

2015

Second Language Vocabulary Learning Through Extensive Reading With Audio Support: How Do Frequency and Distribution of Occurrence Affect Learning?

Stuart Webb

The University of Western Ontario, swebb27@uwo.ca

Anna C-S Chang

Hsing-Wu University

Follow this and additional works at: <https://ir.lib.uwo.ca/edupub>



Part of the [Education Commons](#)

Citation of this paper:

Webb, S.A. & Chang, A, C-S. (2015). Second language vocabulary learning through extensive reading: How does frequency and distribution of occurrence affect learning? *Language Teaching Research*, 19(6), 667-686.

Second language vocabulary learning through extensive reading with audio support: How do frequency and distribution of occurrence affect learning?

Stuart Webb

Western University, Canada

Anna C-S Chang

Hsing-Wu University, Taiwan

Abstract

This study investigated (1) the extent of vocabulary learning through reading and listening to 10 graded readers, and (2) the relationship between vocabulary gain and the frequency and distribution of occurrence of 100 target words in the graded readers. The experimental design expanded on earlier studies that have typically examined incidental vocabulary learning from individual texts. Sixty-one Taiwanese participants studied English as a foreign language (EFL) in an extensive reading program or in a more traditional approach structured around a global English course book. A pretest, posttest, and delayed posttest were administered to all participants. The results indicated that vocabulary gains through reading and listening to multiple texts were high. Relative gains were 44.06% after reading the 10 graded readers and 36.66% three months later. The relationships between vocabulary learning and frequency and distribution of occurrence were found to be non-significant, indicating that frequency was perhaps one of many factors that affected learning.

Keywords

Distribution of occurrence, extensive reading, frequency of occurrence, vocabulary learning

I Introduction

Extensive reading has the potential to develop knowledge of unknown words (Day & Bamford, 1998; Grabe & Stoller, 2002). It also provides learners with opportunities to consolidate knowledge of partially known words each time they are encountered in different contexts (Nation, 2001; Nation & Webb, 2011). In first language (L1) learning, it is generally agreed that vocabulary knowledge can be acquired incidentally through reading abundantly. The more texts that are read, the more opportunities there are to repeatedly encounter the same patterns of letters, words, and collocations. Although there is a considerable amount of research that has examined the degree to which vocabulary is incidentally learned through reading, most studies have examined learning through reading a single text. Gains made through reading a single text may not be consistent with those made through reading multiple texts (Horst, 2005; Schmitt, 2010), nor may reading a single text engage second language (L2) learners in the same manner as an ecologically valid extensive reading situation. The present study aims to expand on the existing research by looking at vocabulary learning through long-term extensive reading. It also looks at the role that the frequency and distribution of occurrence of words in 10 graded readers

play in vocabulary learning through reading while listening.

1 Incidental vocabulary learning through reading

Research has shown that incidental vocabulary learning through reading occurs for both L1 learners (e.g. Jenkins, Stein, & Wysocki, 1984; Nagy, Anderson, & Herman, 1987) and L2 learners (e.g. Dupuy & Krashen, 1993; Pitts, White, & Krashen, 1989). Incidental vocabulary learning is a gradual process that involves the accumulation of knowledge through repeated encounters in context. It may typically begin by learning a word's form in the first few encounters, with the form–meaning connection and collocations acquired after a greater number of repetitions (Schmitt, 2010; Webb, 2007). One factor – frequency of occurrence – has often been examined in relation to learning the form–meaning connection (Horst, Cobb, & Meara, 1998; Pigada & Schmitt, 2006; Rott, 1999; Waring & Takaki, 2003; Webb, 2007). The number of encounters necessary to learn the form–meaning connection has varied between studies. Rott (1999) found that 6 encounters are needed, Horst et al. (1998) suggest 8 encounters, Webb (2007) suggests 10 encounters, and Waring and Takaki (2003) found 20 encounters may be necessary. Overall, the research has shown that frequency of occurrence of target vocabulary affects incidental vocabulary learning.

Several studies have examined the correlation between learning the form–meaning connection and the number of times each word occurred in the text. Saragi et al. (1978) found a correlation of .34, Horst et al. (1998) reported a correlation of .49, Vidal (2011) found a correlation of .47, and Webb (2007) reported correlations of .43 and .23 for recall and recognition of meaning, respectively. Horst (2005) suggests that this line of research has not captured the essence of incidental vocabulary learning through extensive reading because the research has looked at the effects of repetition in a single text rather than in multiple texts. Encountering unknown words n times in a single text may lead to a different degree of vocabulary learning than encountering unknown words the same number of times in multiple texts. Moreover, the distribution of occurrences of words or the number of different texts that the unknown words are encountered in may have an effect on incidental learning gains.

2 Incidental vocabulary learning through extensive reading

The extent of incidental learning of the form–meaning connection through reading a single L2 text is relatively small. Pitts, White, and Krashen (1989) found participants in English as a second language (ESL) learned 6.5%–8.6% of target items through reading a 6700 word text. Zahar,

Cobb, and Spada (2001) reported that ESL participants' vocabulary knowledge increased by 7.2% through reading a graded story. Horst, Cobb, and Meara (1998) found relative vocabulary gains of 22% through reading a graded reader, and Waring and Takaki (2003) reported that participants in English as a foreign language (EFL) could recognize the form–meaning connection of 42% of target words and recall the meanings of 18% after reading a modified graded reader.

Few studies have investigated vocabulary learning through reading multiple texts, and these studies are limited by the fact that they have either used standardized vocabulary tests to measure students' vocabulary learning rather than measuring specific words that were encountered during reading, or used relatively small samples of participants that limit the generalizeability of the findings.

In a series of three experiments, Lee (2007) found that Taiwanese EFL learners involved in extensive reading over a 12-week period had similar gains in vocabulary growth as participants who learned through traditional instruction. However, in two year- long experiments participants who were involved in extensive reading made greater gains in vocabulary knowledge than those who had traditional instruction. The results of these experiments provide some support for the value of extensive reading on vocabulary learning. However, it is difficult to attribute the results entirely to extensive reading because the test used to measure vocabulary

growth, the Vocabulary Levels Test (VLT), did not measure knowledge of any specific words encountered in the texts and thus did not capture the extent learners could possibly gain from the study content.

