The Retention of L2 Vocabulary in an EFL Incidental Learning Setting : A Pilot Study of Four Different Word-retrieval Tasks during Reading

関西外国語大学 研究論集 第80号（2004年8月）
Journal of Inquiry and Research, No.80 (Aug. 2004)

The Retention of L2 Vocabulary in an EFL Incidental Learning Setting: A Pilot Study of Four Different Word-retrieval Tasks during Reading[1)]

Osato Shiki

1. Introduction

Vocabulary plays a crucial role in second-language (L2) reading skills. The fact that the vocabulary size that enables students to start reading naturally in L2 is above the threshold level of 3,000 to 5,000 words has been empirically supported in the field of L2 reading (Coady, 1997; Qian, 1999; Nation, 2000). The importance of vocabulary has also been recognized in L2 reading research designed in the light of cognitive psychology. In the cognitive perspectives of L2 reading, the decisive role of vocabulary is often equated with the importance of bottom-up processing[2)]. The major arguments in this respect are based on the findings that there is a limitation in working memory capacity[3)] (Daneman & Carpenter,1980; Daneman & Merikle, 1996; Harrington & Sawyer, 1992; Just & Carpenter, 1992), which explains that for better comprehension of L2 texts within the capacity limitation of working memory, bottom-up processing needs to be automatized so as to allow more memory capacity to be allocated for various types of higher order processing involving use of reading strategies (LaBerge & Samuels, 1974; Perfetti, 1985; Stanovich, 1992; Akamatsu, 2000). This view of looking at L2 reading processes implies that mastering an extensive number of L2 words can make a significant contribution to heightening the degree of automaticity at the bottom-up level since vocabulary is the basis to determine the quality of bottom-up processing. The present study supports this cognitive perspective of vocabulary learning, and, thus, explores how L2 vocabulary can best be learned.

How can teaching facilitate learners acquiring L2 vocabulary? This is surely one of the issues that has been much debated in L2 pedagogy. One of the issues related to L2 vocabulary learning is whether the types of learning are intentional or incidental. According to Hulstijn

(2001), "incidental vocabulary learning refers to the learning of vocabulary as the by-product of any activity not explicitly geared to vocabulary learning, with intentional vocabulary learning referring to any activity aiming at committing lexical information to memory" (p.271). For first-language (L1) vocabulary learning, researchers agree that a large number of L1 words are learnt incidentally (Nagy et al., 1987). For L2 vocabulary learning, it is understandable that direct and explicit instruction of vocabulary (e.g., teaching about meanings and grammatical functions of new words and memorizing words for vocabulary test) plays an important role and can successfully teach vocabulary (Zimmerman, 1997). Hulstijn (2001), however, stresses that even in the case of learning L2 "a large number of words cannot have been learnt solely by means of explicit vocabulary instruction; rather, most words are learned in an incremental way through repeated encounters during extensive reading" (p.271). Empirical studies also indicate that L2 learners have successfully learnt vocabulary incidentally through extensive reading (Day et al., 1991; Pitts et al., 1989). Accordingly, for learning under the limitations of EFL settings where students have difficulty increasing contacts with English words in their daily lives, it is highly recommended for students to take advantage of the "incidental" nature of vocabulary acquisition so as to expand their vocabulary size.

In general, in EFL reading instruction, students often read a variety of materials; thus, incidental vocabulary learning has already been taking place and facilitating students retaining new words as a by-product of the reading task. There are various resources and multiple fashions to retrieve meanings of words, but it can be predicted that there are certain ways of retrieval of meanings for new words, which are more effective for the retention of those words. Although there have been numerous previous studies of retention of L2 vocabulary with regard to effects of extensive reading programs on L2 vocabulary acquisition, there are, however, few studies aimed at the effects of incidental L2 vocabulary learning through regular reading tasks in EFL classrooms. The main purpose of this pilot study was, thus, to explore what factor or what specific way of retrieving vocabulary might best facilitate the retention of vocabulary learnt incidentally for university students in Japan in EFL reading instruction.

