CrossMark

Studies in Second Language Acquisition, 2015, **37**, 651–675.
doi:10.1017/S0272263114000606

HOW DOES PRIOR WORD KNOWLEDGE AFFECT VOCABULARY LEARNING PROGRESS IN AN EXTENSIVE READING PROGRAM?

Stuart Webb
Western University

Anna C.-S. Chang
Hsing-Wu University

Sixty English as a foreign language learners were divided into high-, intermediate-, and low-level groups based on their scores on pretests of target vocabulary and Vocabulary Levels Test scores. The participants read 10 Level 1 and 10 Level 2 graded readers over 37 weeks during two terms. Two sets of 100 target words were chosen from each set of graded readers and were tested on three occasions. The results showed that the relative gains from pretest to immediate posttest were 63.18%, 44.64%, and 28.12% for the high-, intermediate-, and low-level groups, respectively. There was little decay in knowledge on the Term 1 three-month delayed posttest; relative gains ranged from 21.05% for the low-level group to 59.01% for the high-level group. The learning gains in Term 2 were consistent with those from Term 1. The results indicate that prior vocabulary knowledge may have a large impact on the amount of vocabulary learning made through extensive reading.

This study was supported by a research grant from the Ministry of Science and Technology (former National Science Council, 99-2410-H-266-003). We would like to express our thanks to Frank Boers, Victoria University of Wellington, New Zealand, for his useful suggestions while we were preparing an earlier draft of the article. We would also like to thank the anonymous *SSLA* reviewers for their helpful comments.

Correspondence concerning this article should be addressed to Anna C.-S. Chang, 101 Section 1, Feng-Liao Road, Linkou, New Taipei, Taiwan. E-mail: annachang@livemail.tw

© Cambridge University Press 2015

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press

Second language (L2) vocabulary learning through reading has been a frequently researched topic in the last two decades. Much of this research has looked at the extent to which words are learned incidentally through reading a single text (e.g., Day, Omura, & Hiramatsu, 1991; Dupuy & Krashen, 1993; Horst, Cobb, & Meara, 1998; Pellicer-Sánchez & Schmitt, 2010; Pitts, White, & Krashen, 1989). However, vocabulary learning is a gradual process with incremental gains in word knowledge occurring through repeated encounters with unknown or partially known words. Schmitt (2010) thus argues that only studies that employ longitudinal designs can accurately measure vocabulary learning. There have been very few longitudinal studies that have investigated vocabulary learning through extensive reading. This line of research has shown that reading multiple texts contributes to L2 vocabulary learning (Cho & Krashen, 1994; Horst, 2005; Pigada & Schmitt, 2006), and the size of these gains is typically larger than those of studies that have looked at learning through reading a single text. However, these studies (i.e., Cho & Krashen, 1994; Horst, 2005; Pigada & Schmitt, 2006) have been of relatively small scale with few participants.

One question that remains to be explored is how learners with different levels of word knowledge progress in their vocabulary learning through extensive reading. Learners with greater vocabulary knowledge are likely to have better reading comprehension because they are likely to know more words in a text (Hu & Nation, 2000; Laufer & Ravenhorst-Kalovski, 2010; Qian, 1999; Schmitt, Jiang, & Grabe, 2011). Similarly, those learners who know more words in a text may have greater potential for vocabulary learning because they may focus more attention on the remaining unknown words, and they are likely to have better understanding of the surrounding running words that can be used to help infer unknown word meanings. Further research investigating vocabulary learning through reading multiple texts over a long period of time by learners at different levels of word knowledge would provide a useful and ecologically valid measurement of the effects of extensive reading.

The primary aim of the present study was to expand on the earlier research by looking at the L2 vocabulary learning gains of learners with different levels of vocabulary knowledge and to determine the extent to which vocabulary knowledge correlates with learning gains. A secondary aim was to examine how students progressed in their L2 vocabulary learning through reading 20 graded readers (10 Level 1 and 10 Level 2 graded readers) from the start of an extensive reading program (a 37-week period: two 16-week terms with a 5-week break between terms) to the final 3-month delayed posttest, 49 weeks later. To our knowledge, there has been no research that has examined vocabulary learning progress through the different text difficulty levels of graded readers. This study should (a) provide a more accurate assessment of

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press

the contributions that extensive reading makes to vocabulary learning and (b) shed light on its efficacy for English as a foreign language (EFL) learners with varying vocabulary sizes.