Three studies examined the degree to which L2 learners learned target words through extensive reading (Cho & Krashen, 1994; Horst, 2005; Pigada & Schmitt, 2006). Cho and Krashen (1994) looked at the vocabulary learning of four participants who read a series of books over four months. Target vocabulary was identified by the words that were underlined or written in notebooks during reading. Knowledge of the form–meaning connection of these words was measured through individualized vocabulary tests after the extensive reading treatment. The acquisition rates were 56%, 80%, 69%, and 43% respectively for each participant, for an overall raw average of 62%.

Horst (2005) used an innovative design that involved analysing the vocabulary in 20-page excerpts of 37 graded readers to identify target words. These excerpts represented two thirds of the materials that could be read by 17 adult ESL learners who took part in the six-week study. A vocabulary confidence level test with three rating options (*know, not sure, do not know*) was used to measure knowledge of 100 items, among which 50 words were from 1001–2000 word list, and 50 were less frequent words. The results indicated that the participants' knowledge improved for 76% of the higher frequency words and 62% of the lower-frequency

words, with an overall increase of 69%. A 35-item vocabulary knowledge scale was also used to measure knowledge of the lower frequency words that were indicated as being unknown on the pretest. The results indicated that 51% of these items were learned to some degree. Horst's (2005) experimental design and findings are valuable and provide powerful evidence of the positive effect of extensive reading on vocabulary learning. Horst reports that the results were limited by the fact that the analysis of the vocabulary in the text was imprecise; only portions of the texts were analysed so it is difficult to provide a very precise indication of learning. Because the participants were learning in an ESL context, there was also the potential for them to learn the target words through outside sources to some degree. Further research following up Horst's design with a more complete analysis of the texts in an EFL context would shed greater light on the influence of extensive reading on vocabulary learning.

Pigada and Schmitt (2006) carried out an in-depth study with one French language learner that investigated the degree to which three aspects of vocabulary knowledge were gained through reading four graded readers over one month. A one-on-one interview was used to measure knowledge of 133 words. The results revealed that knowledge of 66 words (50%) increased for one type of word knowledge, 13 (10%) for two types, and 8 (6%) for three types. Overall, there were gains in knowledge of 65% ($66 + 13 + 8/133$) of the words. In contrast to the two

other studies that examined how knowledge of specific words was affected by reading multiple texts, the participant in this study neither consulted the unknown words in a dictionary nor took notes about the unknown words. However, the participant was aware that he would be tested after reading, which might have raised his attention toward vocabulary learning. Although the overall pick-up rate was 65%, this has to be interpreted cautiously because much of these gains were for knowledge of written form, not form–meaning connection.

Compared to the gains documented in studies of incidental vocabulary learning through reading single texts, the acquisition rates were very high in the studies that looked at learning through multiple texts. However, one limitation of the longitudinal studies that may account in part for the higher gains was the lack of control for possible outside learning and a learning effect from taking tests at different retention intervals (completion of an earlier test may have a positive effect on subsequent tests despite learners not participating in treatments). Due to the small number of studies, small number of participants, lack of delayed posttests to measure long-term retention in any of the studies, and need for control for outside learning, further research is warranted. One aspect of vocabulary learning that has not been examined in any of the studies is how the distribution of encounters across texts (total number of books each target word was encountered in) affects learning. For example, the word *drop* is

encountered six times in total in three graded readers from the Oxford Bookworm series. If one reads the three books, he or she will encounter the word three times in *The Children of the Forest*, two times in *Robinson Crusoe*, and one time in *Anne of Green Gables*. Research on spacing of encounters indicates that a larger distribution of encounters would lead to greater learning than smaller spacing (e.g. Cepeda, Vul, Rohrer, Wixted, & Pashler, 2008). Studies that have examined learning through reading a single text may thus represent a relatively small interval between encounters and could therefore underrepresent the size of gains that might occur through more spaced encounters in multiple texts. There is no research that looks at the effects of frequency of occurrence and distribution of occurrence across texts on incidental vocabulary learning. Research that examines both frequency and distribution of occurrence in extensive reading would shed light on incidental vocabulary learning in a more ecologically valid context.

3 *Vocabulary learning through audio-assisted reading*

All of the above studies, with the exception of Horst, Cobb, and Meara (1998), looked at the effects of silent and unassisted reading on incidental vocabulary learning. However, the practice of assisted reading using oral rendition of the texts (simultaneous listening and reading) has had a long history in the development of first language literacy, and also as a remedial

approach for children who have reading difficulties (Beers, 1998; Carbo, 1978; McMahon, 1983). Two recent studies examined the practice of L2 reading while simultaneously listening to an aural version of the text. Brown, Waring, and Donkaewbua (2008) looked at three modes of input: reading only, reading while listening, and listening only, on vocabulary learning with 35 Japanese learners studying three graded readers. The results showed that the participants learned the most words in the reading while listening mode, followed by reading only and then listening only. Webb and Chang (2012a) also found evidence supporting the value of audio-assisted reading with 82 Taiwanese students. Similar to the results of Brown et al., the participants who received audio-assisted repeated reading gained greater vocabulary knowledge than those who were involved in unassisted repeated reading. One reason for greater vocabulary learning through audio-assisted reading is that reading while listening contributes to superior comprehension than reading alone (Webb & Chang, 2012a). Superior comprehension may allow readers to pay greater attention to unknown words. A second reason is that the audio support helps learners segment the texts into larger chunks, allowing learners to have greater working memory to comprehend the texts and increase the potential to infer unfamiliar words (Brown et al., 2008). A third reason is that learners may have greater knowledge of the spoken form of some words and encountering these words in both written and spoken form

may help to link form to meaning (Webb & Chang, 2012a).

4 *The present study*

Taken together, previous research has shown that L2 readers may incidentally acquire vocabulary through reading, and the learning rate could be improved if L2 readers were assisted by listening to audio recordings while reading. However, the majority of the studies either lacked ecological validity by limiting reading to a single text, had a relatively small number of participants, or lacked delayed posttests to measure retention. Furthermore, although frequency of occurrence has been found to be an important factor affecting vocabulary learning through reading a single text, it is not clear how frequency and distribution of occurrence affect vocabulary learning through reading multiple texts. Therefore, to fill the gaps in these areas, the following research questions will be addressed in the present study:

1. To what extent is vocabulary learned through reading while listening to 10 graded readers?
2. What is the relationship between frequency of occurrence in the graded readers and vocabulary learning?
3. What is the relationship between distribution of occurrence in the

graded readers and vocabulary learning?

