



The Reading Matrix © 2011
Volume 11, Number 3, September 2011

Bridging the Gap between Receptive and Productive Vocabulary Size through Extensive Reading

Yuka (Maruyama) Yamamoto
Toyo Eiwa University

ABSTRACT

It is well established that extensive reading promotes the incidental learning of L1 and L2 receptive vocabulary; however, little is known about its effectiveness on productive gains in vocabulary knowledge. This paper investigates the extent to which extensive reading combined with writing tasks promotes productive vocabulary growth of Japanese university students. To examine changes in vocabulary size, three tests were administered: the Vocabulary Levels Test (Nation, 1990), the Productive Vocabulary Levels Test (Laufer & Nation, 1999), and the VocabProfile (Cobb, 2010). Contrary to expectations, no significant increase was shown in the amount of productive vocabulary size. However, the data suggested that one of the benefits of doing extensive reading with writing tasks is to retain receptive and productive vocabulary knowledge.

INTRODUCTION

Numerous studies on ESL/EFL research on incidental vocabulary learning through extensive reading have been conducted over the years (e.g., Elley & Mangubhai, 1981; Krashen, 1993; Laufer, 2009; Nagy, Herman, & Anderson, 1985; Saragi, Nation, & Meister, 1978). Recent studies (Grabe & Stoller, 1997; Nation, 2001, 2009; Waring & Nation, 2004; Waring & Takaki, 2003) have reported that large *receptive vocabulary* gains and long retention through *incidental vocabulary* acquisition cannot be expected as had once been thought. As a result, researchers have started to question the effectiveness of extensive reading on vocabulary acquisition (e.g., Nation, 2001; Waring & Nation, 2004; Waring & Takaki, 2003). However, very little is known about the effectiveness of extensive reading on the incidental learning of *productive vocabulary*, although it is often claimed (e.g., Nation, 2001, 2009; Nation & Ming-Tzu, 1999; Waring & Takaki, 2003) that graded readers play a significant role in helping students expand productive as well as receptive vocabulary. Waring and Takaki (2003) stated that “the benefits of reading a graded reader should not only be assessed by researching vocabulary gains and retention, but by looking at how graded readers help develop and enrich *already known* vocabulary” (p. 130, italics original). The purpose of this study is to find out whether there would be an increase in receptive and productive vocabulary as a result of incidental learning through extensive reading and productive activities.

Arguments against the Effectiveness of Extensive Reading

The importance of extensive reading for receptive vocabulary development has been widely acknowledged in the field of language teaching (e.g., Elley & Mangubhai, 1981; Grabe & Stoller, 1997; Krashen, 1993; Nagy, Herman, & Anderson, 1985; Nation, 2001, 2009; Saragi, Nation, & Meister, 1978). However, whether L2 incidental (implicit) vocabulary acquisition has the identical effectiveness to that of L1 incidental vocabulary learning is questionable. Scholars (e.g., Horst, Cobb, & Meara, 1998; Nation, 2001, 2009) have argued that extensive reading may not be the most efficient way for acquiring a large volume of receptive vocabulary within a limited time. Instead, they believe explicit vocabulary learning, in which attention is paid to linguistic forms, is more effective than implicit vocabulary learning, in which attention is paid to the meaning of the text and on developing reading fluency. For instance, if a learner reads for an hour, they will learn only about 3 to 6 words (Waring & Nation, 2004). Therefore, even if a student reads for 30 minutes every day for 13 weeks (one college semester), the student would read for 52.5 hours, but only potentially learn 157 to 315 words. Furthermore, Waring and Takaki (2003) found that after three months, nearly half of the words learned from reading graded readers had been lost, and thus, only a limited amount of *new* vocabulary had been retained.

They also showed that learners need to meet the word at least eight times to get a 50% chance of recognizing a word after three months; words that were met less than five times had a zero chance (Waring & Takaki, 2003). This indicated that there is a strong connection between word frequency of unknown words and vocabulary learning. Other studies also support that words that are repeated more frequently are learned better: In the Clockwork Orange Study, Saragi et al. (1978) concluded that in order for words to be acquired, a reader needs to encounter the same word at least 10 times, and in a replication of the Clockwork Orange Study, Horst et al. (1998) found that words that are repeated eight times or more were likely to be learned.

Other Aspects of Extensive Reading

If we were only to take a look at the size of the newly acquired receptive vocabulary, it would be hard to expect a dramatic increase within a limited time and without any large scale exposure to new vocabulary. Previous studies suggest that researchers may have overestimated the power of incidental vocabulary acquisition (e.g., Waring & Takaki, 2003; Zahar, Cobb, & Spada, 2001). However, the major studies discussed so far paid attention only to acquiring *new* receptive vocabulary. Several researchers have stated that teaching new words is only one aspect of teaching vocabulary (Horst et al., 1998; Nation, 2001; Nation & Ming-Tzu, 1999). Instead, they have suggested that extensive reading serves as a means of deepening the knowledge of previously met vocabulary. Nation (2001) calls extensive reading a “gradual process of one meeting with a word adding to or strengthening the small amounts of knowledge gained from previous meetings” (p. 155). By encountering words already acquired in their reading, learners can deepen their understanding of them. Elley and Mangubhai (1981) demonstrated that learners surpass a vocabulary threshold, which then allows them to turn receptive vocabulary into vocabulary they can produce.

Laufer (2003) compared the effects of reading alone to productive word-focused tasks such as writing original sentences using target words, and completing sentences using target words on overall vocabulary gains. The results indicated that there was a greater increase in

vocabulary size through word-focused tasks than through reading alone, and that the vocabulary items learned through productive word-focused tasks were retained longer than learning from reading, even with using the dictionary. In her earlier work on new vocabulary words, Laufer commented that, “If not ‘pushed’ to use them, they may never be activated and therefore remain in passive vocabulary only” (1998, p. 267). Nation (1990) also said that “if productive learning is important, then the development of the quality of learning ... is important. Intensive practice in using vocabulary in speech and/or writing is therefore a useful activity” (p. 5).