BACKGROUND

Learning Vocabulary through Reading

It is a common belief that, once children are able to read well, most first language (L1) words are learned incidentally through reading. Research has shown that both L1 (e.g., Jenkins, Stein, & Wysocki, 1984; Nagy, Anderson, & Herman, 1987; Nagy, Herman, & Anderson, 1985) and L2 (e.g., Day et al., 1991; Dupuy & Krashen, 1993; Pitts et al., 1989) learners incidentally learn words through reading. The major difference between L1 and L2 vocabulary learning gains made through reading is that L2 learners tend to have much less L2 written input and, therefore, fewer opportunities to repeatedly encounter unknown and partially known words. Research has shown that the frequency of encounters with unknown words affects learning; the potential to learn words increases as the number of times words are encountered increases (Chen & Truscott, 2010; Horst et al., 1998; Pigada & Schmitt, 2006; Rott, 1999; Waring & Takaki, 2003; Webb, 2007).[1] Thus, L2 learners' incidental vocabulary learning gains are dependent on the amount of L2 input. Extensive reading has been promoted as a means to increase the amount of L2 input and, in turn, incidental vocabulary learning gains.

Surprisingly, there is relatively little research on L2 vocabulary learning through extensive reading. The vast majority of L2 reading studies have looked at incidental learning through reading a single text (e.g., Day et al., 1991; Dupuy & Krashen, 1993; Horst et al., 1998; Pitts et al., 1989; Waring & Takaki, 2003). Taken together, these studies tend to show that although L2 words are learned incidentally through reading, gains are relatively small (Horst et al., 1998; Pitts et al., 1989; Zahar, Cobb, & Spada, 2001). Studies that have investigated vocabulary learning through reading multiple texts have indicated that extensive reading may lead to larger gains in vocabulary knowledge than are reported in the studies examining learning through reading a single text. Cho and Krashen (1994) found that the vocabulary learning gains of four L2 learners who read several books over 4 months ranged from 56% to 80% of tested vocabulary, with an overall raw average gain of 62%. Horst (2005) found that vocabulary knowledge increased for 51% of target words on a vocabulary knowledge scale and 69% of target words on a vocabulary confidence level test for 17 adult English as a L2 (ESL) learners in a 6-week extensive reading program. In a case study of one French language learner who read four graded readers

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press

over 1 month, Pigada and Schmitt (2006) found gains in knowledge for 65% of the target words.

There are several reasons why a more robust vocabulary learning rate has been demonstrated in the extensive reading studies. First, unknown and partially known words may be encountered between texts, with later encounters consolidating previously gained knowledge. Second, delayed posttests were not administered to measure durable learning in the three extensive reading studies. Immediate posttests may not provide an accurate measurement of vocabulary learning because vocabulary knowledge is liable to decay over time (Nation & Webb, 2011; Schmitt, 2010). Third, it is important to note that the vocabulary learning gains made through the extensive reading studies may not be strictly incidental because there was potential for the learners in these studies to gain knowledge of words through dictionary lookups and discussion and writing related to their readings (Cho & Krashen, 1994; Horst, 2005). This approach has ecological validity because it is consistent with extensive reading programs in which learners are not prohibited from discussing or writing about the stories they read and are not forbidden to look up words in dictionaries. Thus, the larger gains in these studies may be due in part to vocabulary learning gains that were both incidental and deliberate.

It is not clear whether the vocabulary learning rates achieved in these three studies can be generalized to a longer extensive reading program and to texts at different levels of difficulty. It is possible that the learning rate may increase because a greater amount of text read may lead to a greater number of encounters with unknown or partially known words (Nation & Wang, 1999). Learners may also become more skillful in reading and may develop improved comprehension through reading multiple texts. This may allow them to devote more attention to unknown words and may increase the potential for vocabulary learning. Nation and Wang (1999) suggest that learners read a minimum of five books at each graded reading level at a rate of about one book per week to ensure that learners meet words again in another text before they have been forgotten. To our knowledge there is currently no research that examines students' vocabulary learning from one graded reading level to the next. The present study aims to shed light on this issue.