II Method

2 Participants

The participants consisted of 82 15- and 16-year-old secondary school students (year 10) learning English as a foreign language in two classes in Taiwan. All of the participants had received three years of formal compulsory English education prior to this study, were taking the same courses at the same grade level, and were at a similar L2 proficiency level. The two intact classes were randomly assigned to one of two learning conditions. Sixty-one of the participants were assigned to an experimental group that learned English through extensive reading and the remaining 21 participants made up a control group. A version of the VLT (Schmitt, Schmitt & Clapham, 2001) containing 1000¹, 2000, and 3000 levels, was administered to participants. The VLT results showed that the experimental participants scored 44/90 and 35/90 for the control group. The VLT scores indicate that the experimental and control groups knew around 1465 and 1150 of the most frequent 3000 words, respectively. The students in both groups each had four hours of English instruction per week that were entirely devoted to English language learning. However, the four hours for the experimental group were devoted to extensive reading, while the control group received form-focused instruction.

The purpose of the control group was to determine the extent to which learning might occur outside of the treatment; this allowed us to control for the possibility of a learning effect from taking the pretest, as well as the possibility that outside learning might occur. Therefore, we shall call this a control condition rather than a comparison condition because it was simply a design feature that helped to provide a more precise measure of learning from the treatment. If there were no pretest–posttest gains for the control group, any gains for the experimental group could be attributed to the treatment.

3 *Study materials*

The following 10 level one graded readers and their corresponding professionally recorded audio versions from the Oxford Bookworm series were used as the learning material for the experimental group:

Love or Money; The Elephant Man; The Monkey's Paw; A Little Princess; The Witches of Pendle; The Phantom of the Opera; Remember Miranda; The Coldest Place on Earth; The Withered Arm; Goodbye Mr. Hollywood.

The vocabulary in the books was analysed with the RANGE software (Nation & Heatley, 2002) together with the 1000 and 2000 word lists from West's (1953) General Service List and Nation's (2006) proper

nouns list. The distribution of the vocabulary in the three lists is presented in Table 1. The total number of running words in the 10 books was 56,188, and 87.17% of these were from the 1000 word level, 5.28% were from the 2000 word level, and 3.77% were proper nouns. Because proper nouns are signaled by the capitalization of the first letter and represent a relatively small learning burden, these words are often considered as known words for non-beginners such as those in the present study (Nation, 2006; Nation & Wang, 1999; Nation & Webb, 2011; Webb & Macalister, 2013). The analysis of these texts was comparable to Nation and Wang’s (1999) and Webb and Macalister’s (2013) analysis of graded readers. Although the cumulative coverage figures indicate that these books might be somewhat difficult for

Table 1. Lexical profile of the graded readers.

Word list	Tokens	Percentage	Cumulative coverage including proper nouns
1000	48,527	87.17	90.94
2000	2964	5.28	96.22
Lower frequency words	2579	4.59	100.01
Proper nouns	2118	3.77	
Total	56,188		

the students, Nation and Wang (1999) suggest that level 1 graded readers may still be at an appropriate level for the early stages of extensive reading ‘because learners will come to these with widely differing vocabulary knowledge both in terms of size and particular words known’ (p. 361), and graded readers also provide some support for unknown vocabulary

through the use of pictures. More importantly, these books had been read by previous students with a similar lexical profile and were found to be quite readable. To reflect typical extensive reading program conditions, the participants had access to dictionaries and were allowed to ask questions after reading. While reading, they were also assisted by the aural rendition of the texts.

4 *Target words*

One hundred target words were quasi-randomly selected from the 10 graded readers (sampling was quasi-random because proper nouns and function words were excluded from selection). Typically the lowest frequency words are chosen as target words in incidental vocabulary learning studies (see, for example, Horst, Cobb, & Meara, 1998). However, random sampling of vocabulary for target items may provide a useful assessment of vocabulary learning at different frequency levels and better capture the essence of extensive reading. There is often an assumption in studies of vocabulary learning through reading that high- frequency words are known based on test scores. However, this may underrepresent learning because it is likely that most high-frequency words are only known to a certain degree, from unknown to different degrees of partial knowledge (Webb, 2012). The frequency of occurrence of target words ranged from 1 to 70, and 37/100 words occurred more than 10 times. The distribution of occurrence ranged

from 1–9 books; 73 words were encountered in one text, 16 words were encountered in two texts, and 11 words were distributed between 3–9 texts. Although the majority of words had a distribution of one, the quasi-random selection procedure provided an accurate reflection of distribution of encounters in grading reading schemes. Moreover, the number of words that had a distribution that was greater than one (27) was still larger than the number of target items in many studies of vocabulary learning through reading. Thus, the sample size for distribution of occurrence was seen to be sufficient to examine this variable in the study.

Because the words were randomly chosen, some words were likely to be known to some degree by some of the participants. Although this was likely to reduce the sample size, it had the benefit of encouraging engagement on the pretest and posttest for both groups; a test measuring knowledge of almost entirely unknown words may not have led

Table 2. Example test item taken from the pretest.

___ rope	a. 主任主管	(director)
___ chandelier	b. 吊燈	(chandelier)
___ crash	c. 面具	(mask)
___ director	d. 歌劇	(opera)
___ mask	e. 烏龜	(turtle)
___ passage	f. 走廊	(passage)
___ opera	g. 舞台	(stage)
___ shadow	h. 影子	(shadow)
___ stage	i. 撞碎	(crash)
___ torture	j. 繩索	(rope)
	k. 折磨	(torture)

to meaningful effort and accurate scores. The 100 target words were

made up of 31 words from the 1000 word level, 36 from the 2000 level, 3 from the 3000 word level, and 30 from outside of those lists.

The control group's study material was examined to determine the extent to which the target words might be encountered in their classes. There were four words (*enormous, opera, beach, hang*) that each appeared once in their course book. Therefore, based on the presence of target words in the materials of the two groups, there was greater potential for the experimental group to learn 96 of the target items.

5 *Dependent measures*

A single test rather than multiple-tests was used to measure vocabulary learning. Using more than one vocabulary test can lead to a learning effect that contributes to improved performance on subsequent tests, and may also tip off participants about the aims of the study. In the present study, a bilingual matching test was used to measure knowledge of the target words on a pretest, posttest, and delayed posttest. At each retention interval, the word order in each block was randomized. In the test, the 100 target words were divided into 10 blocks of 10. The L1 meanings of the 10 target words and one distractor were provided in each block. Students had to choose the correct Chinese L1 meaning for each target word. An example taken from the pretest is shown in Table 2.