2. Literature Review

How vocabulary is learnt incidentally has been approached mainly from memory research perspectives that try to uncover how information is effectively retained to long-term memory[4]. Studies on memory at the beginning were conducted mainly to see the effect of repetition of in-

formation in short-term memory and postulated that the time of rehearsal in short-term memory is the main factor that facilitates the transfer of information from short-term memory to long-term memory. Later, Craik and Lockhart (1972) introduced the term, "depth of processing", emphasizing the degree to which knowledge is processed as a factor leading to better retention. The theory was supported by a study conducted by Craik and Watkins (1973) and a consecutive study by Craik and Tulving (1975), which claim that overt rehearsal in short-term memory is insufficient to enhance long-term memory performance; rehearsal, no matter how sufficiently done, needs to be elaborative in richer semantic-associative fashion to improve the level of retention.

The rationale that the semantic association is a driving force for retaining information has been widely accepted in L2 pedagogy, and the process has recently been mentioned more with the term "elaboration" with emphasis over the effect of "inferring meanings of words" on learning vocabulary (e.g., Redman & Ellis, 1990; Nation, 1994). According to Laufer and Hulstijn (2001), elaboration means "the more attention that is paid to the formal and semantic aspects of words and the richer the associations that are made with existing knowledge..., the higher are the chances that the new information will be retained" (p.1). In other words, for better retention of vocabulary, information processing needs to be elaborative in the way that it involves meaning assessment tasks that drive problem solving and inferences skills since the higher order processing such as making inferences requires more mental efforts (Hulstijn, 1992, p.113). In the incidental learning study conducted by Hulstijn (1992), this hypothesis was proved by statistic analysis. In Hulstijn's study, one group was given the multiple choices (MC) procedure in which synonyms and distractors were listed to help the subjects retrieve the meanings of target words and facilitate them inferring (or using more mental efforts) without wrong guesses while the other group were simply given synonyms or L1 translations of the target words, and the former group did better on the recall test of those target words given unexpectedly afterward.

Essentially, incidental learning is elaborative because the learning by itself encourages learners to infer from contexts and involve their prior knowledge. However, negative aspects are often pointed out: the major claim is that unskilled L2 readers often draw wrong inferences (Bensoussan & Laufer, 1984). The result indicated in the study using MC procedure conducted by Watanabe (1997) was congruent to this negative view. It showed that low proficiency Japanese learners could acquire the target words only when they were given L1 translations to understand them during reading but not when they made inferences with MC procedure, unlike

in the case of high proficiency Japanese learners. Many Japanese learners at university level, who possibly fall into unskilled learner category, may not be able to guess meanings correctly and therefore still may need confirmations of meanings of vocabulary with L1 translations. Fraser's (1999) finding can provide a solution for this problem with the assumption that inferring meanings of words needs to be performed with confirmations in L1 provided afterward since inferring meanings itself may not be enough to learn vocabulary for low-level students.

3. Method

3.1. Research Questions

As indicated in the literature review, it is hypothesized that retrieval of meaning involving making inferences from contexts (or prior knowledge) with confirmations in L1 afterward is more elaborating since it facilitates the exposure of information to the prior knowledge network in long-term memory; thus, the reading activity with guessing vocabulary with confirmations in L1 promises better retention of vocabulary. To study whether making inferences and confirmations in L1 are more beneficial to retention of L2 vocabulary, the present pilot study compared 4 different word-retrieval tasks: 1) guessing vocabulary from contexts with L1 confirmations afterward, 2) translating words into Japanese consulting a bilingual dictionary, 3) being given confirmations in L1 (being provided with an exact Japanese equivalent to each of the target words in that specific context), 4) using confirmations in L2 (being given English synonyms to the target words). The study analyzed effects of these retrieval tasks on the retention of vocabulary learnt incidentally through a reading task. Since task 1) and 2) were more likely to involve both making inferences and confirmations in L1 than task 3) and 4) which involved only confirmations with L1 (L2 synonyms for task 4), and since task 1 was more likely to make inferences than task 2 (task 1 directly told students to infer meanings, and task 2 indirectly asked them to make inferences by telling them to choose a Japanese equivalent for each of the target words from several translation choices on the dictionary), the present pilot study attempted to address the following research questions:

1) Is it possible for EFL students to retain words in the incidental learning setting?
2) Is making an inference plus confirmations with L1 a facilitating factor for vocabulary retention?

   If so, task 1 facilitates the retention of vocabulary better than the other tasks.

   Additionally, task 2 may facilitate the retention of vocabulary better than task 3 and 4.