The Relationship between Prior Word Knowledge and Vocabulary Learning

Prior word knowledge is a factor that affects vocabulary learning during the reading of a single text. Horst et al. (1998) examined the

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press

relationship between vocabulary knowledge, as indicated by scores on the Vocabulary Levels Test (VLT; Nation, 1983), and incidental learning gains through reading and listening to a single graded reader. They found that there was a medium-sized relationship between the two variables (r = .31 at the 2,000-word level, and r = .36 at the 5,000-word level) and reported that vocabulary gains tended to be larger and more consistent for participants with greater vocabulary knowledge. Zahar et al. (2001) also examined the relationship between participants' initial vocabulary knowledge and the amount of vocabulary gained through reading one graded story. One hundred and forty-four ESL students were classified into five levels according to their scores on the VLT (Nation, 1983). After reading the text, groups at different vocabulary knowledge levels demonstrated different amounts of vocabulary learning. Of the five groups, the lowest and highest groups had the smallest gains. The reason the highest level group had small gains might be due to the fact that they knew the most words, which left a relatively small number of unknown words to be learned. Zahar et al. concluded that learners with less vocabulary knowledge may have more difficulty learning words incidentally and need to meet unknown items many more times to learn them than those with a larger vocabulary size. Zahar et al.'s study was replicated by Tekmen and Daloglu (2006) with Turkish EFL students. The results were consistent with the earlier study in that the participants with the highest scores on the VLT had significantly greater vocabulary gains than those with the lowest scores, and the effect size was large.

Taken together, these studies suggest that vocabulary knowledge may affect the ability to learn L2 words incidentally through reading. The reason for this may be that students with greater vocabulary knowledge are likely to have greater text comprehension (Hu & Nation, 2000; Laufer & Ravenhorst-Kalovski, 2010; Qian, 1999; Schmitt et al., 2011), and this may allow them to pay greater attention to unknown words. Whether the findings can be generalized to long-term extensive reading, however, is not clear. All of the studies that examined how prior word knowledge affected learning looked at vocabulary learning during the reading of a single text in a short-term study rather than multiple texts in a longitudinal study.

Research exploring the relationship between learners' preintervention word knowledge and learning through extensive reading may provide some indication of the extent to which high- and low-level students develop L2 literacy skills. Stanovich (1986) provided strong evidence that individual differences in reading ability affect L1 literacy development. He highlighted prior vocabulary knowledge as one of the factors that affect reading skills; learners with greater vocabulary knowledge are likely to have superior reading comprehension, to read more, and to learn more vocabulary through reading than

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press

learners with less vocabulary knowledge. Because students in extensive reading programs are likely to have varying levels of vocabulary knowledge, it is important to prevent the progress of lower level students from lagging far behind that of higher level students. If students with a smaller vocabulary size learn relatively few words through extensive reading over a long period of time, then they may have difficulty in moving up to subsequent graded reading levels. In a corpus-driven study of the vocabulary in different text types, Cobb (2007) found that, even if learners are successful in learning words incidentally through extensive reading, their vocabulary knowledge would still be insufficient to read authentic texts. Measuring the vocabulary learning progress of higher and lower level learners may provide a gauge for the efficacy of L2 extensive reading programs in this regard. Findings from the present study may provide some indication as to whether extensive reading programs are effective for all learners or whether they need to be revised to provide greater support for lower level learners.

The Present Study

The aims of the present research are to investigate (a) the extent to which EFL learners with larger and smaller vocabulary sizes learned vocabulary through reading 20 graded readers and (b) whether vocabulary learning gains are consistent across extensive reading text difficulty levels. The study should also serve as a useful first step in filling a large gap in the research on L2 reading by measuring vocabulary learning through reading graded readers at different text difficulty levels over a much longer period of time than in earlier studies. Moreover, a longitudinal study of vocabulary learning through reading with learners at different vocabulary levels may provide a useful representation of how effective extensive reading is in developing vocabulary knowledge for a range of EFL students.

The following research questions were addressed in the present study:

1. To what extent do EFL learners with different vocabulary knowledge levels learn vocabulary through reading and listening to 10 Level 1 and 10 Level 2 graded readers?
2. Were the vocabulary learning gains of EFL learners who read and listened to Level 2 graded readers consistent with those made through reading and listening to Level 1 graded readers?
3. What is the relationship between prior word knowledge and vocabulary learning gains made through extensive reading with audio support?

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press

METHODOLOGY

Participants

The participants were 60 sixteen-year-old secondary school students (equal to Year 10) from two classes taking the same EFL course in Taipei, Taiwan.[2] All participants had two 100-min English classes per week, and the same instructor taught all of their classes. At the start of the study, they had all received 3 years of formal English language instruction. The participants were from a test-oriented learning background and had little experience reading in English. The two groups took part in an extensive reading program, and their classes were focused entirely on learning English through extensive reading.