The pretest and the VLT were administered to both groups one week before the treatment began. They completed the posttest one week after the participants in the experimental group read the 10th graded reader. The delayed posttest was completed three months after that. All participants were unaware of the nature of the study and that there would be any posttests measuring vocabulary learning.

6 *Procedure*

The experimental group read and listened to one graded reader once in class each week. All participants read the same graded reader. Due to national holidays and other school activities, the treatment was completed over 13 weeks. Because these students had never experienced reading independently, in the beginning the reading process went slowly and some guidance was given by the instructor. For example, while reading and listening to the first or second chapters of a book, the teacher usually wrote the names of the characters or places on the board, stopped the audio recording, and made sure the relationship between the characters and where the story took place were understood. When participants reached the point of fully immersing themselves in the story, the instructor did not stop the audio recording and let them finish the story straight through. Most of the time students were able to finish a book within a class unless the class was interrupted by

unpredictable factors. Reading the book again during the weekend was encouraged but not required because there was sufficient time to finish it in the classroom.

Post-reading activities involved discussions about the storylines, characters, and cultural content, or reading aloud segments of the book. All of these activities were student-led and voluntary. There was no formal teaching of any target vocabulary, but at times the instructor noted a few non-target words to raise students' interest. Keeping a learning journal and writing book reports after reading were encouraged but not required.

There was a 5-week break after the first posttest that was followed by seven weeks of classes before the delayed posttest was administered to the participants. During the seven weeks of classes the participants in both groups continued learning in the same conditions; the experimental group read and listened to 3 new level 2 graded readers and the control group learned through course book based activities.

7 Data analysis

Because the participants had varying levels of knowledge of the target items on the pre-test, relative rather than absolute learning gains were examined. Relative gains take into consideration the varying opportunities

between participants for increases in knowledge, whereas absolute gains do not. Relative learning gains were calculated for both groups using the formula $[(\text{posttest score} - \text{pretest score}) / (\text{number of test items} - \text{pretest score}) \times 100]$. A similar formula was used to calculate the retention rate $[(\text{delayed posttest score} - \text{pretest score}) / (\text{number of test items} - \text{pretest score}) \times 100]$. These results provided the information necessary to answer the first research question.

To answer the second and the third research questions, the frequency of all the target words in the graded readers was counted in two ways: (1) total number of encounters of target words in all books (frequency of occurrence), and (2) total number of books each target word was encountered in (distribution of occurrence). This data was examined in relation to the number of experimental participants who learned the target words. The number of correct responses for a target word were tallied in the pretest, the posttest and the delayed posttest. This allowed us to calculate the absolute gains (the difference between the scores on the pretest and the posttests), and the relative gains for each target word. The relationships between the frequency of occurrence and distribution of occurrence and relative gain for each target word were analysed using Spearman's rho rather than Pearson product moment correlations because the data was not normally distributed.

Table 3. Descriptive statistics for vocabulary learning rates.

	Group	Mean	SD	Minimum	Maximum
Pretest	Extensive	49.66	16.41	17	83
	Control	32.00	9.17	17	51
Posttest	Extensive	69.34	22.69	24	100
	Control	36.43	7.19	27	50
Delayed posttest	Extensive	68.82	26.31	4	100
	Control	27.67	10.12	10	47
Absolute gain (pre to post)	Extensive	19.72	15.16	-12	60
	Control	4.43	9.45	-20	20
Relative gain (pre to post)	Extensive	44.06	36.54	-63.16	100
	Control	5.19	15.11	-40.82	28.57
Absolute gain (pre to delayed)	Extensive	14.69	17.70	-26	68
	Control	-4.33	11.31	-23	18
Relative gain (pre to delayed)	Extensive	36.66	38.57	-42.59	100
	Control	-7.70	17.10	-37.10	25.35

III Results

1 Vocabulary learning rates and retention rates

The descriptive statistics (mean scores, standard deviations, absolute and relative gains) are presented in Table 3. Let us look at the results of the extensive reading group first; as shown, at pretest these participants scored 49.66/100 words correctly, with scores ranging from 17 to 83. After reading the 10 books, the experimental group had a mean score of 69.34/100 words. Overall, these participants gained an average of 19.68 words, with an average of about 31 words remaining unknown. A *t*-test indicated that the gain from pretest to posttest was statistically significant, $t(60) = -10.16$, $p < .001$. Six participants in the experimental group had scores of 100, with gains ranging from 23–60 words; however, five participants' scores decreased from 2 to 12 words. Relative

gains were calculated because they reveal a more accurate measure of learning than absolute gains when there is a large difference in potential gains (Horst, Cobb, & Meara, 1998; Shefelbine, 1990). The mean relative gain on the immediate posttest was 44.06%. The range in gains varied from -63.16 to + 100. The delayed posttest measured retention of words learned. As shown in Table 3, the relative gain from pretest to delayed posttest was 36.66%, indicating an average decrease of 7.40%. The difference between posttest and delayed posttest scores was not significant (α set at .025), $t(60) = 2.22, p = .03$.

The control group scored 32/100 on the pretest, with scores ranging from 17 to 50. The mean score on the posttest was 36.43/100, indicating a gain of 4.43 words. The relative gain from pretest to posttest was 5.19%. A *t*-test shows that there was no statistically significant difference from pretest to posttest (α set at .025), $t(20) = -2.15, p = .04$. The mean score on the delayed posttest was 27.67/100, which was lower than the pretest score and indicated a relative gain of -7.70%. The decrease between pretest and delayed posttest was not statistically significant, $t(20) = -1.76, p = .09$.

Table 4. Frequency of occurrence and relative gain.

Frequency of occurrence	Number of words	Mean of relative gain from pretest to posttest (SD)	Mean of relative gain from pretest to delayed posttest (SD)
-------------------------	-----------------	---	---

1–5	47	40.82 (21.57)	28.94 (29.47)
6–10	16	53.45 (19.17)	17.36 (52.01)
11–15	12	44.51 (21.83)	– 17.98 (153.28)
16–20	9	24.94 (16.38)	–3.04 (56.31)
21–25	6	47.45 (32.87)	2.73 (66.69)
Above 25	10	47.63 (25.52)	7.20 (64.09)
Total	100	42.93 (22.66)	14.80 (67.45)

2 *Correlations between frequency and vocabulary learning and retention*

The relationship between relative gain and overall frequency of occurrence in the 10 graded readers was investigated using Spearman's rho correlation coefficient. No words were correct for all of the experimental group in the pretest; therefore 100 words were examined.

