—Article—

The Effect of Retrieval and Elaboration Practices on L2 Vocabulary Learning[1]

Chieko KAWAUCHI
Kurume University
Email: kawauchi_chieko@kurume-u.ac.jp

Abstract

This study focuses on the effects of two types of vocabulary practice, retrieval and elaboration, and compares these with learners' own ways of vocabulary learning. According to Frits, Acton, Voelkel, and Etkind (2006), retrieval practice simply involves retrieving a target word a few times within short intervals to consolidate the word's meaning. On the other hand, elaboration practice utilizes semantically related information, such as part of speech, base root, and alternative meanings, which are often used in class and considered to be more helpful than rote learning. Forty-nine low-intermediate students learned a total of 50 unfamiliar words either by retrieval, elaboration, or their own ways of learning, and their immediate and delayed effects on active recall and passive recall (Laufer & Goldstein, 2004) were investigated. At the end of the term, a questionnaire was given to note students' perceptions of these practices. Results of the immediate active recall (translation from L1 to L2) showed that vocabulary gain through retrieval was highest, significantly higher than that through elaboration, which was in turn significantly higher than learners' own ways of rote learning. However, results of passive recall (translation from L2 to L1) indicated no significant differences between the three types of practice either for immediate or delayed gains, suggesting that passive vocabulary could be learned effectively by whatever method. Results of the questionnaire supported the effectiveness of retrieval practice, showing that students felt retrieval to be easier and more effective, while also finding it more helpful for immediate recall than either elaboration or their own ways of learning. Some issues on elaboration-based practice are also discussed.

Introduction

One of the major challenges in learning English is to increase vocabulary, not only receptive but also productive vocabulary. Leech (1976, p. 203) states, "Whereas we have learned the grammatical rules of English in all essentials by the age of five, we continue the process of acquiring vocabulary and new uses of vocabulary right the way through our lives." It goes without saying that this is also true for all L2 learners. It is not, however, known how learners increase their vocabulary, particularly during EFL classroom activities.

A large scale survey by Schmitt (1997) reveals that Japanese learners of English most frequently use rote learning, such as writing or repeating words again and again, and consider

it most helpful when they try to learn a word's meaning. In fact, Tinkham (1989) demonstrated that Japanese learners of English who used rote learning scored significantly higher in their recall and recognition of new words in another language than did their American counterparts. Many previous studies have examined the effectiveness of rote learning by comparing it with the keyword method. For example, Ellis and Beaton (1993) indicated that the keyword method was effective for receptive vocabulary learning, but that repetition used in rote learning is a superior strategy for productive vocabulary learning. Focusing on productive vocabulary, Kawauchi (2011) found that rote learning was as good as the keyword method for immediate effect, but that the keyword method was significantly better for delayed effect. Thus, the effectiveness of rote learning is not universally acknowledged.

Recently, however, some other methods have been investigated in the psycholinguistic area, namely elaboration and retrieval (Frits et al., 2006). Frits et al. (2006) implemented elaboration by providing not only the target word's meaning but also related semantic information such as base root, part of speech, affixes, and useful expressions relating to the target word. They assumed that alternative and component meanings would provide a form of elaborate rehearsal and should be more helpful and interesting than rote learning. Elaboration is partly related to the study of "word-webs" by Aitchison (1994, p. 89) who states that "Word meaning is probably learned by noting the words which come alongside." Hashemi and Gowdasiaei (2005) also claim that if words in our mental lexicon are related in an associative network, then presenting items in lexical sets might facilitate the word-learning process. On the other hand, there is a counterargument that providing related word information might cause semantic interference (Erten & Tekin, 2008; Waring, 2008).

The study of retrieval is related to research into the most effective method for practice, such as massed practice, interval practice, and expanding practice. Retrieval involves retrieving target words a few times within a short time interval to consolidate a word's meaning. In one of their experiments, Frits et al. (2006) compared the effectiveness of three types of learning, retrieval, keyword method, and rote learning, using various foreign words from Turkish, Japanese, Hebrew, etc. They showed that retrieval and the keyword method substantially improved performance over rote learning, but there was no significant difference between them in either the immediate (three minutes later) or delayed (three days later) tests. In the other experiment, they added elaboration and combined retrieval and keyword method, and then examined not only receptive but also productive vocabulary knowledge. The results were revealing. Retrieval alone showed significantly higher results than the keyword method alone in recalling productive vocabulary. Combined retrieval and keyword method was not significantly different from either retrieval alone or keyword method alone. Elaboration alone was found to be the least effective for receptive and productive vocabulary both in the immediate and delayed tests. They concluded that retrieval practice could offer

an alternative technique for learning vocabulary and might be applied by a wider range of learners to a wider range of vocabulary.