The participants were given the 1,000, 2,000, and 3,000 levels of a bilingual version of the VLT (Schmitt, Schmitt, & Clapham, 2001) that is freely available on Paul Nation's Web site (http://www.victoria.ac.nz/lals/about/staff/paul-nation).[3] The results showed that the participants had an average raw score of 48 out of 90 on the three levels combined. These scores indicated that the participants knew around 1,600 of the most frequent 3,000 words (a correct response on the VLT represents knowledge of 33.3 words; thus, the extent of knowledge of the different levels can be calculated as the score, 48, multiplied by 33.3). The test scores indicated that the participants probably had sufficient vocabulary knowledge to begin extensive reading of 400-headword graded readers.

The participants were classified as high (n = 16), intermediate (n = 25), or low (n = 19) level according to their scores on a pretest measuring knowledge of the target vocabulary. The VLT scores for each group were also examined to ensure that each group varied in prior vocabulary knowledge. The results showed that the total VLT scores corresponded with scores on a pretest that was designed to measure knowledge of the form-meaning connection of the target words. The high-, intermediate-, and low-level participants had average raw scores of 58 (26, 18, and 14), 48 (23, 14, and 11), and 42 (21, 11, and 10) out of 90, respectively (the numbers in parenthesis represent the scores on the 1,000-, 2,000-, and 3,000-word levels, respectively). Participants' pretest scores in each level were also consistent with the overall VLT scores, which meant that those who scored higher on the VLT also scored higher on the pretest.

Study Materials

A total of 20 graded readers from the Oxford Bookworm series (www.oup.com/bookworms) and their audio versions were used in the study

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press

(see Table 1). In the first term, the extensive reading participants read Level 1 graded readers, and in the second term, they read Level 2 books. There were 50 copies of each book, which enabled the instructor to use the graded readers as class texts. The reason for using common texts rather than letting individual students select their own books was that participants were beginner readers who had no experience reading independently. Class texts allowed the instructor to provide greater support for the students and more opportunities to clarify understanding.

The extensive reading program included audio versions of the graded readers to improve learning potential. First language research has indicated that reading while listening, compared to reading alone, may lead to superior comprehension and increased motivation to read (Koskinen et al., 2000). Studies of incidental learning through reading while listening to single texts (Brown, Waring, & Donkaewbua, 2008; Horst et al., 1998) and multiple texts (Webb & Chang, 2012) indicate that audio support can have a positive effect on vocabulary learning. Second language learners may learn more words through reading while listening than through reading alone (Brown et al., 2008; Webb & Chang, 2012).

Target Words

Two sets of target words were quasirandomly selected from the two sets of graded readers. Sampling was quasirandom rather than random because proper nouns and function words were excluded. Quasirandom sampling allowed the selection of words with different numbers of occurrences in the texts and varying frequency levels. This may provide a more accurate representation of vocabulary learning gains than frequency-based selection of items. One hundred target words were selected from

Table 1. Graded readers used in the study

Term 1	Term 2
The Elephant Man	*Anne of Green Gables*
The Monkey's Paw	*Dead Man's Island*
Love or Money	*The Death of Karen Silkwood*
A Little Princess	*New Yorkers*
Goodbye Mr. Hollywood	*Robinson Crusoe*
The Witches of Pendle	*The Piano*
The Phantom of the Opera	*Earrings from Frankfurt*
Remember Miranda	*The Pit and the Pendulum*
The Coldest Place on Earth	*Henry VIII and His Wives*
The Withered Arm	*The Children of the New Forest*

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press

the Level 1 graded readers, and another 100 target words were selected from the Level 2 books.

The target words were analyzed using the Range software (Nation & Heatley, 2002) together with Nation's (2012) twenty-five 1,000-word lists from the British National Corpus/Corpus of Contemporary American English (BNC/COCA) to determine the lexical frequency of the items. Table 2 shows the proportion of items in each set that occur in the different BNC/COCA lists. The first set (see Appendix A) included 33 words from the 1,000-word level, 32 words from the 2,000-word level, and 35 words that were outside of these lists. The frequency of occurrence of the Term 1 items ranged from 1 to 70 encounters in the set of graded readers (M = 12.34, SD = 13.91). In the second set of 100 words (see Appendix B), 20 words were from the 1,000-word level, 29 words were from the 2,000-word level, and 51 words were outside of these lists. The frequency of occurrence of the Term 2 items ranged from 1 to 22 encounters in the set of graded readers (M = 5.21, SD = 4.22).