Purpose of the Study

As outlined above, researchers are questioning the effectiveness of extensive reading to gain large amounts of new receptive vocabulary. Since reading is a receptive skill, it is natural to assume that students are likely to develop receptive vocabulary rather than productive vocabulary. However, not much research has been done to find out if extensive reading is responsible for increases in productive vocabulary. Thus, the current study investigates the extent to which graded readers promote incidental receptive and productive vocabulary learning in combination with *pushed-output* activities (Swain, 1985). Laufer’s study (1998) distinguished vocabulary knowledge into three types: (1) *receptive* knowledge, knowledge to understand the most frequent meaning, (2) *controlled* productive knowledge, knowledge to recall words using hints, and (3) *free* productive knowledge, knowledge to produce words through activities such as writing assignments. Therefore, this study addresses two research questions:

1. To what degree does extensive reading promote the incidental learning of receptive, controlled productive, and free productive gains in vocabulary knowledge?
2. To what extent do pushed-output tasks enhance students’ controlled productive and free productive vocabulary knowledge?

METHOD

Participants

This study involved 67 students from one co-educational class at a private university in Tokyo, Japan. They were first-year students majoring in Social Welfare and Tourism, enrolled in an English for Academic Purposes reading program. In the study, there were two groups: 34 students were in the Reading Group (control group) and 33 students were in the Extensive Reading Group (experimental group). Both groups of this study were placed in lower-intermediate classes based upon a departmental English-placement test (conducted at the beginning of the semester, taken by 4,344 students from all the departments), which consisted of listening and reading items. The test was considered a valid and reliable measure with a Cronbach alpha of .84. The mean scores out of 50 points ranged from 26.5 (Department of Science) to 32.4 (Department of Law). The mean scores for the total number of students who belonged to the Department of Social Welfare ($N = 313$) and Tourism ($N = 357$) were 27.8 and 31.8, respectively, which indicated that their overall L2 proficiency was lower-intermediate.

Procedure

In both the Reading Group and the Extensive Reading Group, over 80% of the class time was equally spent on intensive reading and explicit vocabulary instruction. The Extensive Reading Group students were also required to, in addition to the regular coursework, read at least five books that they chose from a collection of graded readers. The students were instructed to read for at least 30 minutes every day outside of class, and were asked to read one book within two weeks—a total of five books—for the entire 13-week semester.

When selecting the appropriate level of a book, one strong indicator is the students' current vocabulary size (Nation, 2001, 2009). Thus, on the first day of class students took a receptive-vocabulary test, and, based upon the results, they were told to select their books from the Oxford Bookworms series (which mainly consists of condensed versions of classic novels) which has six levels (Level 1 [400 headwords] to Level 6 [2,500 headwords]). According to Nation (2009), the total word families are roughly the same as the graded readers' level in the Oxford Bookworms series. A level which is slightly lower than the present vocabulary knowledge, “*i*-1,” was recommended. On the second day of class, the instructor checked whether the students bought the right level of books. Then, to check whether students had read at least 30 minutes of their graded readers, they were asked to submit a reading log by filling in the title of the book, the total number of pages they read, and the total time they spent reading.

To ensure that students would be doing ‘incidental’ learning rather than ‘intentional’ learning, as the degree of attention may vary based on the types of tasks assigned, the students were not informed that they would be tested on vocabulary afterwards. In other words, though they were aware they were participating in a research study, they did not know that their vocabulary would be checked via a written book response to see how well they had understood the text. Hulstijn (2003) distinguishes between incidental and intentional learning as whether or not learners are told in advance that they would later be tested. For instance, if learners are given instructions that they would be tested on new words, they will consciously try to store information when they encounter new words. As a result, this would become intentional learning rather than incidental learning. Therefore, students were engaged in extensive learning to understand the text (meaning), although the intention of the study focused on learning something else (vocabulary).

After students finished reading a graded reader outside of class (i.e., about one graded reader every two weeks), they wrote a short book response (about 100 words in length) in English in order to “process words elaborately and repeatedly” (Hulstijn, 2001, p. 272). This pushed output enhanced their productive vocabulary knowledge. Each book response was written in class and included a short summary of the book as well as a reaction to the content of the book.

Research Instruments

Students' vocabulary levels were checked by three unannounced vocabulary-level tests which were administered twice, once at the beginning and once at the end of the semester. In other words, the three tests were given twice. The first test was the Vocabulary Levels Test (Nation, 1990, 2001; Schmitt, 2000; Schmitt, Schmitt, & Clapham, 2001), which measured students' receptive vocabulary knowledge. There are five sections in this test (the 2,000-word level, the 3,000-word level, the university-word level, the 5,000-word level, and the 10,000-word level), but only three sections (the 2,000-, the 3,000-, and the 5,000-word levels) were used in the present study. The university-word level contains specialized vocabulary of university texts and

was therefore excluded from the test. The 10,000-word level was not used because it has been shown that average Japanese university students only know approximately 3,000 to 4,000 words (Mochizuki & Aizawa, 2000).

The other two tests, which were used to assess their productive vocabulary growth, were the Productive Vocabulary Levels Test (Laufer & Nation, 1999) and the English Web VocabProfile (Cobb, 2010). The former test was used to examine the learners' controlled-productive vocabulary ability, whereas the latter was used to assess their free-productive vocabulary ability. The Productive Vocabulary Levels Test has four versions (from A to D). In this study, version A was used for the pre-test and version B was used for the post-test. Like the receptive-vocabulary test, this test is divided into five levels, from the 2,000-word level up to the 10,000-word level. Again, only three levels (the 2,000-, 3,000-, and 5,000-word levels) were used.

The three tests were chosen based on Laufer's study (1998), which measured the sizes of receptive, controlled-productive, and free-productive vocabularies. Each section was comprised of six words and three definitions, and in each section the test takers were asked to match the words on the left with the definitions given on the right. For example:

| | |
|----------|-----------------------------|
| original | |
| private | <u>1</u> first |
| royal | <u>2</u> not public |
| slow | <u>6</u> all added together |
| sorry | |
| total | |

For each level, there were 18 items. Another type of question provided the first several letters of the target word within a sentence, and students were asked to complete the word. For example:

I'm glad we had this opp____ to talk. [Answer: *opportunity*]

To measure free-productive vocabulary, an Internet application called VocabProfile (Cobb, 2010) was used. In order to analyze the results, students were asked to read in class the first chapter of the course textbook *Reading & Vocabulary Development 3: Cause & Effect* (published by Heinle & Heinle) on world population growth, and then write a reaction paper in English based on their readings. The same passage was used in the post-test. It was assumed that the 13-week gap between the pre-test and the post-test was sufficient to preclude distortion or invalidation of the VocabProfile results.

