Group

<table>
<thead>
<tr>
<th></th>
<th>$N$</th>
<th>$M$</th>
<th>$SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ED</td>
<td>18</td>
<td>4.28</td>
<td>1.90</td>
</tr>
<tr>
<td>PD</td>
<td>18</td>
<td>2.94</td>
<td>2.13</td>
</tr>
</tbody>
</table>

Note. Max=8.
ED aided low proficiency students greatly with text comprehension. The students scored better in the ED condition than in the PD condition on the noticing tests 1 and 2. In contrast, the students scored better in the PD condition on the matching and definition-supply tests (see Appendix E). However, none of the differences were statistically significant.

In summary, the ED facilitated dictionary consultation. Moreover, the increased frequency of dictionary consultation brought about improvement in reading comprehension, especially for the low-proficiency group. However, dictionary type did not make significant differences in word retention.

Discussion

In this section, the results of the present study are discussed in comparison with those of the researcher’s previous one (Kobayashi, 2007). The results of the two studies are, when relevant, also compared with other studies.

Participants and Reading Task

The participants in the previous study, conducted in 2007, were 22 Japanese students from two universities. They were freshmen majoring in various subjects. Their mean scores in the Reading Comprehension section of the Test of English as a Foreign Language (TOEFL) was 21.32 out of 50. Overall, they seemed to have slightly higher levels of proficiency than the participants in the present study.

Table 8 presents the comparison of the texts used in the present study (SOS and Nonviolence) with that used in the 2007 one, titled New Ways to Die (New Ways). Although all three texts were adapted from those published in Shukan ST and similar in genre, they differed in difficulty and length. SOS and Nonviolence contained a larger proportion of unknown words than New Ways (7% versus 5%), although the grade levels were lower and the lengths were shorter. Therefore, the texts might have been more difficult for the students.

The participants in the present study had 30 minutes to read the text and answer comprehension questions. In contrast, the students of the 2007 study were given 40 minutes to read the text and answered comprehension questions orally afterward.

Considering the above conditions, the reading task assigned in the present study was probably more challenging for the students than that of the previous study; that is, they had to read texts with a higher proportion of unknown words within a shorter period of time. This speculation is confirmed by the students' lower reading comprehension scores, as shown below.

<table>
<thead>
<tr>
<th>Grade Level</th>
<th>SOS</th>
<th>Nonviolence</th>
<th>New Ways to Die</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length</td>
<td>386 words</td>
<td>396 words</td>
<td>517 words</td>
</tr>
<tr>
<td>Unknown Word</td>
<td>27 word (7%)</td>
<td>29 words (7%)</td>
<td>26 words (5%)</td>
</tr>
</tbody>
</table>

Lookup Behavior

Table 9 presents the frequencies of dictionary consultation for the present study as well as the 2007 one. Since the total numbers of words differ, the frequencies are converted to percentages to make
them comparable. Overall, the participants in the present study looked up a smaller proportion of unknown words than those in 2007. The students of the present study might not have been able to look up as many words as they wished, due to the higher proportion of unknown words and the time constraint.

In both studies, the students looked up more words in the ED condition than in the PD condition. This gives strong support to EDs' positive impact on the frequency of dictionary consultation.

Table 9. Frequency of Dictionary Consultation of the Present Study and the 2007 Study

<table>
<thead>
<tr>
<th>Dictionary Type</th>
<th>Word Looked Up</th>
<th>This Study</th>
<th>2007 Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>ED</td>
<td>PD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>65%</td>
<td>41%</td>
</tr>
</tbody>
</table>

Reading Comprehension

Table 10 presents the reading comprehension scores of the present study as well as the 2007 study, converted to percentages. Reading comprehension scores of the present study fall short of those of the 2007 one, in both the ED and PD conditions. This verifies that the reading tasks assigned in the present study were more challenging for the students.

In this study, the students understood text considerably better in the ED condition than in the PD condition. In contrast, the scores of the previous study's ED and PD groups did not differ greatly. This seems to show that students benefit from EDs, especially when reading challenging texts with a high density of unknown words and a time constraint, like the participants in the present study did.