Retrieval practice is, in fact, not known to be an activity used in language classes. In contrast, elaboration is frequently seen when a new word is introduced. In fact, many researchers support semantic associations of elaboration, saying that teaching new vocabulary in semantically related sets is an effective method of vocabulary instruction (Hashemi & Gowdasiaei, 2005; McCarthy, 1990).

In Frits et al. (2006), the effectiveness of elaboration and retrieval was investigated under a pure experimental condition on native speakers of English who were required to learn various foreign language lexical items, but it is not yet clear how these practices are feasible in regular EFL language classrooms like those found in Japan. This study addressed the following two research questions.

(1) How effective are retrieval and elaboration practices in gaining receptive and productive vocabulary in immediate and delayed recall, compared to learners' own ways of learning?
(2) How do Japanese EFL learners perceive the effectiveness of retrieval, elaboration and learners' own ways?

Method

Participants

Forty-nine low-intermediate students from two intact classes participated in this study. They were non-English-major freshmen who had been placed in the mid-level course as a result of our placement test. However, in the course of the 13-week longitudinal study, 11 students failed to complete all the necessary tests (a pre-test, immediate tests, and delayed tests) and practices (own ways, retrieval, and elaboration), so that the remaining 38 students were focused on here.

Target Vocabulary

A total of 50 unfamiliar words were used as target vocabulary selected from the textbook, *Prism Rose* (Kiggell, 2009), which was required for use in the semester-long classes. These words were divided into five sets of 10 words each. Each of these sets was studied in class as the lesson for the week, and most of them were listed as important words in that lesson of the textbook. The first set was assigned to learners' own way of studying (Own Way), and the remaining four sets were assigned to either retrieval (Retrieval) or elaboration (Elaboration), thus providing two sets for each practice. These four sets were counterbalanced in the two classes.

The target words for five test sets are shown in Table 1. The frequency level from JACET 8000 is also indicated in parentheses. When words were not included, (-) is used.

Table 1

Target Words and Frequency Levels

Set 1	beat (L1), inevitable (L4), extremely (L2), flame (L3), attention (L1), navel (-), publicity (L4), compliment (L6), criticism (L3), profile (L4)
Set 2	personalize (-), reflect (L3), distribute (L4), essential (L3), various (L1), resident (L4), individuality (L5), corporate (L4), display (L2), fee (L4)
Set 3	recommend (L3), pleased (L2), entirely (L2), recent (-), harm (L2), legendary (L7), absolutely (L2), appear (L1), director (L5), choose (L1)
Set 4	store as verb (L1), unveil (L6), compress (L7), lucrative (L7), functional (L4) share (L1), gadget (-), step (-), tune (L3), directly (L5)
Set 5	Breed (L3), soothing (L8), lovable (-), comforting (-), commonplace (L7), intelligent (L3), eyelash (-), smelly (-), grunt (L7), research (L1)

As shown above, frequency levels vary: 44% of them (22 words) belong to the Level 1 to Level 3, which students were supposed to have learned before entering college, while the remaining 56% of words were considered unknown or unfamiliar to them. Since the target words in each test set differ in frequency, it is necessary to take this into consideration, when the data are analyzed.

Procedure

In order to eliminate the possible effect of practice time, the amount of time spent for each practice was controlled. A total of 10 minutes including explanation and learning time were provided for each practice. First, students were asked to learn the first set of 10 words using their own ways without any instruction. They were given a worksheet listing 10 words along with Japanese definitions, one space for notes, and one space for the ways they learned. The teacher read out each word and its Japanese definition, and then told students to memorize it within 30 seconds. To confirm their learning, they were also told to write out the ways they memorized the word. They were not allowed to proceed to the next word until told to do so.

For Elaboration, another worksheet listing 10 words along with their Japanese definitions was provided like for Own Way, but this time three pieces of information were also added to each target word. The additional information was categorized into three types: (1) semantic information, such as other meanings, (2) morphological information, such as part of speech and base root, and (3) useful phrases relating to the target word, all of which were considered to be in the same semantic domain. These were read by the teacher, and students were told to read the definition and additional information repeatedly and remember the target word. Thirty seconds were provided for each word, and students were not allowed to proceed to the next word until told to do so.