To assess their productive-vocabulary size, excluding other factors, students were not allowed to refer to anything, such as the reading passage or a dictionary. Students were told to write a summary of the passage and a reaction to the issues presented in the passage. All their reaction papers were typed into Microsoft Word by the researcher. Then, the VocabProfile calculated and divided the words of each composition into four categories according to their frequency: (a) the most frequent 1,000 English words, (b) the second-most frequent thousand English words, (c) academic English words, and (d) the remainder, or less frequent words; that is, words that are not in any of the above lists. Because the VocabProfile processes all of the words in a student's reaction paper, the total percentage of all the four categories adds up to 100 percent.

Results

The total pre-test scores (Table 1) of the Extensive Reading Group and the Reading Group showed no significant differences between the two groups ($t = 1.63, p = .11$) indicating that they were both about the same vocabulary level before the treatment. Consequently, differences on the post-tests cannot be attributable to differences on the level of two groups.

After the administration of each pre-test and the post-test, the Reading Group ($n = 34$) and the Extensive Reading Group ($n = 33$) were compared on the three vocabulary scores. To investigate the students' vocabulary size growth at the beginning and end of the semester, paired t -tests were analyzed.

Reliability

The reliability was checked on all major instruments used in this study using the Spearman-Brown prophecy formula (split-half method). In the study, the overall reliability of the receptive vocabulary test was .86 and for the controlled productive vocabulary test was .88. The overall reliability of the post-tests was also calculated, which resulted in a fairly high reliability figure of .83. For the controlled-productive vocabulary test, the result was .84, also relatively high. Since the free-productive vocabulary test "does not consist of discrete items" (Laufer, 1998, p. 261), it was impossible to analyze the internal consistency.

Receptive Vocabulary Size

The data was not normally distributed for the receptive vocabulary level test (e.g., 2,000-word level, skewness = -1.81, kurtosis = 4.22, Table 1) and for the controlled-productive vocabulary level test (e.g., 5,000-word level, skewness = 1.35, kurtosis = 2.18; Table 2). However, in order to compare the present test results to those conducted by Laufer (1998), t -tests were used throughout the present study.

For the receptive vocabulary test, the maximum score for each frequency level is 30 and the total maximum score is 90. The differences between the pre-test and post-test were evaluated by a paired-sample t -test. However, multiple tests (in this case three for 2,000-, 3,000-, and 5,000-word levels) using the same group of participants increases the possibility of finding a significant difference between groups by mere chance alone (which is called a Type I error). In order to avoid such an error, a Bonferroni adjustment was made, and the traditional p value of .05 was set to .0167 (.05 divided by 3, the total number of multiple tests) (Tabachnick & Fidell, 2001). The results are shown in Table 2.

Overall, the results indicated that the total number of receptive vocabulary words on average was fairly similar between the pre- and post-tests. Means of the Extensive Reading Group were 64.64 on the pre-test and 65.42 in the post-test. Means of the Reading Group were 60.15 in the pre-test and 61.97 in the post-test. Analyzing this data using the t -test showed that indeed no significant difference was found between the pre-test and the post-test for either group (Extensive Reading Group, $t = .75, p = .46$, two-tailed; Reading Group, $t = 1.00, p = .32$, two-tailed).

Table 1. Pre-Test Descriptive Statistics by Vocabulary Type ($N = 67$)

| | Extensive Reading Group ($n=33$) | | | | | Reading Group ($n=34$) | | | |
|-----------------------|------------------------------------|----------|-----------|----------|----------|--------------------------|-----------|----------|----------|
| | | <i>M</i> | <i>SD</i> | Skewness | Kurtosis | <i>M</i> | <i>SD</i> | Skewness | Kurtosis |
| Receptive | 2,000 | 26.85 | 1.91 | -0.55 | -0.23 | 24.18 | 5.07 | -1.81 | 4.22 |
| | 3,000 | 19.76 | 4.73 | -0.04 | -0.87 | 19.5 | 4.65 | -0.23 | -0.85 |
| | 5,000 | 18.03 | 4.11 | -0.13 | -1.05 | 16.47 | 5.61 | -0.03 | -0.88 |
| | Total | 64.64 | 8.98 | -0.04 | -0.68 | 60.15 | 13.14 | -0.79 | 0.44 |
| Controlled Productive | 2,000 | 13.06 | 1.77 | -0.27 | 1.05 | 12.21 | 2.56 | -0.41 | -0.55 |
| | 3,000 | 8.4 | 2.55 | 0.77 | 1.52 | 7.18 | 2.55 | -0.31 | -0.36 |
| | 5,000 | 3.73 | 2.36 | 0.35 | -0.63 | 2.68 | 2.24 | 1.35 | 2.18 |
| | Total | 25.18 | 4.93 | 0.43 | 0.63 | 22.06 | 5.21 | -0.51 | -0.41 |
| Free Productive | K1 | 88.56 | 4.62 | | | 88.55 | 3.9 | | |
| | K2 | 3.62 | 1.54 | | | 1.04 | 2.64 | | |
| | AW | 4.42 | 3.16 | | | 3.67 | 2.01 | | |
| | Other | 3.43 | 2.76 | | | 3.71 | 2.36 | | |

Note: K1 = the first 1,000 words, K2 = the second 1,000 words, ns = not significant, AW = academic

Table 2. Comparison of Receptive Vocabulary Size ($N = 67$)

| | Extensive Reading Group ($n = 33$) | | | | | | Reading Group ($n = 34$) | | | | | |
|-------|--------------------------------------|-----------|-----------|-----------|------------|-----------------|----------------------------|-----------|-----------|-----------|------------|-----------------|
| | Pre-Test | | Post-Test | | Difference | | Pre-Test | | Post-Test | | Difference | |
| | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | Mean Diff. | <i>t</i> -value | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | Mean Diff. | <i>t</i> -value |
| 2,000 | 26.85 | 1.91 | 26.33 | 1.88 | -.52 | -1.27 | 24.18 | 5.07 | 25.97 | 2.32 | 1.79 | 2.38 |
| 3,000 | 19.76 | 4.73 | 21.90 | 3.36 | 2.15 | 3.05* | 19.50 | 4.65 | 21.68 | 3.87 | 2.18 | 3.00* |
| 5,000 | 18.03 | 4.11 | 17.18 | 4.40 | -.85 | -1.35 | 16.47 | 5.61 | 14.32 | 5.36 | -2.15 | -2.35 |
| Total | 64.64 | 8.98 | 65.42 | 7.88 | .79 | .75 | 60.15 | 13.14 | 61.97 | 9.29 | 1.82 | 1.00 |

Note: * $p < .0167$, two-tailed.