The correlation between relative gain in the posttest and the frequency of occurrence was found to be negatively low and non-significant ($r = -.03, p = .78, n = 100$). No variance was shared by the two variables.

To determine if the relative gains increased as the number of encounters increased, the frequency of occurrence of the 100 target words was classified into six categories (1–5, 6–10, 11–15, 16–20, 21–25, +25 encounters). Table 4 shows that the relative gains on the posttest tended to increase as the number of encounters increased. A one-way analysis of variance indicated that there were no significant differences between each category, $F(5, 100) = 2.17, p > .05$). Figure 1 charts the relative gains

according to the six frequency categories. The correlation between relative gains on the delayed posttest and frequency was not statistically significant ($r = -.17, p = .09, n = 100$). The two variables shared about 3% of their variance in a negative way.

3 Correlations between distribution of occurrence across texts and vocabulary learning and retention

This section examines the relationship between distribution of occurrence across texts and vocabulary learning and retention. Table 5 shows that the correlation between the distribution of occurrence across texts and relative gain at the posttest was low ($r = .04, p = .73, n = 100$), with the two variables sharing 1.6% of their variance. The same can be said for the correlation between the two variables after three months ($r = -.41, p = .67$). Because there were relatively few words that had a large distribution, these results should be interpreted with caution.

Figure 2 shows that the mean relative gain on the posttest increases as the distribution of occurrence increased except for 4 words appearing in 5–9 texts. As shown in Table 5, the relative gains increased gradually from one occurrence of 42.68, to four occurrences of 55.97. However, there was no fixed pattern shown in the delayed posttest. The mean

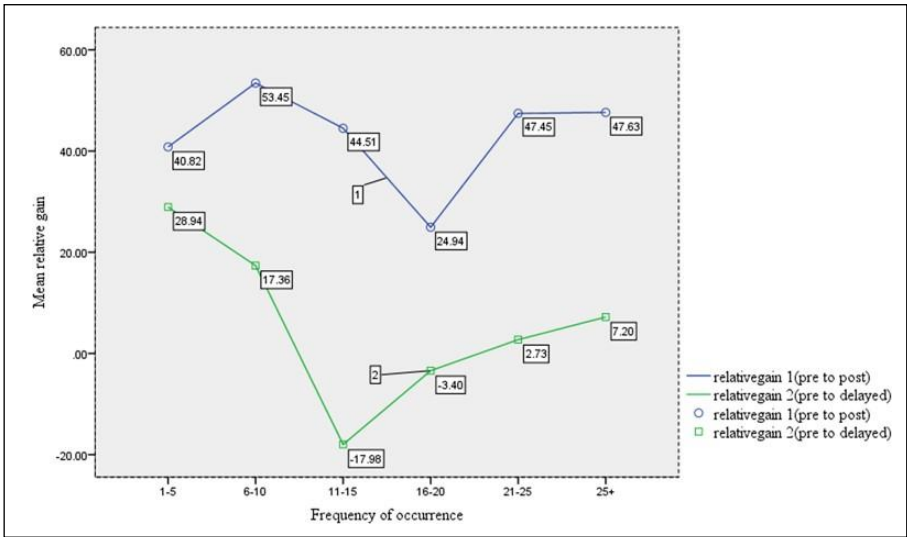


Figure 1. Mean relative gains at different numbers of encounters.

Table 5. Mean of relative gains on the posttest and delayed posttest according to the distribution of occurrence.

Distribution across texts	Number of words	Mean of relative gain from pretest to posttest	SD	Mean of relative gain from pretest to delayed posttest	SD
1	73	42.68	23.28	20.84	40.58
2	16	42.16	20.22	-.39	134.67
3	3	51.30	5.84	33.64	17.27
4	4	55.97	34.64	-14.70	101.62
5, 6, 9	4	31.42	22.83	-19.40	79.58
Total	100	42.93	14.80	14.80	67.45

relative gain on the delayed posttest dropped sharply, in particular for distributions of 4 and 5, 6 and 9 texts.

IV Discussion

1 Vocabulary learning rates

In answer to the first research question, the results indicated that the experimental group made sizeable gains in receptive knowledge of the

form–meaning connection through reading and listening to multiple graded readers. The participants’ scores increased by 19.68 words from 49.66 to 69.34. The relative learning gain immediately after the extensive reading program was 44.06%, which is slightly lower than previous small-scale studies that looked at vocabulary learning through extensive reading. Cho and Krashen (1994) found gains of 62%, and Horst (2005) found a gain of 69% on a vocabulary confidence test and 51% on a vocabulary knowledge scale. The results of the delayed post- test showed that the relative learning gains three months after the treatment were 36.66%. These findings contrast those of Waring and Takaki (2003), who found that recognition of form–meaning connection was 24.4% and recall of meaning was 3.6% three months after participants read a single graded reader.

The results of the present study suggest that relatively large vocabulary learning gains may occur through reading and listening to multiple texts. The size of the gains in this study contrast the relatively small gains found in earlier research that has examined learning through reading a single text (Horst, Cobb, & Meara, 1998; Pitts, White, & Krashen, 1989; Waring & Takaki, 2003; Zahar, Cobb, & Spada, 2001).