For Retrieval, students received another worksheet of 10 words, but these words were divided into two parts, five words in each. The first five words appeared on the front page along with Japanese definitions, and 15 blanks were also provided at the bottom of the same page. Similarly, the remaining five words and 15 blanks were written on the opposite page. First, the teacher read the first target word and its meaning and asked students to memorize it within 20 seconds, as in Own Way. The second and third words were repeated in this way. Then the teacher read out the Japanese definition of the first target word and asked students to recall and write it down in each blank at the bottom of the worksheet. Students were asked to retrieve each word and write it three times after a group of three or four words were read out, thus filling out 15 blanks altogether. The same procedures were repeated for the remaining words. These procedures for Retrieval and Elaboration were adapted versions of Frits, et al. (2006). The worksheets used in these practices are shown in Appendix A.

At the end of the semester, a questionnaire was given that asked the students' perceptions on the three kinds of practice, Own Way, Elaboration, and Retrieval. A total of 15 questions were asked, using the Likert scale (1－5). All the questions for the questionnaire are displayed in Appendix B.

Pre-Test, Immediate Test, Delayed Test, and Scoring Criteria

Students' initial vocabulary knowledge of 50 target words was examined in the pre-test by using two types of tests adapted from Laufer and Goldstein (2004). The first type given in the first week was active recall (translation from Japanese into English) which was aimed to ascertain their productive vocabulary knowledge, and the other test given in the second week was passive recall (translation from English into Japanese) to confirm their receptive vocabulary knowledge. For active recall, the letters from the first half of each target word were provided as hints and to prevent learners from using other possible words. The results of the pre-test were used as a baseline for the study. Then, the three kinds of practice were given in the following 10 weeks. Detailed data collection is shown in Appendix C.

When each practice was explained, students were told that they would be tested after the practice. The immediate test was composed of active recall and passive recall with the same test format as the pre-test with one exception — when active recall was tested, the first half of each word was not provided as hints. Active recall was carried out first, followed by passive recall. Delayed tests with the same format were also carried out one week later with no prior announcement to see the effect of longer retention of the target words.

In scoring the active recall test, correct words were given 1 point, and answers with minor spelling mistakes such as "directer" for "director," or part of speech errors such as "legend" instead of "legendary," were given 0.5 point. In scoring the passive recall test, 1 point was given for the correct definition, and 0.5 point for answers with minor mistakes such as those involving part of speech errors.

Results

The effects of Own Way, Elaboration, and Retrieval on Immediate and Delayed Recall

Before comparing the three types of practice, it is necessary to show what kind of strategies students employed when they described their learning methods for Own Way. A total of 228 strategies were written out and categorized by the author, with a two-week separation. The rate of agreement was .95. Nearly 91%, or 197 strategies, were categorized as typical rote learning. The most frequently used strategy, which accounted for 71%, was spelling out the target word a number of times, sometimes with Japanese definitions. The other rote learning method involved pronouncing the target word many times or pronouncing it while writing. Therefore, it is fair to say that students tended to rely on rote learning when they were required to memorize words within a certain time frame. This finding also supported the survey by Schmitt (1997).

As shown in Table 1 earlier, the five test sets of target words vary in frequency level. Therefore, it is considered to be essential to examine whether or not these sets are equivalent, or how students responded to these words in the pre-test. Table 2 shows the results for passive recall (PR) and active recall (AR) in the pre-test.

Table 2

Scores for Pre-Test on Five Sets of Target Words

		Set 1	Set 2	Set 3	Set 4	Set 5
PR	Mean	3.29	2.96	4.28	2.58	3.16
	SD	1.44	1.87	1.92	1.24	1.91
AR	Mean	1.62	1.96	2.86	1.04	1.16
	SD	1.16	1.43	1.50	0.78	1.61

On the whole, passive recall ranges from 2 to 4 words out of 10, while active recall is much lower, from 1 to 2 words, implying that the target words are very difficult to produce. One-way ANOVA was carried out to see whether these five test sets are significantly different or not. The results yielded significant differences both in passive recall ($F(4, 185) = 5.25$, $p < .01$, $\eta^2 = .10$) and active recall ($F(4, 185) = 11.44$, $p < .01$, $\eta^2 = .20$). These findings strongly suggest that it is more appropriate to focus on vocabulary gain, rather than on the test scores above, so as to indicate how much students had learned the vocabulary through the three types of practice. To put it another way, the differences between the pre-test scores and the following immediate and delayed test scores can be interpreted as being indicative of learning resulting from each practice type.

The difference between the pre-test and the immediate test will be hereafter called "immediate gain," and the difference between the pre-test and the delayed test will be referred

as "delayed gain." Table 3 shows the descriptive statistics for all the immediate gains (Gain 1) and delayed gains (Gain 2) obtained from each practice. The results are based on 38 students who completed all the necessary tests and practices. Figure 1 also illustrates overall changes of vocabulary gains.