Dependent Measures

Matching tests were used to measure prior and posttreatment knowledge of each set of target words. The participants were not informed about any of the tests. The tests were administered to participants in all groups 1 week before the start of the extensive reading program, 1 week after the extensive reading program finished in Term 1, and approximately three months after the end of the Term 1 extensive reading

Table 2. The number of target words occurring in Nation's (2012) BNC/COCA word lists

Word list	Set A	Set B
1,000	33	20
2,000	32	29
3,000	9	13
4,000	10	8
5,000	2	14
6,000	6	2
7,000	1	0
8,000	2	5
9,000	0	2
10,000	1	3
11,000	0	1
Offlist	3	3

Note. Offlist refers to words not included in Nation's (2012) 25 BNC/COCA 1,000-word lists.

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press

program. A second matching test was administered to participants at the same intervals in Term 2 to measure knowledge of the second set of target words.

In the matching tests, each set of target words was divided into 10 blocks. Each block had approximately 10 target words and one distractor, and the participants had to choose the correct L1 meaning of each target word. An example of one block taken from the pretest in Term 2 is shown in (1). The glosses provided in (1) did not appear in the actual pretest. The order of items was changed at each retention interval to reduce the potential of a learning effect from one test to subsequent tests. After the participants completed the tests, all test papers were collected to ensure that the students did not take away the test papers to study the words.

(1)

1. ____ forester	a. 水桶 “bucket”
2. ____ calf	b. 馬廄 “stable”
3. ____ colonel	c. 上校 “colonel”
4. ____ fortnight	d. 火焰 “flame”
5. ____ oatmeal	e. 燕麥片粥 “oatmeal”
6. ____ parliament	f. 國會, 議會 “parliament”
7. ____ stable	g. 小牛 “calf”
8. ____ cart	h. 馬車 “cart”
9. ____ flame	i. 兩星期 “fortnight”
10. ____ bucket	j. 林務員 “forester”
	k. 夜晚 “night”

Procedure

Each term was approximately sixteen weeks long; however, in several of the weeks, classes were canceled due to other school activities. In the first week of Term 1, the VLT was administered to all participants, followed by the 100-item target vocabulary pretest. The participants were told that their course involved reading stories and that no formal instruction would be given. At first, extensive reading was challenging for these students, and support was given to aid their comprehension. This involved the instructor stopping the audio on occasion to check the students’ understanding of the stories and writing the names of the characters or places on the board. The instructor avoided teaching language elements and never focused on target vocabulary. Instead the instructor guided students’ attention to the story line. After reading approximately four books, these participants appeared to be more comfortable with the extensive reading program, their comprehension appeared to be adequate, and some participants reported that they were enjoying it.

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press

Typically, the students completed one graded reader throughout 1 week of classes. The participants completed student-led postreading activities common to extensive reading programs. These included writing about the stories in learning journals, reading parts of the stories aloud, doing role plays related to the stories, discussing the stories, and looking up words in dictionaries. The students were not instructed to complete any particular activities, so all postreading activities were fully voluntary. Because there was potential for students to look up target words in dictionaries after reading the text, it is important to note that any vocabulary learning gains may not have been entirely incidental. However, because dictionaries may occasionally be used by students in extensive reading programs, their use in this study is ecologically valid. Moreover, disallowing dictionaries over two terms would have resulted in failure to meet institutional ethical protocols.

One week after reading the last graded reader, and again 3 months after that, the students completed the 100-item vocabulary posttests. There was a 5-week break after Term 1, and then the classes started again. Term 2 followed the same testing and learning schedule, with a pretest given in the first week followed by the extensive reading treatment and then the posttests at the same retention intervals.

Data Analysis

SPSS (Version 20) for Windows was used to analyze the data. First, the participants were classified into three levels—high, intermediate, and low—according to the pretest scores for the target words in Terms 1 and 2. In Term 1, 19 students who scored between 17 out of 100 and 40 out of 100 were placed in the low-level subgroup, 25 students who scored between 41 and 60 were placed in the intermediate group, and 16 students who scored above 60 were placed in the high-level group. In Term 2, those who scored between 10 and 20 were placed in the low subgroup (n = 26), those who scored between 21 and 30 were put in the intermediate subgroup (n = 18), and those who scored above 30 were placed in the high subgroup (n = 16).