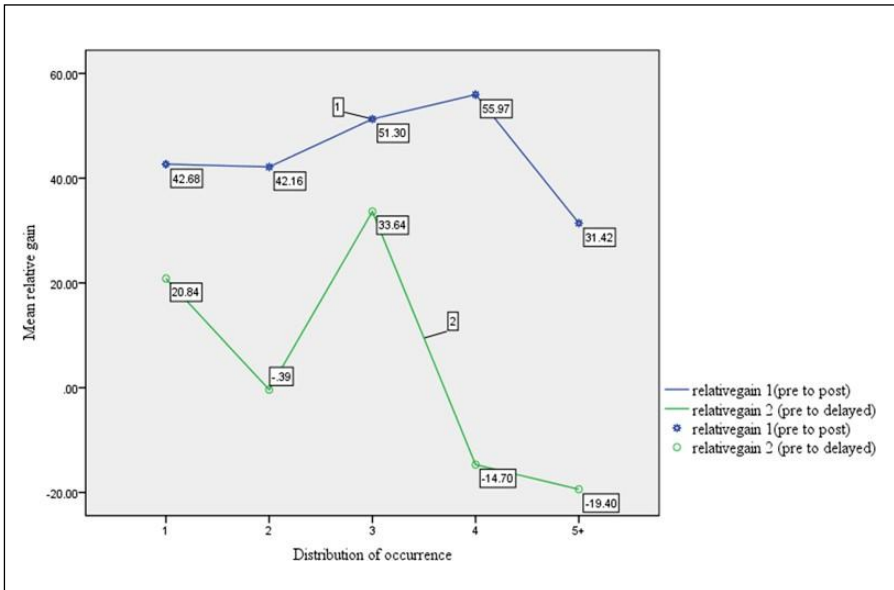


Figure 2. Distribution of occurrence and relative gain (as a percentage).

There are five reasons why reading and listening to multiple texts may have a greater impact on vocabulary knowledge than reading a single text. First, readers in this study were assisted by the aural rendition of the texts. Aural support during reading has been found to have a positive effect on L2 vocabulary learning (Brown et al., 2008; Webb & Chang, 2012a). Second, the participants in this study may have become more accustomed to reading while listening to graded readers than those who took part in studies involving reading a single text. Familiarity with extensive reading may have supported comprehension and helped to keep interest high, leading to greater attention paid to unknown words in the text. Third, because the experimental group read the same graded reader in each class,

there was greater potential for class discussion about the content. Discussion may have led to negotiation of unknown word meaning and occasional use of the unknown words. Productive activities (Webb, 2005) and negotiation of word meaning (Newton, 1995) have both been found to facilitate learning. Fourth, the design of the present study allowed for participants to use learning journals and dictionaries after reading and listening to stories. Although this is consistent with extensive reading programs and earlier studies (Horst, 2005; Cho & Krashen, 1994), it means that vocabulary learning in this study may not be purely incidental. However, the findings reflect what occurs in most extensive reading programs, where learning may at times be incidental and at times deliberate. Fifth, reading multiple graded readers provides greater opportunity for repetition and in turn consolidation of knowledge of unknown and partially known words. Although the correlations between frequency and distribution of occurrence indicated a minimal relationship with vocabulary gains, these findings and those of earlier studies suggest that frequency does play a role in learning. In this study, the effects of frequency may have been outweighed by other factors.

The results also suggest that retention of words learned through extensive reading may be much higher than previously indicated (Waring & Takaki, 2003). This may in part be due to the participant's continued involvement in the extensive reading program after the posttest. During

the interval between posttest and delayed posttest, the experimental group read and listened to a further three graded readers. Twenty-eight of the target words appeared between 1–50 ($M = 7$) times in these books with a distribution of occurrence ranging from 1–3 ($M = 1.54$). A t -test that compared the mean relative gains on the delayed posttest of these words and the words that were not encountered in the three texts indicated that there was no difference between the two sets [$t(87) = .54$, $p = .59$]. This was perhaps due to the fact that many of these items were known on the pretest. However, it also suggests that in this study encountering the words further may have had little effect on retention. The findings on the delayed test may better reflect vocabulary learning through extensive reading than if reading had ended because an aim of such programs is to encourage learners to continue reading on their own. Studies that involve no further reading or use pseudowords may misrepresent the potential vocabulary learning gains on delayed posttests because there should be the assumption that reading (and further encounters with target items) will continue post treatment.

2 The relationship between frequency of occurrence and vocabulary learning

In answer to the second research question, the correlation between frequency and relative gain was not significant ($r = .03$). The frequency

of occurrence of the target words in the texts was very broad (1–70 encounters), which may represent the frequency of encounters with words in extensive reading. The non-significant correlation between the two variables in the present study contrasts those of earlier studies (Horst, Cobb, & Meara, 1998; Saragi et al., 1978; Vidal, 2011; Webb, 2007). This is likely due in part to the larger interval between reading and testing. In earlier studies, tests were typically administered immediately after the completion of reading. However, in the present study the interval between encountering target words and testing ranged from 1–13 weeks. Because many target words were tested weeks after being encountered, the effect of frequency of occurrence might have been diluted. Instead other factors that affect retention such as interference, redundancy, and overshadowing that may influence the degree to which the target words were attended to may have played a larger role (Ellis, 2006). It may be that while frequency of occurrence clearly leads to short-term learning, it simply takes learners to a point at which time learner attention is required to consolidate learning, or durable learning requires reaching a threshold of repeated encounters that was not met in this study. With the large difference between the present findings and those of earlier studies, further research examining the effects of frequency on vocabulary learning in extensive reading is clearly warranted.

The result of the present study seem to support Nation and Wang's (1999) claim that there is no set number of repetitions that will guarantee learning, and that the relationship between repetition and word learning is likely complicated by other factors (Saragi et al., 1978). Variables that may also affect incidental vocabulary learning are proficiency (Zahar, Cobb, & Spada, 2001), illustrations in text (Horst, Cobb, & Meara, 1998), background knowledge (Pulido, 2004), and the amount of information present in the context that can be used to infer word meaning (Webb, 2008). Factors in this study that may also have affected learning were aural support, discussion of the stories, and the use of learning journals and dictionaries after reading. Together with increased familiarity with extensive reading, these factors may have led to increased comprehension, which in turn may have positively influenced vocabulary learning (Pulido, 2004).

3 The relationship between distribution of occurrence and vocabulary learning

In answer to the third research question, there was not a significant correlation ($r = .04$) between the distribution of occurrence and vocabulary learning at both retention intervals. The main reason for the lack of a statistically significant correlation might be the large number of target words (73) that appeared in only one text. This was rather surprising

because Nation and Wang (1999) found that 85.5% of the headwords from Level 1 of the same graded reading series were encountered 20 or more times in a corpus of 42 graded readers. They found that many of these headwords occur more often at the next level leading Nation and Wang to suggest that learners should only read 5–9 books at the lowest level and instead should read more at the higher levels.