Table 10. Reading Comprehension Scores of the Present Study and the 2007 Study

<table>
<thead>
<tr>
<th>Dictionary Type</th>
<th>Reading Comprehension</th>
<th>This Study</th>
<th>2007 Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>ED</td>
<td>PD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>60%</td>
<td>47%</td>
</tr>
</tbody>
</table>

Word Retention

Table 11 presents the vocabulary test scores of the present study as well as the 2007 one, converted to percentages. Since only the matching test and definition-supply test were administered in the 2007 study, the scores of these two types of tests are compared. The students of the present study retained far fewer words than those of the previous one, regardless of dictionary type. This seems to be largely due to the research designs used. In the previous study, after completing the reading task, the students were required to discuss the LPS used to deal with each unknown word. This seemed to give them additional learning opportunities. Another possible reason is the difficulty of the reading task assigned in the present study: due to the more demanding material, the students might have used most available attentional resources understanding the text, leaving few to remember the words.

In both studies, the word retention rates did not differ greatly between the ED and PD condition. This contradicts the depth of processing hypothesis. For the 2007 study, the measures might not have been sensitive enough to capture a small degree of word learning. Accordingly, the noticing tests 1 and 2 were used in this study, in addition to the definition-supply test and the matching test. However, no significant differences were found between the ED and PD conditions in any of the four measures. A possible reason is discussed in the next section.

One of the possible reasons that no difference was found between the ED and PD condition in this study is the floor effect; that is, since the students scored very low in both the ED and PD conditions,
the comparison became somewhat questionable. The results of the present study need to be reexamined in future research in which participants can perform better.

Table 11. Vocabulary Test Scores of the Present Study and the 2007 Study

<table>
<thead>
<tr>
<th>Dictionary Type</th>
<th>This Study</th>
<th>2007 Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ED</td>
<td>PD</td>
</tr>
<tr>
<td>Matching</td>
<td>7%</td>
<td>11%</td>
</tr>
<tr>
<td>Definition-Supply</td>
<td>1%</td>
<td>2%</td>
</tr>
</tbody>
</table>

**Mediating Factors**

The results of the present study highlight the important role that learners' proficiency levels play in determining the effects of EDs on reading comprehension. The low-proficiency group understood the texts better in the ED condition than in the PD condition, but the high-proficiency group did not. This suggests that EDs particularly benefit low-proficiency students.

The effectiveness of the ED for each group seems to be related to the proportions of unknown words. The proportion of unknown words was 9% for the low-proficiency group and 5% for the high-proficiency group. For the low-proficiency group, the ED seemed to decrease the proportion of unknown words and thus improve text comprehension. In contrast, the high-proficiency group, who knew many of the surrounding words, might have been able to guess word meaning from context. This speculation is supported by Nation's (2001) assertion that knowledge of 95% or more of words makes it possible to understand the text and guess word meaning from context (Nation, 2001). Therefore, the increased frequency of consultation in the ED condition might not have led to better comprehension for the high-proficiency group.

Similar results were obtained elsewhere. These studies also used texts difficult for students' proficiency levels (e.g., Aust et al., 1993; Koga, 1995; Leffa, 1992; Osaki et al., 2003). For example, in Leffa's (1992) study, beginning students read authentic texts. In Koga's (1995) study, intermediate students read a text taken from the Graduate Record Examination (GRE). In both studies, the ED group outperformed the PD group in text comprehension. In contrast, other studies that used less difficult texts (e.g., Iso & Osaki, 2004; Koyama & Takeuchi, 2007) did not find significant differences between the ED and PD conditions.

This corroborates the results of the comparison between this researcher's studies, as discussed above. The students of this study read texts with 7% of unknown words with a time limit. In contrast, the students of the previous study read a text with 5% of unknown words without a strict time limit. As a result, the ED benefited the students of this study, but was not so helpful to the students in 2007.

Taken together, text difficulty, especially the proportion of unknown words, seems to influence the effects of EDs on reading comprehension. EDs seem to improve students' reading comprehension greatly when they are reading challenging texts, that is, texts with more than 5% of unknown words, especially when there is a time constraint. This suggests that the use of EDs should be highly recommended on such occasions.

In the present study, the PD did not facilitate word retention in either the high- or low-proficiency groups. Moreover, as discussed above, there was no relation between test formats and the effects of dictionary type on word retention, either. This suggests that dictionary type does not affect word retention, regardless of such factors as learners' proficiency levels and test
Do Pocket Electronic Dictionaries Influence Word Retention and… — 11 —

formats. This may be because dictionary type alone does not determine what kind of lexical processing the word receives and thus whether it is retained or not. Other factors, such as learners’ previous experiences and the characteristics of words, also come into play in determining the nature of lexical processing. More importantly, dictionary consultation is not the only LPS, and learners deal with unknown words through other strategies, such as guessing from context.