Because real words were used as items, it was expected that the participants would score correctly for some items on the pretest. Therefore, relative gains—rather than actual gains—were examined because relative gains provide a more accurate comparison among groups when there is a difference in potential gains (Horst et al., 1998; Shefelbine, 1990). Students' immediate relative learning gains were calculated using the following formula: [(correct raw posttest score – pretest score) / (total words tested – correct raw pretest scores)] * 100. A similar formula was used to calculate the delayed relative gain:

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press

[(correct delayed posttest raw score – pretest score) / (total words tested – pretest score)] * 100.

A one-way between-groups ANOVA with post hoc Bonferroni multiple comparisons was used to determine whether the sizes of the relative gains were different between the three subgroups (high, intermediate, and low). This analysis was used to answer the first research question. To answer the second research question, paired-samples *t* tests were used to compare the relative gains for each subgroup between Terms 1 and 2.

Preliminary analysis indicated that no violation of the assumptions of linearity, normality, and homoscedasticity had occurred, so Pearson product-moment correlation coefficients were used to investigate the relationship between preintervention vocabulary knowledge and relative gains. The results of this analysis were used to answer the third research question.

RESULTS

The descriptive statistics of vocabulary test results from Term 1 and Term 2 are presented in Tables 3 and 4. On the Term 1 pretest, the high-, intermediate-, and low-level groups had mean scores of 72, 49, and 32 out of 100, respectively. At the end of Term 1, each of the three groups gained 19, 21, and 19 points and had posttest scores of 91, 70, and 51, respectively. The relative gains between pretest and posttest were 63%, 45%, and 28% for the high-, intermediate-, and low-level groups, respectively.

An ANOVA was performed to compare whether the three groups differed significantly in their learning gains. The results showed that there was a significant difference in the relative gains from the pretest to the posttest, $F(2, 57) = 4.40$, $p = .017$, and the effect size was medium, $\eta^2 = .13$. Post hoc comparisons using the Bonferroni test indicated a statistically significant difference between the high- and low-level groups, $p = .013$ (effect size by Cohen's $d = 0.98$). No significant differences were found between the low and the intermediate groups, $p = .31$, or between the intermediate and the high groups, $p = .37$.

Mean scores on the 3-month delayed posttest for Term 1 were 88, 67, and 46 for the high-, intermediate-, and low-level groups, respectively. Each group's mean scores decreased by approximately, 3, 4, and 6 points from the gains found in the posttest. The high-level group retained more than the intermediate- and low-level groups. The relative gains between pretest and delayed posttest were 59% for the high-level group, 37% for the intermediate-level group, and 21% for the low-level group. An ANOVA indicated that there was a statistically significant difference among the relative gains of the three groups, $F(2, 57) = 5.03$,

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press

Table 3. Descriptive statistics of vocabulary test scores for Term 1

Participant subgroups	Pretest	Posttest	Delayed posttest	Absolute gain (pre–post)	Absolute gain (pre–delayed)	Relative gain (pre–post)	Relative gain (pre–delayed)
High (n = 16)	71.75 (6.70)	90.63 (8.60)	88.06 (11.28)	18.88 (12.04)	16.31 (10.45)	63.18 (42.28)	59.01 (35.94)
Intermediate (n = 25)	49.48 (6.23)	70.32 (20.82)	66.52 (21.50)	20.84 (16.00)	17.04 (17.08)	44.64 (36.48)	37.23 (37.31)
Low (n = 19)	32.21 (6.02)	51.26 (17.51)	46.11 (22.65)	19.05 (17.36)	13.89 (20.83)	28.12 (25.42)	21.05 (31.85)
Total (n = 60)	49.95 (16.39)	69.70 (22.74)	65.80 (25.38)	19.75 (15.29)	15.85 (16.75)	44.36 (36.79)	37.91 (37.64)

Note. The values reported in the pretest, posttest, delayed posttest, and absolute gain columns refer to mean scores. The values in the relative gain columns refer to percentages. Standard deviations have been provided in parentheses.