Another reason for the low distribution of occurrence is that the stories selected for this study were not from the same genre (Coady, 1997). For example, *The Elephant Man* is nonfiction, *Monkey's Paw* is horror, and *Love or Money* is a mystery. If texts were selected from the same genre, it is more likely that genre specific vocabulary such as *inspector*, *blood*, and, *judge* from the mystery genre would reoccur (Hwang & Nation, 1989; Schmitt & Carter, 2000). Apart from the above two reasons, a number of other variables that may affect vocabulary learning to some degree have been identified by previous studies. These variables include phonological features, semantic content, word class, and word length (see Ellis & Beaton, 1993). However, investigating the effects of these variables was beyond the scope of the present study. One of the purposes of reading graded readers is to develop high frequency vocabulary and consolidate knowledge of partially known words, not to develop vocabulary for particular topics or genres. The results would suggest that a larger sample of texts from different levels may be needed to

provide a more accurate assessment of the effects of distribution. Another approach would be to look at distribution in narrow reading, because these texts are likely to have a large number of words with varying distributions.

The findings in this study together with those of Cho and Krashen (1994), Horst (2005), Pigada and Schmitt (2006), and Lee (2007) suggest that vocabulary learning gains made through extensive reading with audio support are likely much larger than studies examining learning from single texts have indicated. The research provides stronger evidence of the value of (1) extensive reading in L2 learning, and (2) audio support during extensive reading. One limitation of extensive reading in the L2 context is that the amount of vocabulary learned is closely tied to the amount of reading (Cobb, 2007; Laufer, 2003). Because L2 learning programs often result in relatively small amounts of vocabulary growth over many years of study (Webb & Chang, 2012b), progressive approaches to learning vocabulary involving greater amounts of L2 input need to be undertaken. A second limitation is that there is little comprehensible reading material that can be used to develop mid-frequency vocabulary. However, this has recently been rectified with the development of mid-frequency adapted novels (Nation, 2014). The findings indicate that greater emphasis should be placed on the development of extensive reading programs.

4 Pedagogical implications

The findings provide further support for incorporating extensive reading into L2 learning programs. When taken together with earlier studies of extensive reading (Cho & Krashen, 1994; Horst, 2005; Pigada & Schmitt, 2006), the results suggest that the proportion of words that are learned through reading is higher when many texts (rather than a single text) are read. Because the majority of L2 learners struggle to learn vocabulary (Webb & Chang, 2012b), greater emphasis on extensive reading in the classroom, and developing motivated readers may provide the best path for lexical development. Moreover, with the availability of audio support in most current graded reading schemes, reading while listening rather than reading alone should be the primary approach to extensive reading today. This is supported by the impressive gains in studies of vocabulary learning through reading while listening (Horst, Cobb, & Meara, 1998; Webb, Newton, & Chang, 2013), as well as those that have shown that audio-assisted reading contributes to greater vocabulary learning than reading alone (Brown, Waring, & Donkaewbua, 2008; Webb & Chang, 2012a).

The lack of a frequency effect in this study may indicate that while repetition may play a large role in the short-term (when participants read a single text, repeated encounters with words in that text may lead to

greater attention paid to those words, and in turn greater knowledge demonstrated on immediate posttests), in the long-term other factors may also play an important role. It may be that frequent encounters with unknown or partially known words in a single text leads to an immediate gain in vocabulary knowledge, but it is the use of these words by the students after reading that contributes to more durable long-term gains. Research has indicated that using words in speech or writing can have a positive effect on vocabulary learning gains (Joe, 1998; Newton, 2013). After reading, some of the frequently encountered words might be used and retained, while others might not be used and are forgotten. Thus, post-reading activities such as discussion, book reports, and keeping learning journals in extensive reading programs might enhance lexical development and should be encouraged. Furthermore, although the use of common readers rather than having students choose their own texts may potentially reduce enjoyment (Day & Bamford, 1998), it may increase the chances that students recycle and use new vocabulary. Research examining the use of target vocabulary in post-reading activities is clearly warranted.

5 *Limitations*

Several limitations of the present study should be considered when interpreting the data. First, the learning gains were limited to receptive

knowledge of form–meaning connection; neither productive knowledge of target words nor other aspects of vocabulary knowledge were assessed. Receptive knowledge of form–meaning connection is perhaps the most important aspect of vocabulary knowledge in extensive reading because comprehension of the meanings of individual words affects comprehension of the text as a whole (Schmitt, Jiang, & Grabe, 2011). However, scores on receptive vocabulary tests are likely to be higher than those on productive vocabulary tests (e.g. Webb, 2005, 2009), so it would be useful for future studies to look at how extensive reading affects productive knowledge and different aspects of word knowledge. Second, this study adopted audio-assisted extensive reading due to the popularity of audio books, the participants’ language level, and the positive impact that audio support can have on incidental vocabulary learning (Brown et al, 2008; Webb & Chang, 2012a). Because vocabulary learning gains are likely to be larger for audio-assisted reading than reading alone, the results should not be generalized to reading without audio support. Third, the results may not reflect those of extensive reading programs that allow each student to select their own text. Because participants could discuss the stories, there may have been greater negotiation of the meanings of target words and productive use of target words than might happen if different students were reading different texts. The design of the present study involved reading and listening which made the selection of different texts by different

students impossible in this learning context. Finally, although the two groups in this study had similar L2 learning profile (age, years of L2 study) and were taking the same course in the same grade, the VLT scores did indicate that the experimental group had greater vocabulary knowledge than the control group. Thus, it may be useful to replicate this study using two groups that are also matched for prior vocabulary knowledge.

V Conclusions

This study shows that L2 learners may gain sizable word knowledge from the regular reading and listening of graded readers. The percentage of vocabulary learning through extensive reading appears to be much higher than that of reading a single text. However, vocabulary gains did not correlate significantly with frequency and distribution of occurrence. This was unexpected because previous studies have found at least a moderate correlation between vocabulary learning and frequency (Horst et al., 1998; Webb, 2007). With the high learning gains in the present study, it is apparent that frequency of occurrence is only one of the many factors that affect the extent to which vocabulary is retained (Ellis & Beaton, 1993). In this study, audio support, students' note-taking, consulting dictionaries, and after-reading discussion by the students may have affected word learning to some degree. However, the relatively

large vocabulary learning gains in this and earlier studies examining the effects of reading multiple texts provide strong support for audio-assisted extensive reading.

Acknowledgements

We would also like to extend our thanks to the anonymous *LTR* reviewers for their constructive comments.