Looking at each group individually, the Extensive Reading Group showed no significant change in the receptive vocabulary size overall, except at the 3,000-word level, showing a significant growth (Extensive Reading Group, $t = 3.05$, $p = .005$, two-tailed; Reading Group, $t = 3.00$, $p = .005$, two-tailed). The results could be interpreted that students in both groups had room for improvement at 3,000-word levels. The mean score for the Extensive Reading Group was 21.90 and for the Reading Group was 21.68 out of 30, a mastery level of about 70%.

Another thing to note is the decline at the 5,000-word level in both groups. Even though both groups did not show any statistically significant loss, this may not be surprising as the students were exposed to minimal input at the 5,000-level words in class. Alberding (2007), who is familiar with the textbook used by the participants, analyzed the second chapter of the course textbook. Table 3 summarizes how many word forms and word families are allotted to the 1,000, 2,000, Academic Word List (AWL), and Off-List word levels by using a computer software program, *Range* (Nation & Heatley, 1994).

Table 3. Textbook and Graded Reader's Level Analysis

| | | K1 Words | K2 Words | AW | Off-List Words |
|----------------|---------|----------|----------|-------|----------------|
| Textbook | | 76.70% | 8.96% | 4.18% | 10.15% |
| Graded Readers | Level 5 | 25.0% | 44.7% | 11.3% | 19.0% |
| | Level 6 | 22.1% | 35.2% | 21.2% | 21.4% |

Note: K1 = the first 1,000 words, K2 = the second 1,000 words, AW = Academic Words

For analysis, a chapter from Unit 2 was selected randomly from the textbook. Based on the chapter analysis, students were exposed mostly to the first 2,000 words in class, which accounted for 85.66% of the total words they encountered. Only 14.33% of words were above the 2,000-word level. Furthermore, Alberding (2007) found that only 26 words appeared more than three times (11.5%) in the reading passage, and only 18 words appeared four or more times (8%) in the textbook. Yet, it is surprising that both groups showed a significant improvement at the 3,000-word level. Because the computer program only provides information on the K1, K2, AWL, and Off-List words, it is not clear which level beyond the 2,000-word level belongs to the Off-List words. However, based on the results, it can be assumed that most of the words probably belonged to the 3,000-word level, not the 5,000-word level or beyond. In addition, the graded reader's level analysis at Level 6 (Table 3; Claridge, 2005) suggests that students were almost equally exposed to each level, although the proportion beyond the 2,000-word level (i.e., 3,000- and 5,000-word level) is unclear.

Using the formula created by Laufer (1998) to estimate the number of word families known by a learner, word family estimations were calculated based upon the total raw scores of receptive and controlled productive tests. First, even though students were tested on three levels (2,000, 3,000, and 5,000 levels), the 1,000 and 4,000 levels were accounted for statistically. For the 1,000 level, students were assumed to have the same score as for the 2,000 level. For the 4,000 level, the average score for the 3,000 and 5,000 words was taken. Then, the total score was added and multiplied by 5,000 (as students were tested up to a size of 5,000 word families) and divided by 150 (30 items per level for 5 levels—1,000, 2,000, 3,000, 4,000, and 5,000). For example, if a student scored 20 out of 30 at the 2,000-word level, 15 out of 30 at the 3,000-word level, and 5 out of 30 at the 5,000-word level, it is estimated that a student has 20 out of 30 at the 1,000-word level and 10 out of 30 at the 4,000-word level. The total score adds up to 70. This figure is multiplied by 5,000 and divided by 150 which roughly equals to 2,333 word families.

The total mean scores of the Extensive Reading Group's pre-test were 64.64 (about 3,680 word families) and 65.42 (about 3,709 word families) in the post-test (Tables 2 and 4). The total mean scores of the Reading Group were 60.15 (about 3,411 word families) and 61.97 (about 3,531 word families), respectively (Tables 2 and 4). Comparing this figure to typical Japanese university students whose abilities range from 3,138 (Saida, 2006) to 3,769 English words (Mochizuki & Aizawa, 2000), the data suggests that the present students were well within the range of average intermediate university students.

Table 4. Growth of Receptive Vocabulary Size ($N = 67$)

| | Extensive Reading Group ($n = 33$) | | | | Reading Group ($n = 34$) | | | |
|----------|------------------------------------------------------|--|-----------|--|--------------------------------------------|--|-----------|--|
| | Pre-Test | | Post-Test | | Pre-Test | | Post-Test | |
| | <i>Estimate in Word Families</i> | | | | | | | |
| Total | 3,680 | | 3,709 | | 3,411 | | 3,531 | |
| % Change | | | +1% | | | | +4% | |

Overall, students did retain their lexical knowledge after the 13-week treatment, and both groups even showed a tendency for *growth* in their lexical knowledge. There was a 1% increase in the Extensive Reading Group and 4% in the Reading Group.

Controlled Productive Vocabulary Size

For the controlled-productive vocabulary test, the maximum score for each frequency level was 18 and the total maximum score was 54. Both groups showed a declining trend at every level except at the 3,000-word level (Table 5). A Bonferroni adjustment was made and the traditional p value of .01 was set to .0034 (.01 divided by 3, 2,000, 3,000, and 5,000-word levels) (Tabachnick & Fidell, 2001).