Table 4. Descriptive statistics of vocabulary test scores for Term 2

Participant subgroups	Pretest	Posttest	Delayed posttest	Absolute gain (pre–post)	Absolute gain (pre–delayed)	Relative gain (pre–post)	Relative gain (pre–delayed)
High (n = 16)	36.19 (5.31)	77.00 (13.05)	68.25 (16.66)	40.81 (14.00)	32.06 (14.46)	63.64 (21.09)	51.15 (24.61)
Intermediate (n = 18)	25.00 (2.72)	55.28 (28.76)	57.83 (29.62)	30.28 (27.55)	32.83 (28.77)	40.94 (37.48)	44.20 (39.05)
Low (n = 26)	15.69 (2.71)	32.77 (18.70)	32.19 (13.72)	17.08 (19.24)	16.50 (14.75)	20.10 (22.52)	19.36 (17.00)
Total(n = 60)	23.95 (9.13)	51.32 (27.66)	49.50 (25.53)	27.37 (22.88)	25.55 (21.15)	37.96 (34.48)	35.29 (30.33)

Note. The values reported in the pretest, posttest, delayed posttest, and absolute gain columns refer to mean scores. The values in the relative gain columns refer to percentages. Standard deviations have been provided in parentheses.

p = .010, and the effect size was large, η^2 = .15. Bonferroni multiple comparisons showed that a significant difference was found only between the high- and low-level groups, p = .007. The effect size for the difference was large (Cohen's d = 1.12). No significant differences were found between the low and the intermediate groups, p = .41, or between the intermediate and the high groups, p = .18. Figure 1 illustrates the differences in relative gains among the groups in Term 1.

In Term 2, the high-, intermediate-, and low-level groups had scores of 36, 25, and 16 out of 100 on the pretest, respectively. At the end of Term 2, they had mean scores of 77, 55, and 33, indicating gains of approximately 41, 30, and 17 points, respectively. The relative gains between the pretest and the posttest were 64% for the high-level group, 41% for the intermediate-level group, and 20% for the low-level group (see Figure 2). Although the relative gains were quite similar to those of Term 1, the actual gains revealed a large discrepancy among the three groups.

An ANOVA (α level set at .017) indicated that the difference between the three groups' relative gains from the pretest to the posttest in Term 2 was statistically significant, $F(2, 57) = 12.52$, $p < .001$, and the effect size was very large, η^2 = .31. Bonferroni multiple comparisons showed that a significant difference was found between the low and high groups, $p < .001$. Because the significance level was set at .017, there were no differences between the low and intermediate groups, p = .05, or between the intermediate and high groups, p = .06.

Mean scores on the 3-month delayed posttest for Term 2 were 68, 58, and 32 for the high-, intermediate-, and low-level groups, respectively. The relative gains between pretest and delayed posttest were 51% for the high-level group, 44% for the intermediate-level group, and 19% for the low-level group. An ANOVA indicated that there was a statistically significant difference among the relative gains of the three groups,

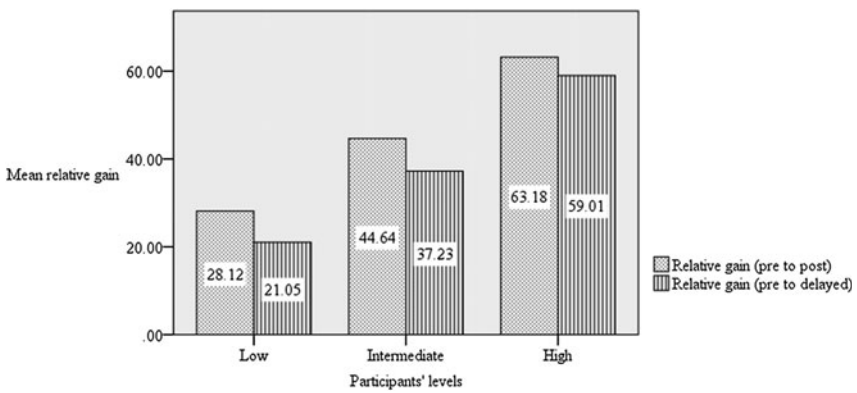


Figure 1. Relative gains for the low-, intermediate-, and high-level groups in Term 1.