Table 3

Vocabulary Gains for Passive Recall and Active Recall

	Passive Recall (PR)			Active Recall (AR)		
	Own Way	Elaboration	Retrieval	Own Way	Elaboration	Retrieval
Gain 1 Mean	5.79	5.99	5.93	5.58	6.47	7.42
SD	1.52	1.49	1.58	2.22	1.62	1.17
Gain 2 Mean	4.11	4.73	4.62	1.21	3.24	3.51
SD	1.83	1.74	1.49	1.37	1.67	2.24

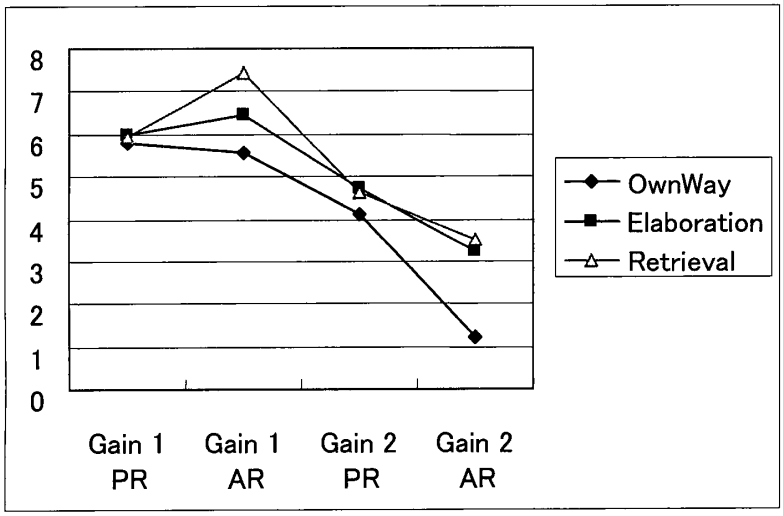

Figure 1. Overall results of vocabulary gains by three types of practice.

Compared to the results of the pre-test in Table 2, these results above show clear gains of target vocabulary through the three types of practice. To examine the effects of practice type, two-way ANOVA (repeated) was conducted, using vocabulary types (passive recall and active recall) and practice types (Own Way, Elaboration, and Retrieval) as independent factors. Following is an explanation of the results for immediate gains, as shown in Table 4.

Main effects were found in both practice types ($F(2, 74) = 9.75, p < .01, \eta^2 = .07$) and vocabulary types ($F(1, 37) = 9.17, p < .01, \eta^2 = .03$). There was, however, an interaction between them ($F(2, 74) = 10.84, p < .01, \eta^2 = .03$). The analysis for the interaction indicated that there was a main effect of practice types on active recall ($F(2, 74) = 19.36, p < .01$), indicating that the effects of practice types were significantly different in active recall.

No main effect of practice types on passive recall was found ($F(2, 74) = 1.01ns$). There was also a main effect of vocabulary types on Retrieval ($F(1, 37) = 36.82, p < .01$), showing that vocabulary gains for passive recall and active recall were significantly different in Retrieval (active > passive). (See Appendix D for detailed results).

Table 4

Immediate Gains of Vocabulary (AR/PR) by Three Types of Practice

Source	*SS*	*df*	*MS*	*F*	η^2
Subjects	252.1095	37	6.8138		
Practice types	52.7453	2	26.3727	9.75**	.07
s x Practice	200.2622	74	2.7062		
Vocabulary types	20.9967	1	20.9967	9.17**	.03
s x Vocabulary	84.7424	37	2.2903		
Practice x Vocabulary	24.6431	2	12.3215	10.84**	.03
s x Practice x Vocabulary	84.1228	74	1.1368		
Total	719.6220	227			

**$p < .01$

Finally, the Bonferroni *post hoc* comparisons revealed that there were significant differences in active recall: Own Way was significantly different from Elaboration and Retrieval, and Elaboration was also significantly different from Retrieval. Thus, it is fair to say that for active recall Retrieval was the most effective method, followed by Elaboration and Own Way (Retrieval > Elaboration > Own Way).

Table 5 shows the two-way ANOVA (repeated) results for delayed gains.