3) Which method of retrieving meanings of words, using L1 translations or L2 synonyms, can facilitate retention better for low proficiency learners? And is there a difference when a making-inferences task is involved or not?

3.2. Subjects

The subjects were three classes of first year Japanese college students majoring in English in 2003. The subjects included both male and female students, and were homogeneous in their educational background, having studied English for 6 years in junior high school and high school. Three intact classes participated in the study. Subjects from two of the intact classes were divided into the 4 experimental groups, and subjects from the final class became a control group. After removing those who failed to fulfill some parts of the tasks required for this experiment, 64 were eventually used for the analysis. The number of subjects for each group was: G1 = 13, G2 = 13, G3 = 11, and G4 = 11, and G5 (control group) = 16. A TOEIC-mini test was used to measure their proficiency in English. The comparison of the mean scores of five groups indicated that the 5 groups were relatively equivalent and comparable in terms of English proficiency measured by TOEIC-mini test, F (4, 59)=1.11, p>.05.

3.3. Materials

The passage read by students for the experiment was taken from "*Timed Reading (Level 2)*" published for speed reading practice by Jamestown Publishers. The passage was about two famous rivers, the Nile and Yangzte, and its contents seemed unfamiliar to students. Since it was written with 400 frequently used words, it was expected that making inferences would be relatively easy for the students. 16 words were chosen from the passages and changed to non-words that do not actually exist so as to eliminate the chance that students knew or had encountered the target words previously. These non-words were checked by native speakers of English to determine whether they phonetically and morphologically sounded English.

A set of tests to measure how many words are retained included three vocabulary tests. The first one was the Word Recognition test (Test A), which showed a list of words including the target words and distracters. The test asked subjects to identify the words contained in the passage they read. The second one was the Meaning test (Test B) which simply asked students to give meanings of the target words either in Japanese or English. The last one was a fill-in-a-blank test (Test C), which asked students to fill in the parentheses of the sentences with the target words given in the margin. Test A was expected to measure each student's memori-

zation of orthographic information while test B and C were expected to measure their memorization of semantic information, and test C measured the ability to use the target words appropriately in given contexts.

3.4. Procedure

The present pilot study was conducted in "incidental learning design." The term "incidental learning", though it often misleads due to the difficulty maintaining a coherence of what is incidental, essentially indicates the study that subjects are not told that they will be tested after being given vocabulary instructions or exposed to stimuli (Hulstijn, 2001, pp.267-268). This study also took this position to define incidental learning.

The experiments were conducted during the regular class hours in which data were collected separately from each class. All five groups were asked to read the same passage to answer comprehension questions, and they answered these questions afterward without referring to the passage. Subjects retrieved the meanings of words in different fashions during the reading task. Group 1 was told to infer meanings of the non-words from contexts while reading and later given Japanese translations to check if their inferences were correct. Group 2 was told to translate the contents of the passage into Japanese by consulting a bilingual dictionary created for the purpose of the study for searching non-words in similar fashion to a normal bilingual dictionary. The dictionary gave several translation choices for each of the target words; selecting the best translation out of the lists could encourage the subjects to infer meanings, though it may not have been as hard as when Group 1 inferred meanings from the contexts. Group 3 was simply given the list of Japanese translations corresponding to every non-word in a one-by-one fashion while group 4 was given the list of English synonyms. Group 5 was a control group, which was given no meaning cues at all, and they were not told to guess words either. All the subjects were not told to memorize vocabulary nor informed that they would be given the vocabulary tests later.

After receiving the instructions, all the groups received regular lessons for 20 minutes to eliminate any information that may have stayed in their short-term memory. The amount of 20 minutes was used since Ebbinghaus House claims that 40% of information is lost during that length of time (in Takano, 1995, p.149). Then, suddenly a set of vocabulary tests (Test time 1) were administered without being announced in advance. A week later, the same set of tests (Test time 2) were administered again without advance notice in order to see how much attrition of vocabulary occurred for each group over a week to compare the retention levels of vocab-

ulary among the five groups. The data of these tests were analyzed statistically by a two-way repeated-measure ANOVA (5 groups×2 tests).

4. Results

Table 1 below summarizes the overall scores of Test time 1 and Test time 2 for 5 different groups.