Regardless of test formats and students’ proficiency levels, word retention rates were low in both the ED and PD conditions of this study. This corroborates the results of previous research (e.g., Hulstijn, 1992; Krashen, 1989; Paribakht & Weshe, 1999). Students seem to learn only a small portion of words through reading for comprehension. Even with access to a dictionary, they look up relatively few unknown words in the dictionary and forget many of those they do look up. This underscores the need to supplement reading practice with more focused vocabulary-building exercises.

**Conclusion**

The results of the present study show that EDs increase the frequency of dictionary consultation. In addition, EDs improve reading comprehension, although they do not influence word retention significantly. The findings also show that a major factor that mediates the effects of EDs on reading comprehension is text difficulty, in particular the proportion of unknown words. When students read a challenging text with a high proportion of unknown words, EDs have positive effects on text comprehension. In contrast, such factors as test formats and students’ proficiency levels do not influence the effects of dictionary type on word retention. Regardless of these factors, the students showed a low rate of word retention in both the ED and PD conditions.

There are two major limitations to this study. First, there were only 37 participants. In the future, a larger-scale study should be conducted to reexamine these results. Secondly, both the ED and PD groups scored very low on vocabulary tests, which made the interpretations of the results somewhat questionable. Future studies should use either shorter and less demanding texts or extend the reading time so that students can perform better on the vocabulary tests.

**Notes**

(1) Target words for SOS were calling, condone, lewd, obscene, overwhelming, pediactric, pedophilia, prostitution, raid, rampant, shudder, shut, weird. Target words for Nonviolence were amidst, bereft, emanate, entail, indignation, ominous, oppressive, perilous, preach, proliferation, respite, strife, wrath.

(2) $t(35) = 1.37, p = .18; z = -1.35, p = .177; t(35) = .159, p = .874; z = -1.633, p = .102.$

(3) $t(17) = .724, p = .479; t(17) = -678, p = .507; t(17) = -1.365, p = .19; z = -1.365, p = .461.$

(4) $t(17) = -5.66, p = .579; t(17) = -9.9, p = .381; t(17) = -1.41, p = .177; z = -1.382, p = .167.$

(5) The two universities are both located in western Japan. One is a prestigious national university, and the other is a mid-level private one. The private university is similar in its school culture to the university used as the research site of this study. Twelve students were from the national university, while ten were from the private university.

(6) Of the 22 participants, 10 majored in English, 7 agriculture, and 5 cross-cultural studies. There were 17 female students and 5 male students.

(7) The score of 21.32 corresponds to 44-46 in the Reading Comprehension section (Educational Testing Service, 2003). If a student scores the same in all three sections of the TOEFL, the
total score will be 440-460. According to the conversion formula, this score corresponds to a TOEIC score of 449-456 (Nakata, 2004).

Acknowledgements

I would like to express my deep gratitude to Casio Computer Co. Ltd. for lending me the EDs used in this study.

References


Do Pocket Electronic Dictionaries Influence Word Retention and...


Appendix A. Reading Text SOS

The recent arrest of an elementary school vice principal in Hokkaido on child smut charges has highlighted the tragedy of children being used for pornography and prostitution in Japan. When police raided his home, they found many obscene images of underage girls. He was even selling them online. I shudder to think of how many cases like this have gone unreported.

Pedophilia is rampant, all over the world. Western countries have been taking the problem seriously for at least two decades now. I can only hope that we Japanese finally stand up against this social disease. But I doubt it. Not only is child pornography apparently condoned here, but lewd magazines and comics are in stores where even children can easily see them. Even vending machines sell X-rated stuff on the street.

The other day in a park near my home, I was shocked to see two grade-school boys reading weird magazines that contained pornographic images. "How can this be possible?" I thought. Then I remembered that in Japan pornographic goods are even given away as prizes in some video game arcades.

The blind acceptance of child pornography and prostitution in Japan allows pedophilia to thrive. We must do something to stop pedophilia now before more innocent children become the victims of predators.

A pedophile doesn't have certain looks that distinguish him or her as such. The person could be married with children of his or her own. He or she could be a next-door neighbor, a tutor, or a teacher. Pedophiles tend to choose occupations that put them in close contact with children, such as teaching, counseling, day care, pediatric practice, etc.