Table 5. Comparison of Controlled-Productive Vocabulary Size ($N = 67$)

| | Extensive Reading Group ($n = 33$) | | | | | | Reading Group ($n = 34$) | | | | | |
|-------|------------------------------------------------------|-----------|-----------|-----------|------------|-----------------|--------------------------------------------|-----------|-----------|-----------|------------|-----------------|
| | Pre-Test | | Post-Test | | Difference | | Pre-Test | | Post-Test | | Difference | |
| | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | Mean Diff. | <i>t</i> -value | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | Mean Diff. | <i>t</i> -value |
| 2,000 | 13.06 | 1.77 | 10.45 | 2.55 | -2.61 | -5.03** | 12.21 | 2.56 | 9.00 | 2.85 | -3.21 | -5.93** |
| 3,000 | 8.39 | 2.55 | 9.73 | 2.60 | 1.33 | 3.31** | 7.18 | 2.55 | 7.68 | 2.99 | 0.5 | 1.02 |
| 5,000 | 3.73 | 2.36 | 3.67 | 2.29 | -.06 | -.17 | 2.68 | 2.24 | 1.97 | 2.18 | -0.71 | -1.49 |
| Total | 25.18 | 4.93 | 23.85 | 5.85 | -1.33 | -1.72 | 22.06 | 5.21 | 18.65 | 6.11 | -3.41 | -3.62** |

Note: ** $p < .0034$, two-tailed

However, attrition was significant only in the Reading Group at the 2,000-level ($t = -5.93$, $p = .000001$, two-tailed) as well as the total ($t = -3.62$, $p = .0010$, two-tailed). In other words, while the Extensive Reading Group was able to retain their lexical knowledge over time, there was a significant loss in the Reading Group.

As for the 2,000-word level, the paired t -tests showed a significant difference in both groups (Extensive Reading Group, $t = -5.03$, $p = .00002$, two-tailed; Reading Group, $t = -5.93$, $p = .000001$, two-tailed), but in a direction that was contrary to expectations. The possibility for the significant decrease might be that, as we have already seen in the course textbook analysis (Table 3), only 8.96% of the vocabulary belonged to the K2 words, showing that it was enough to cause the growth of receptive-vocabulary size in the Reading Group, but not enough to make it productive at that level. Because of the lack of exposure and less frequent use of vocabulary words at the 2,000-word level, it could have led to attrition in both groups. As for the 3,000-word level, the result indicated a similar trend as the receptive vocabulary level test. The mean scores of both groups showed improvement; however, this time only the Extensive Reading Group's

score was significantly different ($t = 3.31, p = .002$, two-tailed). The fact that the difference was found only in the Extensive Reading Group could be the result of additional output-oriented activities, such as book responses.

As for the 5,000-word level, although the results showed no statistically significant difference between the pre-test and post-test, the mean scores did show a declining trend, perhaps due to the lack of vocabulary exposure at that level. Another possibility is a “floor effect,” indicating the difficulty of the test. Students scored less than four points out of 18 from the pre-test which made it nearly impossible to score even lower in the post-test.

The Extensive Reading Group on pre- and post-tests decreased from 2,461- to 2,278-word families, which is an average loss of 183 word families (Table 6). The Reading Group had an even more drastic vocabulary loss of 374 word families—from 2,178 to 1,804. The attrition rate was 7% in the Extensive Reading Group and 17% in the Reading Group. The result revealed a substantial difference between the pre- and post-tests only for the Reading Group, due to vocabulary attrition.

Table 6. Loss of Controlled-Productive Vocabulary Size ($N = 67$)

| | Extensive Reading Group ($n = 33$) | | | | Reading Group ($n = 34$) | | | |
|----------|------------------------------------------------------|--|-----------|--|--------------------------------------------|--|-----------|--|
| | Pre-Test | | Post-Test | | Pre-Test | | Post-Test | |
| | <i>Estimate in Word Families</i> | | | | | | | |
| Total | 2,461 | | 2,278 | | 2,178 | | 1,804 | |
| % Change | | | -7% | | | | -17% | |

Free-Productive Vocabulary Size

Student reactions and summaries to textbooks were analyzed through the VocabProfile program to examine free-productive vocabulary growth. The result of the VocabProfile is given in percentages of the words in each frequency level (Table 7). Again, a Bonferroni adjustment was made and the traditional p value of .05 was set to .0125 (.05 divided by 4, K1, K2, AWL-, Off-List-levels, Tabachnick & Fidell, 2001).

Table 7. Summary of Free Productive Vocabulary Size Change Beyond 2K (in Percentage)

| | Extensive Reading Group ($n = 33$) | | | | | | | Reading Group ($n = 34$) | | | | | | |
|-----------|------------------------------------------------------|------|-----------|------|------------|------------|----------|--------------------------------------------|-----------|------|------------|------------|--|--|
| | Pre-Test | | Post-Test | | Mean Diff. | t -value | Pre-Test | | Post-Test | | Mean Diff. | t -value | | |
| | M | SD | M | SD | | | M | SD | M | SD | | | | |
| K1 | 88.54 | 4.62 | 86.38 | 3.13 | -2.16 | -2.64 | 88.58 | 3.90 | 86.61 | 3.68 | -1.97 | -2.55 | | |
| K2 | 3.62 | 1.54 | 4.83 | 1.68 | 1.21 | 2.89* | 4.04 | 2.64 | 5.13 | 2.04 | 1.09 | 1.98 | | |
| AW | 4.42 | 3.16 | 3.88 | 1.53 | -.53 | -.98 | 3.67 | 2.01 | 4.04 | 1.40 | 0.37 | .97 | | |
| Off | 3.43 | 2.76 | 4.91 | 2.16 | 1.48 | 2.51 | 3.71 | 2.36 | 4.22 | 2.27 | 0.51 | 1.00 | | |
| Beyond 2K | 7.85 | 4.66 | 8.80 | 2.60 | .95 | 1.17 | 7.38 | 2.67 | 8.26 | 2.70 | .89 | 1.54 | | |

Note: Mean values represent percentages. Percentages have been rounded to the nearest second decimal place, and thus may not add up to 100. * $p < .0125$, two-tailed. K1 = the first most frequent 1,000 words; K2 = the second most frequent 1,000 words; AW = the academic words of English from the Academic Word List; Off = Off-List words, words that are not in any of the above lists, Beyond 2K = AW + Off-List.