Table 1:

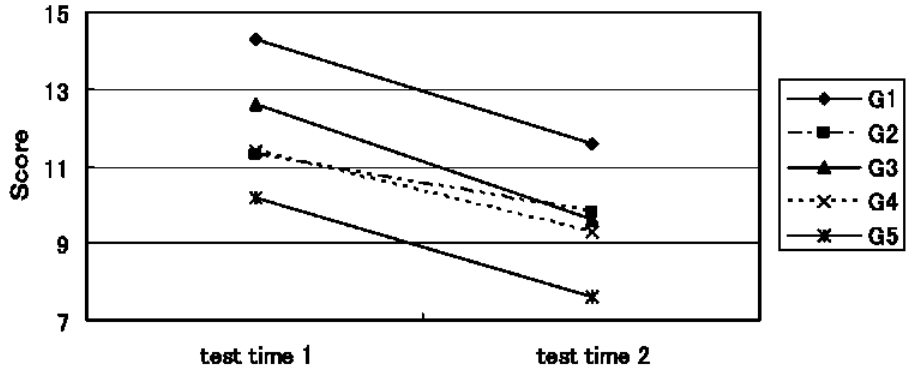
		Test time 1			Test time 2		
		Test A	Test B	Test C	Test A	Test B	Test C
G	N	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	M(SD)
G1	13	14.3 (0.9)	11.4 (2.6)	9.5 (3.4)	11.6 (1.7)	9.6 (4.0)	7.9 (3.9)
G2	13	11.3 (2.0)	7.5 (3.3)	7.1 (3.3)	9.8 (2.7)	6.5 (3.4)	5.9 (3.6)
G3	11	12.5 (2.4)	7.7 (3.4)	5.9 (3.1)	9.5 (2.9)	5.5 (3.3)	3.5 (3.9)
G4	11	11.4 (2.4)	4.6 (2.1)	3.7 (2.3)	9.3 (2.1)	3.2 (1.5)	1.9 (1.6)
G5	16	10.2 (2.7)	1.3 (1.9)	1.7 (2.1)	7.6 (3.7)	1.0 (1.3)	1.1 (1.9)

(Max. 16)

4.1. Test A (Word recognition test)

The results of the two-way ANOVA showed that the main effect of the groups was significant, F (1, 59) = 71.95, p<.001. A multiple comparison by Dunnett resulted in significant differences between G1 and G2, G4, G5 (p<.01) in Test time 1 and between G1 and G5 (p<.01) in Test time 2. The interaction effect between the groups and the tests was not statistically significant.

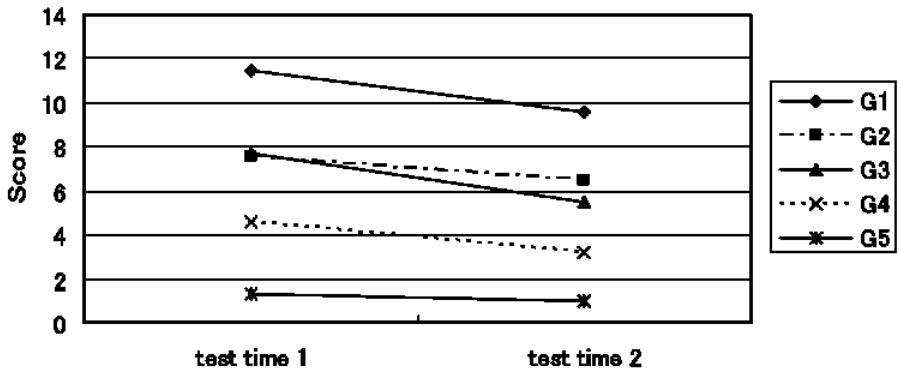
Figure 1 (Test A)