Table 5

Delayed Gains of Vocabulary (PR/AR) by Three Types of Practice

Source	*SS*	*df*	*MS*	*F*	η^2
Subjects	305.6109	37	8.2598		
Practice types	82.7659	2	41.3829	19.07**	.08
s x Practice	160.6189	74	2.1705		
Vocabulary types	185.8609	1	185.8609	93.55**	.18
s x Vocabulary	73.5080	37	1.9867		
Practice x Vocabulary	30.8391	2	15.4196	6.65**	.03
s x Practice x Vocabulary	171.5636	74	2.3184		
Total	1010.7672	227			

**$p < .01$

Like for the immediate gains, there were main effects of practice types ($F(2,74) = 19.07$, $p < .01$, $\eta^2 = .08$) and vocabulary types ($F(1, 37) = 93.55$, $p < .01$, $\eta^2 = .18$), but an interaction was also found ($F(2, 74) = 6.65$, $p < .01$, $\eta^2 = .03$). The analysis of interaction showed a main effect of practice types on active recall ($F(2, 74) = 24.00$, $p < .01$), but not on passive recall ($F(2, 74) = 1.55ns$), just like for immediate gains. There were also main effects of vocabulary types on Own Way ($F(1, 37) = 73.73$, $p < .01$), Elaboration ($F(1, 37) = 25.49$, $p < .01$) and Retrieval ($F(1, 37) = 8.02$, $p < .01$), indicating that passive recall is significantly different from active recall for all the practice types (passive > active). (See Appendix E for detailed results). It can be said that passive recall seems to be consistent whatever practice is used, although active recall greatly declines by the time of the delayed test.

Multiple comparisons by Bonferroni revealed significant differences in active recall: Own Way was significantly lower than both Elaboration and Retrieval, but no significant difference was shown between Elaboration and Retrieval. These results turned out to be Elaboration = Retrieval > Own Way.

In summary, the effects of three types of practice were significant on active recall, and Retrieval was the most effective for immediate gains. Elaboration was also effective, compared to Own Way, but less effective than Retrieval for immediate gains. However, for delayed gains Elaboration was as good as Retrieval, with Own Way being the least effective for active recall. Passive recall, on the other hand, showed no significant differences either for immediate or delayed gains. The implication is that passive recall is likely to be achieved whatever practice might be used.

Questionnaire: How Learners Perceived Each Practice

Students responded to a total of 15 questions regarding these three types of practices. (See Appendix B for the questions). They were asked to choose from the scale of 1 (I never think so) to 5 (I always think so). Table 6 displays ranking and mean scores of (1) how easy it was to use, (2) how effective it was to learn, (3) how enjoyable it was to do, (4) how well the word could be recalled immediately after each practice, and (5) how well the word could be retained one week later.

Students' perceptions clearly indicated that Retrieval was the most favored practice, showing the highest ranking and the highest mean scores for all the questions. Friedman tests were carried out to see if their perceptions differed among the three types of practice, and they revealed that all the questions yielded significant differences. This implies that students perceived that each practice was not equal in the degree of easiness, effectiveness, enjoyment, and recall in the immediate and delayed tests.

Table 6

Questionnaire Results for Ranking, Means, and Friedman Test Results

	(1) Easy	(2) Effective	(3) Enjoyable	(4) Immediate Recall	(5) Delayed Recall
Own Way	2.02	1.77	1.82	2.09	1.82
Mean (*SD*)	3.64 (0.91)	3.45 (0.97)	3.04 (1.09)	3.70 (1.02)	2.61 (0.97)
Elaboration	1.47	1.70	1.88	1.57	1.81
Mean (*SD*)	2.86 (0.80)	3.20 (0.97)	3.02 (0.98)	3.09 (1.05)	2.52 (0.87)
Retrieval	2.51	2.52	2.31	2.34	2.38
Mean (*SD*)	3.95 (0.98)	4.05 (0.86)	3.48 (1.04)	3.86 (1.09)	2.98 (0.92)
Friedman χ^2	30.04**	23.13**	10.05**	21.36**	15.86**

**$p < .01$

In particular, their responses concerning Retrieval were distinctively high for easiness, effectiveness, and immediate recall. In contrast, Elaboration practice was perceived as lowest in ranking, particularly on easiness and immediate recall. This finding partially supports the results of immediate gains for active recall, in which Elaboration was significantly lower than Retrieval. Although Elaboration produced significantly higher gains than Own Way, students' feedback on Elaboration was even lower than on Own Way. This finding might suggest that students' perceptions about Elaboration do not necessarily coincide with learning effect.

Discussion

The results for vocabulary gains show a clear effect of Retrieval on active recall, or productive vocabulary, both in short-term and long-term retention. Also, Elaboration was more effective than Own Way which was found to be the least effective. In contrast, passive recall, or receptive vocabulary, failed to show any significant differences between Retrieval, Elaboration, and Own Way, suggesting that receptive vocabulary could be learned to the same degree whatever practice might be used. This finding is in agreement with that of Kawauchi (2010), which showed that receptive vocabulary was learned equally in activities such as writing original sentences, answering matching questions, and learning vocabulary through CALL.