Laufer and Nation (1999) suggest that the higher the percentage of infrequent words a learner knows, the larger the learner's free-productive vocabulary is. Thus, the VocabProfile ratios beyond the 2,000-word level, the academic words, and the Off-List frequency were selected for analysis. While significant change was observed in the controlled-productive vocabulary size, the free-productive vocabulary size ratios did not change significantly, though mean percentages increased slightly (Table 7).

Because of the characteristics of the free-productive vocabulary test, it was not possible to estimate the productive vocabulary size, but instead the proportions beyond 2,000-word levels were identified. The mean differences showed that the vocabulary growth in the Extensive Reading Group (.95% growth) exceeded that of the Reading Group (.89%), but this difference is not statistically significant.

Ratio between Receptive and Controlled Productive Vocabulary

Using the same analysis undertaken in Laufer's study (1998), the ratio between receptive and controlled productive vocabulary was calculated for each group. The gap between the receptive- and controlled-productive vocabulary for the Reading Group became much wider from 63.9% on the pre-test to 51.1% on the post-test (a 12.8% gap) compared with the Extensive Reading Group which went from 66.9% to 61.4% (a 5.5% difference) after the 13-week treatment (Table 8).

As for the free-productive vocabulary test, the VocabProfile program compares only the proportion of frequent to less frequent words, as opposed to the actual number of word families. Thus, it was not possible to establish a statistical comparison between the VocabProfile and first two vocabulary tests.

Table 8. Comparison of Receptive-Controlled Productive Ratios

| | Extensive Reading Group | | | Reading Group | | |
|-----------|----------------------------------|---------|-------|----------------------------------|---------|-------|
| | R Test | CP Test | Ratio | R Test | CP Test | Ratio |
| | <i>Estimate in word families</i> | | | <i>Estimate in word families</i> | | |
| Pre-Test | 3,680 | 2,461 | 66.9% | 3,411 | 2,178 | 63.9% |
| Post-Test | 3,709 | 2,278 | 61.4% | 3,531 | 1,804 | 51.1% |

Note: R Test = Receptive vocabulary level test, CP Test = Controlled productive vocabulary level test

Correlations

To examine how the three types of vocabulary knowledge relate to each other, correlations were computed among receptive, controlled-productive and free-productive vocabulary tests (beyond the 2,000-word level, Table 9). The results show that there was a correlation between receptive vocabulary size and controlled-productive vocabulary knowledge for both the pre- and post-tests. However, no correlation was found between these two measurements and free-productive vocabulary. This result is in line with Laufer (1998), who found that "learners who could recognize more words than other learners and produce them if forced to, were not necessarily those who would use more infrequent vocabulary in free expression" (p. 265).

Table 9. Inter-Correlations between the Receptive and Productive-Vocabulary Tests Scores

| | Controlled-Productive | | Free-Productive (Beyond 2K) | |
|-----------------------|-----------------------|-----------|-----------------------------|-----------|
| | Pre-Test | Post-Test | Pre-Test | Post-Test |
| Receptive | .70** | .62** | .08 | .10 |
| Controlled-Productive | | | .20 | .19 |

Note: ** $p < .01$, two-tailed. 2K = the second most frequent 1,000 words

DISCUSSION

Summary of Main Results

The aim of this study was to investigate the extent to which incidental vocabulary learning combined with pushed-output activities promote receptive and productive vocabulary knowledge growth over a one-semester EAP (English for Academic Purposes) course. Although the current study lasted only one semester, the results clearly indicate that the combination of incidental vocabulary learning and productive tasks serve as a means to prevent lexical attrition for productive vocabulary. The findings of the study support the assumption that “learners do not learn a lot of *new* words from graded reading, but in fact graded reading helps to deepen and consolidate *already known* language” (Waring & Takaki, 2003, p. 154). This is probably the reason why, even though both groups retained their receptive vocabulary knowledge during the 13-week treatment, only the members of the Extensive Reading Group were able to retain their controlled-productive vocabulary knowledge over time. The Reading Group exhibited a significant attrition rate of 17%, or an estimated loss of 374 word families. This supports the findings (Hansen & Reetz-Kurashige, 1999; Weltens & Grendel, 1993) that productive-vocabulary knowledge decays much faster than receptive-vocabulary knowledge.

The study could also be interpreted to say that as students encountered words that they already knew receptively by reading graded readers, the connection between receptive and productive vocabulary knowledge was made much stronger, especially at the 3,000-word level. This seems reasonable as students were likely to be encountering words mostly at this level based on the analysis of the textbook and graded readers (Alberding, 2007; Claridge, 2005).

As for free-productive vocabulary, no significant statistical difference was found beyond the 2,000-word level in either group over the treatment period. Even though both productive tests measured the same productive vocabulary knowledge, a different trend was observed between groups. Although the Reading Group showed a significant attrition on the controlled productive test, the proportion of frequent and less frequent vocabulary use had not changed in their free writing. This may be explained by looking at the correlations among the tests. For instance, a correlation was found between the receptive vocabulary- and controlled-vocabulary knowledge, but there was no correlation between either of these two measures and free-productive vocabulary. This is in line with Laufer’s (1998) study indicating that “even though their passive and controlled active vocabularies have improved, they use the same proportion of frequent and non-frequent vocabulary in a free writing activity” (p. 263).

The lack of correlation may be explained by the notion of avoidance, in which a learner refuses to use unknown or uncertain words (Melka, 1997). In this study, students may have consciously avoided using words about which they were uncertain. Such uncertainty may be due to a lack of confidence regarding correct spelling, and correct grammatical and word usage. Another possibility is that the topic of the writing assignment, which was world population growth, provided only a narrow opportunity for lexical use. Even if the students tried to use all the acquired vocabulary, not all words would be particularly relevant to the topic.