4.2. Test B (Meaning test)

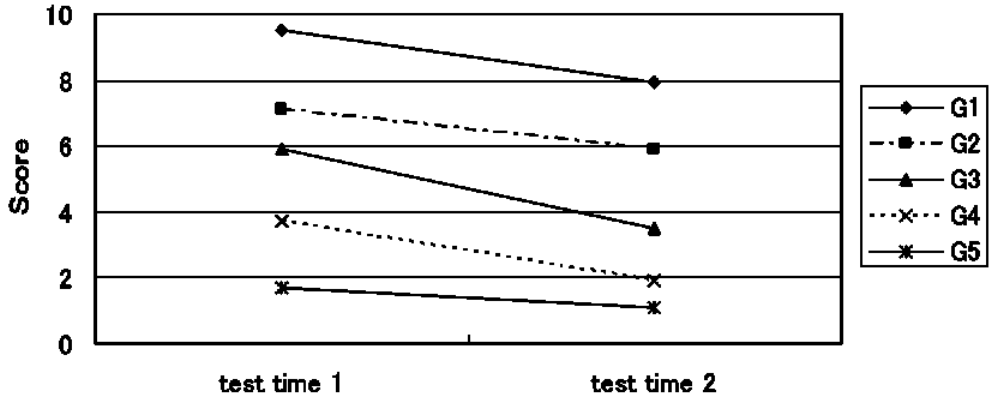
The results showed that the main effect of the groups was significant, $F(1, 59) = 25.59$, $p < .001$. A multiple comparison by LSD indicated that there were significant differences between G1 and each of G2, G3, G4, G5 ($p<.001$); between G2 and each of G4, G5 ($p<.01$); between G3 and each of G4, G5 ($p<.01$) in Test time 1. On Test time 2, each of G1 and G2 was significantly higher than G4 and G5 ($p>.05$), and G3 and G4 were higher than G5 ($p<.01$). The interaction effect between the group and tests was not significant, though it is close to the significant level ($p = .053$).

Figure 2 (Test B)



4.3. Test C (A-Fill-in-a-Blank test)

Figure 3 (Test C)



The main effect of the groups was siginificant, F (1, 59) = 14.14, p<.001. A multiple comparison by LSD resulted in significant differences (p<.05) between G1 and all the other groups, (p<.01) between G2 and G4, and between G2 and G5 respectively, and (p<.001) between G3 and G5 in Test time 1. For Test time 2, a multiple comparison by Bunnett showed that each of G 1 and G2 was significantly higher than each of G4 and G5 (p<.05). Again, there was no significant interaction between the tests and groups, though it is close to the significant level (p = .06).

4.4. Average in Total

Figure 4 (Average)

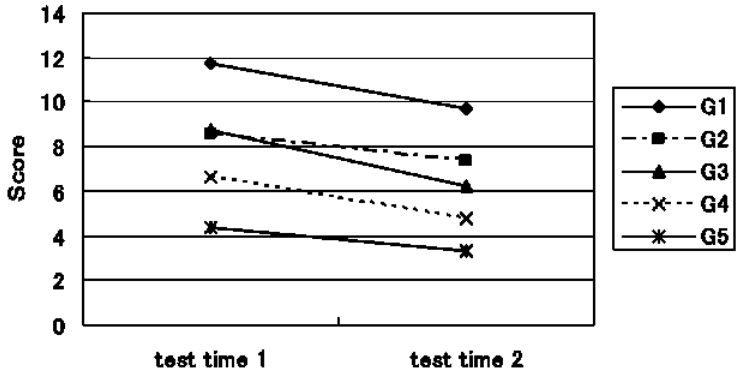


Table 2.

		Test time 1	Test time 2
		Average	Average
	N	M (SD)	M (SD)
G1	13	11.7 (2.0)	9.7 (2.9)
G2	13	8.6 (2.3)	7.4 (2.7)
G3	11	8.7 (2.6)	6.2 (2.9)
G4	11	6.6 (1.5)	4.8 (1.2)
G5	16	4.4 (1.9)	3.3 (1.8)

(Max. 16)

Figure 4 shows the mean differences of the 5 groups for the average scores of all types of the tests (A, B, and C) in Test time 1 and Test time 2. There was a significant interaction between the tests and groups, F (1, 59) = 3.09, p<.05. The main effect of the groups was also significant, F (1,59) = 20.0, p<.001. The multiple comparison by LSD showed that G1 was significantly higher than all the other groups (p<.001) in Test time 1 and 2. There were significant differences between G2 and each of G4, 5 (p<.05), between G3 and each of G4, G5 (p>.01), between G4 and 5 (p>.01) in the test 1. For Test time 2, G2 was higher significantly than G4,5 (p>.01), and G3 was higher significantly than G5 (p>.01). As seen in Figure 4, G2 almost kept the same score between Test time 1 and 2.