Corresponding author:

Anna C-S Chang, Department of Applied English, Hsing-Wu University, New Taipei, Taiwan. Email: annachang@livemail.tw

References

- Beers, K. (1998). Listen while you read. *School Library Journal*, 30, 30–35.
- Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language*, 20, 136–163.
- Carbo, M. (1978). Teaching reading with talking books. *The Reading Teacher*, 32, 267–273.
- Cepeda, N.J., Vul, E., Rohrer, D., Wixted, J.T., & Pashler, H. (2008). Spacing effects in learning: A temporal ridgeline of optimal retention. *Psychological Science*, 19, 1095–1102.
- Cho, K.-S., & Krashen, S.D. (1994). Acquisition of vocabulary from the Sweet Valley Kids series: Adult ESL acquisition. *Journal of Reading*, 37, 662–667.
- Coady, J. (1997). L2 vocabulary acquisition through extensive reading. In: J. Coady & T. Huckin (Eds), *Second language vocabulary acquisition* (pp. 225–237). Cambridge: Cambridge University Press.
- Cobb, T. (2007). Computing the vocabulary demands of L2 reading. *Language Learning and Technology*, 11, 38–63.
- Day, R., & Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge: Cambridge University Press.
- Dupuy, B., & Krashen, S.D. (1993). Incidental vocabulary acquisition in French as a foreign language. *Applied Language Learning*, 4, 55–63.
- Ellis, N.C. (2006). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics*, 27, 164–194.
- Ellis, N.C., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning*, 43, 559–617.

- Grabe, W., & Stoller, F. (2002). *Teaching and researching reading*. Harlow: Pearson Education.
- Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study. *The Canadian Modern Language Review*, 61, 355–382.
- Horst, M., Cobb, T., & Meara, P. (1998). Beyond A Clockwork Orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language*, 11, 207–223.
- Hwang, K., & Nation, P. (1989). Reducing the vocabulary load and encouraging vocabulary learning through reading newspapers. *Reading in a Foreign Language*, 6, 323–335.
- Jenkins, J.R., Stein, M.L., & Wysocki, K. (1984). Learning vocabulary through reading. *American Educational Research Journal*, 21, 767–787.
- Joe, A. (1998). What effects do text-based tasks promoting generation have on incidental vocabulary acquisition? *Applied Linguistics*, 19, 357–377.
- Laufer, B. (2003). Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? *Canadian Modern Language Review*, 59, 565–585.
- Lee, S.-Y. (2007). Revelations from three consecutive studies on extensive reading. *RELC Journal*, 38, 150–170.
- McMahon, M. (1983). Development of reading-while-listening skills in the primary grades. *Reading Research Quarterly*, 19, 38–52.
- Nagy, W.E., Anderson, R.C., & Herman, P.A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal*, 24, 237–270.
- Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

- Nation, I.S.P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63, 59–82.
- Nation, I.S.P. (2014). How much input do you need to learn the most frequent 9,000 words? *Reading in a Foreign Language*, 26, 1–16.
- Nation, I.S.P., & Heatley, A. (2002). Range: A program for the analysis of vocabulary in texts [soft- ware]. Available at: <http://www.victoria.ac.nz/lals/staff/paul-nation/nation.aspx> (November 2014).
- Nation, I.S.P., & Wang, K.M.-T. (1999). Graded readers and vocabulary. *Reading in a Foreign Language*, 12, 355–380.
- Nation, I.S.P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Boston, MA: Heinle.
- Newton, J. (1995). Task-based interaction and incidental vocabulary learning: a case study. *SecondLanguage Research*, 11, 159–177.
- Newton, J. (2013). Incidental vocabulary learning in classroom communication tasks. *Language Teaching Research*, 17, 164–187.
- Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language*, 18, 1–28.
- Pitts, M., White, H., & Krashen, S. (1989). Acquiring second language vocabulary through reading: A replication of the Clockwork Orange study using second language acquirers. *Reading in a Foreign Language*, 5, 271–275.
- Pulido, D. (2004). The relationship between text comprehension and second language incidental vocabulary acquisition: A matter of topic familiarity? *Language Learning*, 54, 469–523.
- Rott, S. (1999). The effect of exposure frequency on intermediate language learners' incidental vocabulary acquisition through reading. *Studies in Second Language Acquisition*, 21, 589–619.
- Saragi, T., Nation, I.S.P., & Meister, C.F. (1978). Vocabulary learning and

- reading. *System*, 6, 72–128.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Basingstoke: Palgrave Macmillan.
- Schmitt, N., & Carter, R. (2000). The lexical advantages of narrow reading for second language learners. *TESOL Journal*, 9, 4–9.
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *Modern Language Journal*, 95, 26–43.
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviours of two new versions of the Vocabulary Levels Test. *Language Testing*, 18, 55–88.
- Shefelbine, J.L. (1990) Student factors related to variability in learning word meanings from context. *Journal of Reading Behavior*, 22, 71–97.
- Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning*, 61, 219–258.
- 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, 130–163.
- Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge on word knowledge. *Studies in Second Language Acquisition*, 27, 33–52.
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28, 46–65. Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language*, 20, 232–245.
- Webb, S. (2009). The effects of pre-learning vocabulary on reading comprehension and writing. *Canadian Modern Language Review*, 65, 441–470.

- Webb, S. (2012). Depth of vocabulary knowledge. In: C. Chappelle (Ed.) *Encyclopedia of applied linguistics* (pp. 1656–1663). Oxford: Wiley-Blackwell.
- Webb, S., & Chang, C.-S. (2012a). Vocabulary learning through assisted and unassisted repeated reading. *The Canadian Modern Language Review*, 68, 276–290.
- Webb, S., & Chang, C.-S. (2012b). Second language vocabulary growth. *RELC Journal*, 43, 113–126.
- Webb, S., & Macalister, J. (2013). Is text written for children appropriate for L2 extensive reading? *TESOL Quarterly*, 47, 300–322.
- Webb, S., Newton, J., & Chang, C.-S. (2013). Incidental learning of collocation. *Language Learning*, 63, 91–120.
- West, M. (1953). *A general service list of English words*. London: Longman, Green & Co.
- Zahar, R., Cobb, T., & Spada, N (2001). Acquiring vocabulary through reading: Effects of frequency and contextual richness. *The Canadian Modern Language Review*, 57, 541–572.