The analyses of the pre-test data also suggested that the receptive-vocabulary size (Extensive Reading Group, 3,680-word families, Reading Group, 3,411-word families) was more than 1.4 times larger than that of the controlled-productive vocabulary size (Extensive Reading Group, 2,461-word families, Reading Group, 2,178-word families). This finding corresponds to earlier research (e.g., Laufer, 1998; Nation, 2001) that productive vocabulary is generally smaller than receptive vocabulary size. The gap became even wider after the 13-week treatment, showing a 40% increase in the Reading Group, which was double the increase of the Extensive Reading Group. This supports the point of view that receptive-vocabulary size develops much faster in comparison to the gradual gain of productive knowledge; the gap between receptive and productive vocabulary takes time to be reduced (Melka, 1997).

To sum up, analyses of the test data indicates that the combination of incidental vocabulary learning with output-oriented tasks serves as a means to retain productive vocabulary knowledge. This supports the view that by encountering receptive words already acquired in their reading, learners can deepen their understanding of the use of vocabulary knowledge (e.g., Nation, 2001; Nation & Ming-Tzu, 1999; Waring & Takaki, 2003). Encountering words repeatedly via reading graded readers might have led to the confirmation of the meaning and function of the words that were already stored in their memory system which may have, potentially, developed into productive-vocabulary knowledge.

Furthermore, with low-frequency words below the 2,000-word level being less encountered, it is possible that words beyond the 5,000-word level decayed much faster compared with other levels. In addition to the low proportion of less frequent words used in the course textbook, the reason for vocabulary loss in the 5,000-word level may have been due to the fact that with increased reading ability, more reading needs to be done in order to encounter less frequent, advanced, unknown words (Waring & Takaki, 2003). However, as the textbook analysis by Alberding (2007) suggests, only 14.33% of the words were above the 2,000-word level, and so there was very little exposure to words beyond the 2,000-word level. This lends further support for the use of graded readers at the university setting. Nation (2001) states, “the larger the vocabulary size, the greater the quantity of language that needs to be processed in order to meet the words to be learned again” (p. 67).

Pedagogic Implications

Despite the limitations of the study, several important implications can be drawn with respect to Japanese-English education in university settings. The results from the pre-test receptive-vocabulary level test showed that students have a receptive vocabulary size of 3,411- to 3,680-word families, indicating that they were typical intermediate-level Japanese university students. Based on Laufer’s (1992) findings, a threshold of 3,000-word families (4,800 lexical items) is enough for *minimal* comprehension of reading. In this regard, students of this study had a sufficient amount of vocabulary size to understand the reading texts. However, this is still far

from the satisfactory level for advanced learning. For instance, Hazenberg and Hulstijn (1996) concluded that in order to comprehend first-year university reading materials, L2 learners need to know a minimum of 10,000-word families. Even more daunting, Okamoto (2007) showed, without any additional treatments, students' vocabulary size declined rapidly after one semester.

In order to retain lexical knowledge and to further develop additional learning, several suggestions can be made. First, as the current study showed, extensive reading practice indeed helped students retain their receptive, controlled, and free-productive vocabulary knowledge. For effective incidental learning to occur, selecting an appropriate level of reading text for individual students is crucial. For instance, the course textbook analysis (Alberding, 2007) used in the present study implied that the textbooks may be lexically too easy for first-year Japanese university students, leading to attrition in productive vocabulary knowledge. This would seem to suggest that encouraging students to do extensive reading, particularly reading graded readers, is a solution to this problem.

Second, along with incidental vocabulary learning, to make the connection between receptive and productive vocabulary stronger, the current study implies that some sort of conscious effort needs to be made on the part of learners. It can take form through multiple exposures to words through practicing words in context and isolation as well as output-oriented tasks.

Limitations and Suggestions for Future Studies

Since the current study only looked at the changes in receptive- and productive-vocabulary knowledge over a very limited amount of time (one 13-week semester), it would be worthwhile to carry out the study for a longer period, especially with regard to productive vocabulary knowledge. First, it takes time for receptive vocabulary to become productive vocabulary, so gains in productive vocabulary are likely to occur over a longer period of training. Second, it is important to see how much vocabulary students can actually retain over time. Waring and Takaki (2003) state, "new words will be fresh in the mind for an immediate post test, thus these scores will be higher than if the test were taken some time later" (p. 150). They found that even if a word was met more than 18 times, it was retained only 10% to 15% of the time. Future studies will need to test how words are better retained.

Second, it could be argued that reading five graded readers was not enough to even be considered as extensive reading; that is, not enough input for incidental vocabulary acquisition to take place. Yet, a statistical difference was found between the Extensive Reading Group and the Reading Group. This suggests that a larger amount of reading may lead to a greater difference between the two groups and better retention.

Third, due to the fact that the study was limited to the breadth of vocabulary knowledge, (i.e., vocabulary size), it would be interesting to determine what type of receptive words students encountered in their reading which actually led to new productive word use in their writing. This study did acquire such data, and analysis is forthcoming. Future research using qualitative analytical methods may also provide greater insight regarding the relationships among receptive, controlled-productive, and free-productive vocabulary knowledge.

The biggest limitation to this study was that through the process of writing book responses, it is questionable as to how much pushed output actually took place. One of the critical components of the output hypothesis is to give learners the opportunity to receive feedback and to push towards "the delivery of message that is not only conveyed, but that is

conveyed precisely, coherently, and appropriately” (p. 249). However, in this study, it is hard to know to what extent each student “noticed the gap” between what they wanted to write and what they were able to write. Even if they were able to notice the gap, it is questionable as to how many of them actually went back to review their writing tasks after receiving written feedback from the instructor.

CONCLUSION

While there has been a long debate over whether implicit (incidental) or explicit learning is more effective, the advantage of doing extensive reading should not only be seen in terms of gaining a massive receptive-vocabulary size. Instead, one of the major benefits of reading graded readers is to retain productive as well as receptive vocabulary knowledge. The current study implies that extensive reading practice might help students confirm the meaning and function of the words that are already stored in their memory systems making the connection stronger, which in turn may potentially develop into productive-vocabulary knowledge. The analyses of the test data indicated that extensive reading practice alone may not be very helpful toward acquiring productive-vocabulary knowledge. However, conscious effort (which leads to noticing) needs to be made on the part of learners by, for example, pushed output practice (Swain, 1985), and “processing words elaborately and repeatedly” (Hulstijn, 2001, p. 272).