5. Discussion

The overall picture of the results suggest that it is possible to learn words in an incidental learning setting where students are not notified they are to be tested later or not told to memorize vocabulary; however, the number of words acquired incidentally by Group 5 was very small compared with those acquired by the other four groups. Group 5 was a control group that was not given any meaning cues; thus, the setting can be said to be the most incidental and implicit, which is quite similar to "extensive reading." This fact implies that unless students are given some meaning cues that may help them confirm the meanings of the words and encourage them to retrieve them, their vocabulary learning is not facilitated. To better acquire vocabulary incidentally, some types of meaning negotiation through retrieval of words should take place. It is possible to learn vocabulary incidentally in EFL settings but not effectively without being given some semantic cues.

Over a one-week duration after Test 1 was given, attrition of vocabulary took place for all the five groups. The results of Test A, B, and C showed no remarkable interaction effects between the groups and Test times, which means that no dramatic results were obtained out of this pilot phase; and in terms of the retention rate of the target words over a week, none of the groups was superior to the others. This can, however, provide pedagogical implications that in a regular reading class, incidentally learned words should be reinforced often by administering a vocabulary quiz soon after students encounter new words. The reinforcement should not be delayed until the final exam since attrition of words likely occurs soon after their learning.

The results of the average scores, however, showed the interaction effect between the groups and Test times, which indicated that the degree of the attrition of the words Group 2 and 5 learned was smaller than that of the other groups. Group 5 should not be taken into this account since the number of the acquired words by the control group was too small to explain how many words they forgot. In Test B and C, which were supposed to measure student memorization of semantic information of the non-words, the interaction effects found in Test B ($p = .053$) and in Test C($p = .06$) were almost approximate to the significant level ($p<.05$), and Figures 2 and 3 showed Group 2 experienced less attrition of the acquired words. Group 1, which was expected to require more mental effort in making inferences to retrieve meanings of the target non-words than Group 2, was initially predicted to be superior to the others for the retention of the words over a one-week duration, but G2 retained better than G1 over a one-week duration. This does not prove the hypothesis introduced earlier for this study. For further research, it is,

however, interesting to draw such an assumption that the reason for the better retention that Group 2 showed over a week is because the translation task somehow facilitated students making inferences and could provide enough confirmations in L1.

In Test B and the average in total on Test time 1, the data showed that the number of words Group 3 retained was larger than that of Group 4. The difference of the retrieving method can be pointed out as a factor that made this difference. Group 4 was given English synonyms and explanations in English for the target non-words while the other group was given the list of the Japanese translations. If these results of this pilot study are reliable and the subjects are considered to fall into a low proficiency group that still needs L1 translations to read English, what Group 4 showed is congruent to Watanabe's study (1997) that semantic information given in English can not facilitate vocabulary learning of low-proficiency students if it is in situation where students are expected to learn words incidentally while reading. This point will obviously need more careful analyses to draw a definite conclusion.

One question that confuses the evaluation regarding these results is that using English to understand meanings of vocabulary may require much more mental effort than using Japanese translations; thus, these results were somehow contradictory to the hypothesis of this pilot study emphasizing the "mental effort" for better retention of vocabulary. In this respect, the term "mental effort" is obviously failing to account for the characteristics of "depth of processing" and effective incidental learning of vocabulary since judgment on what conditions of learning can be clearly said to be involving "mental effort" is very subjective and hence the notion of "mental effort" somewhat loses explanatory power in that sense. One tentative explanation for these results may be drawn from the theory of "the limitation of working memory capacity" (Just & Carpenter, 1992), according to which low-proficiency students have difficulty retaining information while their attention resources being taken for a reading task. In this respect, whether the higher order processing involving "mental effort" turns out to be positive or negative has to do with working memory capacity. In this study, using Japanese translations may have required less effort than using English in retrieving meanings and have worked better to retain words for the students who still needed much attention just for reading English.

This explanation appears interesting since there was no significant difference between Group 2 and Group 3, both of which were given Japanese translations to retrieve meanings, in terms of the number of words to be retained in Tests B, C, and the average in total. That is, using Japanese translations to understand words was a shared factor between Group 2 and 3, which helped them to learn the words to the same degree as a temporary retention. The only

difference was that Group 2 retained knowledge of vocabulary better over a week than Group 3, which was probably owing to the higher degree of mental-effort involvement in the translation tasks than in simply looking at Japanese equivalents of the target words. Thus, this result shows that a difficult task may contribute to a strong memory trace.