Considering the fact that active recall is more difficult to learn than passive recall, (Laufer & Goldstein, 2004; Sasao, 2008), the effect of Retrieval is worthy of note. So the question is: Why was Retrieval effective? One of the major differences at the time of Retrieval was that, while listening to Japanese definitions being orally recited, students had to retrieve and write the target words three times within short intervals. Thus, Retrieval might have required learners to be more conscious of the words and keep their attention more

focused than for Elaboration or Own Way. In fact, there were no specific chances of retrieving target words at the time of Elaboration and Own Way, although some students might have done this during their own rehearsal. Students' feedback also supported this kind of retrieval: Retrieval was most favored for easiness and effectiveness, and it was thought helpful for immediate recall. Another reason for its effectiveness may be due to its similarity to the test format used for active recall. One may argue that it seems to bear a slight resemblance to retrieving target words, and that this might have affected the immediate recalls.

On the other hand, Elaboration practice, which is often seen in the classroom, was not as good as Retrieval. The questionnaire results also indicated that Elaboration was perceived lowest for easiness and immediate recall, suggesting that this practice might not necessarily be a preferable way of learning for students. It seems likely that additional semantic-related information might not facilitate immediate recall to the same degree as Retrieval, but it might have helped in the end, since Elaboration scored as well as Retrieval in the delayed recall.

Another matter of concern with regard to Elaboration is semantic interference caused by cross-associations. Semantic interference has been discussed in various studies (Erten & Tekin, 2008; Finkbeiner & Nicol, 2003; Tinkham, 1997; Waring, 1997). These researchers claim that learning vocabulary in semantically related sets is not effective compared to learning vocabulary in semantically unrelated sets. Erten and Tekin (2008) and Waring (1997) demonstrated the semantic interference effect caused by cross-associations both during the encoding of information into memory and during the retrieval of information in translation. In fact, the present study found several minor mistakes concerning part of speech, particularly in Elaboration. On the other hand, Retrieval was based on retrieving target words a few times within short intervals, which does not include any words semantically related to one another. These findings confirmed the claim that vocabulary items seem to be organized in the mental lexicon around semantic bonds (McCarthy, 1990), but at the same time this is likely to hinder the learning of unfamiliar and/or new words at least in the immediate recall.

Conclusion

The present study investigated the effects of Retrieval, Elaboration, and learners' Own Way on L2 vocabulary learning. The findings showed clear effects of Retrieval and Elaboration on productive vocabulary acquisition both in the immediate and delayed recalls, compared to Own Way. In particular, Retrieval was found to be the most effective for learning productive vocabulary. This was also supported by the questionnaire in which Retrieval was perceived most favorably by students. On the other hand, Elaboration practice, in which additional semantic information was provided, was not as effective as Retrieval in immediate recall, and students' perceptions of this method were least favorable, particularly for easiness and immediate recall. The lower gains in immediate recall suggest that

semantic interference might have occurred by cross-associations in the semantic-related information provided in Elaboration.

For receptive vocabulary, however, practice types do not seem to cause a difference in learning effects. That is to say, receptive vocabulary can be learned effectively even by learners' own ways, which would bring about the same effect as Retrieval and Elaboration. In order to increase productive vocabulary, however, the findings of this study strongly suggest that learners' own ways of learning are not enough, and that some sort of specific activities, such as Retrieval, is necessary.

Some limitations of this study must be mentioned. First, due to some logistic reasons, this study was conducted in regular English classes, where a unified syllabus and textbook were required to be used. Therefore, it was not possible to select words totally unknown to students or to use words strictly on the basis of word frequency, word length, and word transparency, which are influential in learning (Laufer, 1997). Second, since the present study is based on a semester-long research, about 22% of the students failed to complete all of the necessary tasks because of lateness or absences. Future studies will need more test items and more students with different proficiency levels. Finally, although the delayed tests were given unannounced, some learners might have studied at home. This is one of the problems involved in classroom-based study and is difficult to control, but a more carefully arranged research design should be developed.

Oxford and Crookall (1990, p. 9) state that "learners are expected to pick up vocabulary on their own without any guidance." The current study shows that Retrieval could be part of a type of guidance that can be easily used in the regular classroom. In order to facilitate vocabulary learning, particularly that of productive vocabulary, it is about time that we develop some useful guidance methods and make them available in our daily classrooms, where English output is severely limited.

Note

1. This paper is a revised and enhanced version of the author's presentation at the 49th JACET Annual Convention held in Miyagi in 2010.

References

Aitchison, J. (1994). *Words in the mind.* Oxford: Blackwell.