The overwhelming majority of people in these callings are certainly not pedophiles. Nevertheless, we have to be aware of whom our children are spending their time with. Also, we must not allow the Internet to become an electronic babysitter for us. We must always pay attention to what our children are doing when they are online.

The environment in which we live is also the one in which our children grow up, and together we must share the problems facing youngsters. We must get pornographic vending machines off the streets, pornography out of children's comics and video game arcades, and fight to have child pornography completely banned.

Appendix B. Reading Text Nonviolence for a Better World

We are living in perilous times. TV news speaks of the ominous dangers the world is facing. Soldiers and civilians alike are being killed every day in Afghanistan, Iraq, Congo and Sudan. Still, there's military buildup all over the globe and fear of more confrontations. Amidst all these problems, what should be attitude of the average person be?

Speaking to a group of teachers and guardians just after 9-11, I referred to the wonderful song "Imagine," and stated that we should all put pressure on our leaders to adopt the philosophy of non-
violence in resolving strife. "Isn't that a form of escapism?" one man asked.

"Maybe," I replied. "And before the proliferation of violence destroys humanity, we're all going to be looking for the 'exit' signs."

"Imagine all the people, living life in peace." These words of hope were written by John Lennon more than three decades ago, and I believe they are what we presently need to build a better world. Because many people have been bereft of hope for the future, they have become violent and ineffective in life. A continual looking forward to a nonviolent world is not a form of escapism, but a virtue.

Great leaders such as Mahatma Gandhi and Martin Luther King, Jr., who preached nonviolence, had hope. Gandhi was a "warrior," but he never resorted to weapons to fight his battle. The indignation he felt toward the oppressive British colonialists was not mere anger; it was a wrath of love, faith and hope which emanated from the bottom of his heart. That's why he tried to achieve his goal with wisdom rather than machine guns.

A U.S. Gulf War veteran once said; "I only wish for a nonviolent and peaceful tomorrow with my wife and children. Aim at nonviolence, and you will get peace and freedom; aim at war and you will get neither."

Historically, soldiers destroyed others' homes, while desiring peace for their own. Peace has been a temporary relief, a respite between battles entailing much sacrifice. The successive civilizations of the past have all been violent to some degree; there has never been a lasting, peaceful social order. It is impossible to build a peaceful world on the cracked foundation of violence. The cycle of war and peace—can this karma be overcome without the spirit of nonviolence?

### Appendix C. Word Retention by Dictionary Type for Only the Words Looked Up

<table>
<thead>
<tr>
<th></th>
<th>ED Condition</th>
<th>PD Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>M</td>
</tr>
<tr>
<td>Noticing 1</td>
<td>36</td>
<td>0.27</td>
</tr>
<tr>
<td>Noticing 2</td>
<td>36</td>
<td>0.10</td>
</tr>
<tr>
<td>Matching</td>
<td>36</td>
<td>0.12</td>
</tr>
<tr>
<td>Definition</td>
<td>36</td>
<td>0.01</td>
</tr>
</tbody>
</table>

*Note. Max = 1.*

### Appendix D. Word Retention by Dictionary Type for High-Proficiency Group

<table>
<thead>
<tr>
<th></th>
<th>ED Condition</th>
<th>PD Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>M</td>
</tr>
<tr>
<td>Noticing 1</td>
<td>18</td>
<td>4.67</td>
</tr>
<tr>
<td>Noticing 2</td>
<td>18</td>
<td>0.50</td>
</tr>
<tr>
<td>Matching</td>
<td>18</td>
<td>0.72</td>
</tr>
<tr>
<td>Definition</td>
<td>18</td>
<td>0.28</td>
</tr>
</tbody>
</table>

*Note. Max = 13.*

### Appendix E. Word Retention by Dictionary Type for Low-Proficiency Group

<table>
<thead>
<tr>
<th></th>
<th>ED Condition</th>
<th>PD Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>M</td>
</tr>
<tr>
<td>Noticing 1</td>
<td>18</td>
<td>2.94</td>
</tr>
<tr>
<td>Noticing 2</td>
<td>18</td>
<td>1.00</td>
</tr>
<tr>
<td>Matching</td>
<td>18</td>
<td>1.22</td>
</tr>
<tr>
<td>Definition</td>
<td>18</td>
<td>0.06</td>
</tr>
</tbody>
</table>

*Note. Max = 13.*