In conclusion, throughout the data above, it is said that Group 1 could retain the higher number of words than the other groups, especially in the set of tests administered in 20 minutes after being exposed to those words. This implies that in terms of the number of words learned incidentally, making inferences with confirmations in L1 was effective. It would appear that retrieving meanings of words that involves some degree of "depth of processing" or "elaboration" could facilitate their acquisition even in incidental learning settings. However, the defects of the research design should be pointed out about the outstanding results Group 1 experienced.

The data in Test A showed that for the orthographic information of the target non-words Group 1 retained best. As the model of working memory (Osaka, 2002) suggests, if the visual encoding and semantic encoding work separately, the result of Test A suggests that Group 1 paid more attention to those target words than the other groups. The design of this pilot study failed to control for the possible variances in the degree and length of time to which attention was paid to each word. In the research design, however, the total amount of time to be spent on reading the passage was made approximately equal for all the groups, and Groups 2, 3, 4, and 5 were allowed to read the passage as many times as they wanted to in order to pay as much attention as Group 1 did on the target words. Although Craik & Lockhart (1972) pointed out that it is the depth of process, and not the length of time, that accounts for effective retention of information, it is questionable whether the outstanding results Group 1 showed overall in all the tests were solely due to "making inferences" which may have involved much mental effort. The result of Test A implies that the degree of explicitness interfered since Group 1 was told to infer meanings of the target words in advance; thus, they guessed the meanings of the words not without paying attention to reading the whole passage simultaneously. This means that the condition of the experiment for Group1 was not comparable to those of the others. The design failed to measure the main purpose of this study; however, the results of Group 1 imply that "explicit" learning of vocabulary with making inferences can be a strong driving force to facilitate students learning vocabulary, even if they are not told to memorize them beforehand.

6. Limitations and Conclusion

The pilot study carried out this time was planned to test the reliability and validity of the original research design and explore the possibility of whether the hypothesis proposed for this study could be testable or not. The following limitations were noticed and should be remedied for the next phase of the study. First of all, the sample size was too small to generalize the results obtained. Secondly, the validity of TOEIC-mini test was questionable since a variety of abilities may have influenced students' vocabulary retention. It is not clear whether the equivalence of English proficiency proven by TOEIC was appropriate for this study. Since it is difficult to carry out studies in random sampling designs for educational research, quasi-experimental design is often used, but in what way groups of samples are made homogeneous is one of the difficulty that a research design of this kind faces. Furthermore, it should be taken into account that affective factors such as motivation and attitude might have interfered with learning vocabulary during the experiment. For the next research, the defects pointed out at this pilot phase should be considered and remedied; especially effects of making an inference should be tested in an on-line fashion[5] to make it as implicit as possible for subjects to learn vocabulary while guessing its meaning. Finally, several points made in this pilot study are worth being checked more to support the views in cognitive psychology and to explore the possibility of incidental vocabulary learning which is strongly believed to provide great pedagogical implications to EFL students in Japan.

Notes:

1) This article is a revised and modified version of the contents presented by the author as a part of the symposium titled "Cognitive Mechanism of Reading in English" at the 47th Conference of the Kwansai English and American Literary Society of Japan, in Osaka.

2) Bottom-up processing stresses that reading comprehension proceeds sequentially from the isolated units in the lower level such as letters and words to sentences and paragraphs while top-down (higher order) processing emphasizes utilizing the higher levels of comprehension such as readers' background knowledge to draw predictions of textual meanings to be confirmed later with what is actually read.

3) Baddeley (1986) defines that working memory is "the temporary storage of information that is being processed in any of a range of cognitive tasks" (p.34). It consists of a central executive for a general memory capacity and several subsystems, the articulatory loop and the visual-spatial scratch pad,

which can allow one to complete more than two tasks at the same time. Complicated tasks such as reading require one to maintain information temporary while processing another task. The model of working memory emphasizes the availability of the limited working memory capacity, which accounts how smoothly one can process different tasks at the same time. There is an individual difference of working memory capacity, which has been considered to predict quality of one's reading performance (Just & Carpenter, 1992).

4) Long-term memory is a permanent store, usually explained as a declarative knowledge and procedural knowledge, from which one can draw information unconsciously. Short-term memory is for a temporary storage (15-30 seconds). By rehearsal at short-term memory, information can be transferred to long-term memory for its permanent storage.