Ellis, N., & Beaton, A. (1993). Factors affecting the learning of foreign language vocabulary: Imagery keyword mediators and phonological short-term memory. *The Quarterly Journal of Experimental psychology, 46(3)*, 533-558.

Erten, I. H., & Tekin, M. (2008). Effects on vocabulary acquisition of presenting new words in semantic sets versus semantically unrelated sets. *System, 36*, 407-422.

Finkbeiner, M., & Nicol, J. (2003). Semantic category effects in second language word

learning, *Applied Linguistics, 24*, 369-383.

Frits, C., Morris, R., Acton, M., Voelkel, A., & Etkind, R. (2006). Comparing and combining retrieval practice and the keyword mnemonic for foreign vocabulary learning. *Applied Cognitive Psychology, 21*, 499-526.

Hashemi, M. R., & Gowdasiaei, F. (2005). An attribute-treatment interaction study: Lexical-set versus semantically-unrelated vocabulary instruction. *RELC Journal, 36*, 341-361.

Kawauchi, C. (2010). The effects of CALL-alone and CALL-plus: Developing receptive and productive vocabulary. *Language Education & Technology, 47*, 75-92.

Kawauchi, C. (2011). Learning unknown words: The effects of dictionary, association, and own way learning strategies. *JACET Journal, 52,* 71-77.

Kiggle, T. (2009). *Prism rose*. Tokyo: Macmillan Language House.

Laufer, B. (1997). What's in a word that makes it hard or easy. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 140-155). Cambridge: Cambridge University Press.

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning, 54*, 399-436.

Leech, G. (1976). *Semantics*. Harmondsworth: Penguin Books.

McCarthy, M. (1990). *Vocabulary*. Oxford: Oxford University Press, Hong Kong.

Oxford, R., & Crookall, D. (1990). Vocabulary learning: A critical analysis of techniques. *TESL Canada Journal, 7*, 9-13.

Sasao, Y. (2008). Estimating vocabulary size: Does test format make a difference? *JACET Journal, 46*, 63-76.

Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 199-227). Cambridge: Cambridge University Press.

Tinkham, T. (1989). Rote learning, attitudes, and abilities: A comparison of Japanese and American students. *TESOL Quarterly, 23*, 695-698.

Tinkham, T. (1997). The effect of semantic clustering on the learning of second language vocabulary. *Second Language Research, 13*, 138-163.

Waring, R. (1997). The negative effects of learning words in semantic sets: A replication. *System, 25*, 261-274.

Appendix A

Worksheet for Practice

1. Own Way

今から、10 語を書いたリストを配布します。それぞれ、英語とその日本語の意味が書いてあります。各語彙に対して、30 秒時間を与えますので、どんな方法でもいいですので覚えてください。30 秒がたちましたら、合図をしますので、どのように覚えたか、右端の空欄に書いて下さい。メモは使っても使わなくても結構です。

注意：1 語ずつ、「はじめ」と「終わり」の合図をしますので、それがすむまで、先に進まないで下さい。

単語	日本語	メモ用	どのように覚えましたか
1. beat	～を破る		

~10 words in total

2. Elaboration

下の 10 語について、さまざまな説明を加えています。それぞれの単語を何度も発音し、説明文を読みながら単語と意味を結び付けましょう。それぞれ 30 秒間、何度も読んで覚えてください。メモは使っても使わなくても結構です。

注意：1 語ずつ、「はじめ」と「終わり」の合図をしますので、それがすむまで、先に進まないで下さい。

単語	意味	こんなことも知っておくと役に立つ	30 秒で覚えよう
1. recommend	勧める	「推薦する」と言う意味もあり、その後に名詞や that 節がきます。名詞形は recommendation となり、推薦状と言う意味もある。	メモ

~10 words in total

3. Retrieval

今日は、2 つのステップを使って覚えていきます。最初のステップは英単語とその意味を言います。20 秒でその意味を覚えてください。次のステップは、ときどき、前に出てきた単語を、ランダムに出しますので、それに対応する英単語を書きなさい（下の空欄に答えを書く）。

単語	日本語の意味	20 秒で覚えよう
1. gadget	機器	メモ

~5 words on page 1. The remaining 5 words are written on the opposite page.

先生が日本語の意味を言いますので、対応する英単語を書きなさい。

1	2	3

~15 blanks on page 1. The remaining 15 blanks are written on the opposite page.