Thus, in terms of practical pedagogy, it is important to design a curriculum that nurtures the retention of lexical knowledge and further develops additional learning. Extensive reading does indeed seem to be effective in retaining receptive- and productive-vocabulary knowledge. For better retention, creating tasks which focus on different aspects of word knowledge will help learners deepen the level of processing. Since this was a quantitative study with a focus on the breadth of vocabulary knowledge (i.e., vocabulary size of participants), it is worth investigating what is happening in students’ mental lexicon while they are engaged in extensive reading. Future research involving qualitative analysis may provide greater insights regarding the process of how incidental vocabulary learning leads to acquisition, and how receptive-vocabulary knowledge turns into productive-vocabulary knowledge.

Acknowledgements

I would like to express my sincere gratitude to all the participants in this study for their cooperation. I would like to especially thank Dr. Yoshinori Watanabe, Professor Kensaku Yoshida, and Dr. Mitsuyo Sakamoto for their insightful comments and advice regarding the original version of this paper. My appreciation is also extended to Ron Martin for his assistance in editing the work. I am also grateful to Professor John I. Liontas and the anonymous reviewers for their helpful comments and suggestions on the previous version of the paper.

Note

This paper was first presented at the International Applied Linguistics Association (AILA), 2008 (August), Germany.

Yuka (Maruyama) Yamamoto is a Lecturer at Toyo Eiwa University, Japan, and is a Ph.D. candidate in Applied Linguistics at Sophia University, Tokyo, Japan, where she earned an M.A. in Applied Linguistics. Her research interests include extensive reading, lexical studies, and EAP curriculum development.

Email: y-yama@toyoeiwa.ac.jp

REFERENCES

- Alberding, M. (2007, June). *Using software to create learning materials for R & L*. Paper presented at Rikkyo University ELP Faculty Development Seminar, Tokyo, Japan.
- Claridge, G. (2005). Simplification in graded readers: Measuring the authenticity of graded readers. *Reading in a Foreign Language*, 17(2), 144-158.
- Cobb, T. (2010). Web Vocabprofile (Version 2.5) [Computer software and manual], an adaptation of Nation & Heatley's (1994) *Range*. Retrieved November 23, 2010, from <http://www.lex tutor.ca/vp>
- Elley, W. B., & Mangubhai, F. (1981). *The impact of a book flood in Fiji primary schools*. Wellington: New Zealand Council for Educational Research.
- Grabe, W., & Stoller, F. (1997). Reading and vocabulary development in a second language: A case study. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 98-122). NY: Cambridge University Press.
- Hansen, L., & Reetz-Kurashige, A. (1999). Investigating second language attrition: An introduction. In L. Hansen (Ed.), *Second language attrition in Japanese contexts* (pp. 3-18). NY: Oxford University Press.
- Hazenbergs, S., & Hulstijn, J. (1996). Defining a minimal receptive second language vocabulary for non-native university students: An empirical investigation. *Applied Linguistics*, 17, 145-163.
- Horst, M., Cobb, T., & Meara, P. (1998). Beyond a clockwork orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language*, 11(2), 207-223.
- Hulstijn, 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, UK: Cambridge University Press.
- Hulstijn, J. H. (2003). Incidental and intentional learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 349-381). Malden, MA: Blackwell Publishing.
- Krashen, S. (1993). *The power of reading: Insights from the research*. Englewood, CO: Libraries Unlimited.
- Laufer, B. (1998). The development of passive and active vocabulary in a second language: Same or different? *Applied Linguistics*, 19, 255-271.
- Laufer, B. (2003). Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence. *Canadian Modern Language Review*, 59(4), 567-587.
- Laufer, B. (2009). Second language vocabulary acquisition from language input and from form-focused activities. *Language Teaching*, 42(3), 341-354.

- Laufer, B., & Nation, I. S. P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing, 16*, 33-51.
- Melka, F. (1997). Receptive vs. productive aspects of vocabulary. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 84-102). Cambridge, UK: Cambridge University Press.
- Mochizuki, M., & Aizawa, K. (2000). An affix acquisition order for EFL learners: An exploratory study. *System, 28*, 291-304.
- Nagy, W. E., Herman, P. A., & Anderson, R. C. (winter, 1985). Learning words from context. *Reading Research Quarterly, 20*, 233-253.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Boston, MA: Heinle & Heinle.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge, UK: Cambridge University Press.
- Nation, I. S. P. (2009). *Teaching ESL/EFL reading and writing*. NY: Routledge.
- Nation, I. S. P., & Heatley, A. (1994). *Range: A program for the analysis of vocabulary in texts* [computer software]. Available at <http://victoria.ac.nz/lals/staff/paul-nation/nation.aspx>
- Nation, I. S. P., & Ming-Tzu, K. W. (1999). Graded readers and vocabulary. *Reading in a Foreign Language, 12*, 355-80.
- Okamoto, M. (2007). Lexical attrition in Japanese university students: A case study. *JACET Journal, 44*, 71-84.
- Saida, C. (2006). Eigo hyoka [English assessment]. In H. Fukuda (Ed.), *Sougou eigo program: zengaku dounyu to aratanaru chosen* [Ibaraki University integrated English program: Introducing the program to all departments and a new challenge] (Tech. Rep., Vol. 2). Ibaraki, Japan: Ibaraki University.
- Saragi, T., Nation, I. S. P., & Meister, A. (1978). Vocabulary learning and reading. *System, 6*, 70-78.
- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge, NY: Cambridge University Press.
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring that behaviour of two new versions of the vocabulary levels test. *Language Testing, 18*, 55-88.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*. (4th ed.). NY: Allyn and Bacon.
- Waring, R., & Nation, I. S. P. (2004). Second language reading and incidental vocabulary learning. *Angles on the English-Speaking World, 4*, 11-23.
- Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language, 15*(2), 130-163.
- Weltens, B., & Grendel, M. (1993). Attrition of vocabulary knowledge. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 135-156). Amsterdam: John Benjamins Publishing Company.
- Zahar, R., T. Cobb & N. Spada (2001). Acquiring vocabulary through reading: Effects of frequency and contextual richness. *The Canadian Modern Language Review, 57*(4), 541-572.