5) Since the group 1 was told to guess meanings of the target words in advance, it would have been possible that the subjects carried out the word-guessing activity without paying much attention to reading contents of the passage. The task was explicitly carried out in an off-line fashion, which means that the words were not learnt as the by-product of reading the given passage. In order to examine if making inferences is effective for the retention of vocabulary which is incidentally learnt as the by-product of reading, subjects need to carry out words-guessing activity, simultaneously paying attention to comprehending what they are reading. The task needs to be carried out in an on-line fashion.

References

Akamatsu, N. (2000). Automaticity in second-language word recognition. *Journal of Joetsu University of Education, 19(2),* 785-795.

Baddeley, A. D. (1986). *Working memory.* New York: Oxford University Press.

Bensoussan, M., & Laufer, B. (1984). Lexical guessing in context in EFL reading comprehension. *Journal of Research in Reading, 7,* 15-32.

Coady, James. (1997). L2 vocabulary acquisition through extensive reading. In J. Coady & T. Huckin (Eds.), *Second Language Vocabulary Acquisition* (pp. 225-237). Cambridge: Cambridge University Press.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671-684.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology 104* , 268-684.

Craik, F. I. M., & Watkins, M. J. (1973). The role of rehearsal in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 11,* 671-684.

Daneman, M., & Carpenter, P. (1980). Individual differences in working memory and

reading. *Journal of verbal learning and verbal behavior, 19* , 450-466.

Daneman, M., & Merikle, P. (1996). Working memory and language comprehension: a meta-analysis. *Psychonomic Bulletin & Review, 3,* 422-433.

Day, R., Omura, C., & Hiramatsu, M. (1991). Incidental EFL vocabulary learning and reading. *Reading in a Foreign Language, 7,* 541-551.

Fraser, C. A. (1999). Lexical processing strategy use and vocabulary learning through reading. *Studies in Second Language Acquisition, 21,* 225-241.

Harrington, M.,& Sayer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition, 14.,* 25-38.

Hulstijin, J. H. (1992). Retention of inferred and given word meanings: experiments in incidental vocabulary learning. In P. J. L. Arnaud & H. Bejoint (Eds.), *Vocabulary and Applied Linguistics* (pp. 113-125). London: Macmillan.

Hulstijin, J. H. (2001). Intentional and incidental second-language vocabulary learning: a reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and Second Language Instruction,* (pp. 258-286.) Cambridge: Cambridge University Press.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychology Review, 99,* 122-149.

LaBerge, D., & Samuels, S. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6,* 273-323.

Laufer, B., & Hulstijin, J. (2001). Incidental vocabulary acquisition in a second language: the construct of task-based involvement. *Applied Linguistics, 22,* 1-26.

Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal, 24,* 237-270.

Nation, P. (1990). *Teaching and learning vocabulary.* New York: Newbury House.

Nation, P. (1994). *New Ways in Teaching Vocabulary.* Alexandria, Virginia: TESOL, Inc.

Osaka, M. (2002). *Working Memory: the sketchpad in the brain.* Tokyo: Shinyosha.

Perfetti, C. A. (1985). *Reading ability.* New York: Oxford University Press.

Pitts, M., White, H., & Krashen, S. (1989). Acquiring second language vocabulary through reading: a replica of the Clockwork Orange study using second language acquisition. *Reading in a Foreign Language , 5* , 271-275.

Qian, D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review, 56,* 283-307.

Redman, S., & Ellis, R. (1990). *A Way with Words Book 2.* Cambridge: Cambridge University Press.

Stanovich, K. E. (1992). Speculations on the causes and consequences of individual differences in early reading acquisition. In P. B. Gough, et al. (Eds.), *Reading Acquisition* (pp. 307-342). Hillsdale, NJ:

Lawrence Erlbaum Association.

Takano. Y. (1995). *Memory.* Tokyo: University of Tokyo Press.

Watanabe, Yuichi. (1997). Effects of single and multiple-choice glosses on incidental vocabulary learning. *JACET Bulletin, 28* , 177-191.

Zimmerman, C. B. (1997). Do reading and interactive vocabulary instruction make a difference?: an empirical study. *TESOL Quarterly, 31,* 121-140.

（しき・おさと　短期大学部講師）