Appendix B

Questionnaire

1.「自己流」で覚える（最初の週のみ実施）。
2. 単語についてのその他の意味など「説明文をよく読みながら」覚える。
3. 単語を覚えてすぐにその意味の「単語を書くクイズ」で覚える。

次の質問に下記の基準で答えてください。

５：とてもよくあてはまる　　４：よくあてはまる　　３：いくらかあてはまる
２：あまりあてはまらない　　１：まったくあてはまらない

A. 覚えやすさについて					
1 「自己流」は覚えやすい。	1	2	3	4	5
2 「説明文をよく読みながら」は覚えやすい	1	2	3	4	5
3「単語を書くクイズ」は覚えやすい	1	2	3	4	5
B. どの程度、効果的だと思うか					
4 「自己流」は効果的だと思う	1	2	3	4	5
5 「説明文をよく読みながら」は効果的だと思う	1	2	3	4	5
6 「単語を書くクイズ」は効果的だと思う	1	2	3	4	5
C. 覚えるときの楽しさについて					
7 「自己流」で覚えると楽しい	1	2	3	4	5
8「説明文をよく読みながら」覚えると楽しい	1	2	3	4	5
9「単語を書くクイズ」で覚えると楽しい	1	2	3	4	5
D. 直後に語彙を覚えているか					
10「自己流」で覚えると直後ではよく覚えている	1	2	3	4	5
11「説明文をよく読みながら」覚えると直後ではよく覚えている	1	2	3	4	5
12「単語を書くクイズ」で覚えると直後ではよく覚えている	1	2	3	4	5
E. 1週間後に語彙を覚えているか					
13「自己流」で覚えると１週間後でもよく覚えている	1	2	3	4	5
14「説明文をよく読みながら」覚えると１週間後でもよく覚えている	1	2	3	4	5
15「単語を書くクイズ」で覚えると１週間後でもよく覚えている	1	2	3	4	5

Appendix C

Procedures for Data Collection

Week	Class A	Class B
1	Pre-test : Active recall (AR)	Pre-test : Active recall (AR)
2	Pre-test : Passive recall (PR)	Pre-test : Passive recall (PR)
3	Own Way & Immediate AR/PR	Own Way & Immediate AR/PR
4	Delayed Own Way: AR/PR	Delayed Own Way: AR/PR
5	Elaboration 1& Immediate AR/PR	Retrieval 1 Immediate AR/PR
6	Delayed Elaboration 1: AR/PR	Delayed Retrieval 1: AR/PR
7	Retrieval 1& Immediate AR/PR	Elaboration 1 & Immediate AR/PR
8	Delayed Retrieval 1: AR/PR	Delayed Elaboration 1: AR/PR
9	Elaboration 2 & Immediate AR/PR	Retrieval 2 & Immediate AR/PR
10	Delayed Elaboration 2: AR/PR	Delayed Retrieval 2: AR/PR
11	Retrieval 2 & Immediate AR/PR	Elaboration 2 & Immediate AR/PR
12	Delayed Retrieval 2: AR/PR	Delayed Elaboration 2: AR/PR
13	Questionnaire	Questionnaire

Note. AR: Active recall, PR: Passive recall, 1: First practice, 2: Second practice.

Appendix D

Results for Analysis of Interaction in Immediate Gains

Source	*SS*	*df*	*MS*	*F*
A at B1:	3.9221	2	1.9611	1.01ns
(s x A at B1:	143.9529	74	1.9453)	
A at B2:	73.4662	2	36.7331	19.36**
(s x A at B2:	140.4322	74	1.8977)	
B at A1:	0.3289	1	0.3289	0.15ns
(s x B at A1:	78.9211	37	2.1330)	
B at A2:	4.5033	1	4.5033	3.41ns
(s x B at A2:	48.9342	37	1.3225)	
B at A3:	40.8076	1	40.8076	36.82**
(s x B at A3:	41.0100	37	1.1084)	

Note. A=Practice types (A1: Own Way, A2: Elaboration, A3: Retrieval).
B=Vocabulary types (B1: Passive recall, B2: Active recall). $^{**}p < .01$

Appendix E

Results for Analysis of Interaction in Delayed Gains

	SS	*df*	*MS*	*F*
A at B1:	7.0417	2	3.5206	1.55ns
(s x A at B1:	167.8950	74	2.2689)	
A at B2:	106.5636	2	53.2818	24.00**
(s x A at B2:	164.2875	74	2.2201)	
B at A1:	151.5613	1	151.5613	73.73**
(s x B at A1:	76.0541	37	2.0555)	
B at A2:	41.5732	1	41.5732	25.49**
(s x B at A2:	60.3359	37	1.6307)	
B at A3:	23.5656	1	23.5656	8.02**
(s x B at A3:	108.6816	37	2.9373)	

Note. A=Practice types (A1: Own Way, A2: Elaboration, A3: Retrieval).
B=Vocabulary types (B1: Passive recall, B2: Active recall). $^{**}p < .01$