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press

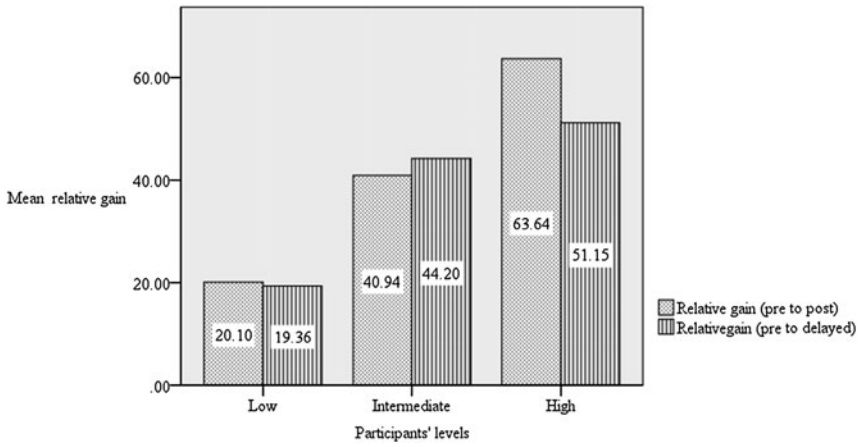


Figure 2. Relative gains for the low-, intermediate-, and high-level groups in Term 2.

$F(2, 57) = 8.14$, $p = .001$, and the effect size was large, $\eta^2 = .22$. Bonferroni multiple comparisons showed that there were significant differences between the low and intermediate groups, $p = .013$, and between the low and high groups, $p = .002$; however, no significant difference was found between the intermediate- and high-level groups, $p = 1.00$.

Independent-samples *t* tests were used to determine whether there were any statistically significant differences between the relative gains of each group in Term 1 and Term 2. The analysis indicated that there was no difference in relative gains between Term 1 and Term 2 across all three groups. That means the learning rates within each level were consistent between Term 1 and Term 2. Overall result patterns show that the relative gains from pretest to posttest ranged from approximately 20% to 28% for the low-level group, 41% to 45% for the intermediate-level group, and 63% to 64% for the high-level group. The relative gains from the pretest to the delayed posttest ranged from approximately 19% to 21% for the low-level group, from 37% to 44% for the intermediate-level group, and from 51% to 59% for the high-level group.

Paired-samples *t* tests were also performed to examine whether there were any significant differences in the relative gains from pretest to posttest and from pretest to delayed posttest within each group. This analysis indicated that there were no significant differences for the three groups in Terms 1 and 2. This means that the words learned immediately after the intervention were mostly retained for 3 months across all groups.

Pearson's correlation coefficients between the extensive reading participants' scores on the tests measuring prior vocabulary knowledge (pretest scores) and their vocabulary learning gains were also examined. In Term 1,

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press

the correlations between pretest-posttest, $r = .74$, $p < .001$, and pretest–delayed posttest, $r = .76$, $p < .001$, were large. Large correlation coefficients were also found in Term 2 between pretest and posttest, $r = .64$, $p < .001$, and between pretest and delayed posttest, $r = .62$, $p < .001$.

DISCUSSION

In answer to the first research question, the results indicated that the higher level participants had significantly larger relative gains than the low-level participants in both terms; these results support the findings of earlier studies (Horst et al., 1998; Zahar et al., 2001). In Term 1, the relative vocabulary learning gain for the high-level group was 63% between the pretest and the posttest and 59% between the pretest and the delayed posttest. In contrast, the relative gain for the low-level group was 28% between the pretest and the posttest and 21% between the pretest and the delayed posttest. The intermediate level's relative gain was 45% from the pretest to the posttest and 37% from the pretest to the delayed posttest. Three months after the end of Term 1, the intermediate-level group's scores decreased by 7%, and the high-level group's score decreased by 4% from the immediate posttest to the delayed posttest. There were no statistically significant differences between the immediate and delayed posttest scores across the three groups. In Term 2, the relative vocabulary learning gain for the high-level group was 63% between the pretest and the posttest and 51% between the pretest and the delayed posttest. In contrast, the relative gains for the low-level group were 20% between the pretest and the posttest and 19% between the pretest and the delayed posttest. The gains of the intermediate-level group sat between the high- and low-level groups. They gained 40% between the pretest and posttest and 44% between the pretest and delayed posttest in Term 2.

The impressive gains made by the high-level participants provide strong evidence that extensive reading can be an effective approach to developing L2 vocabulary knowledge. However, the differences between the gains of the high- and low-level participants indicate that large gains in vocabulary knowledge cannot be generalized to all learners in extensive reading programs. Instead, the findings suggest that there is likely to be a Matthew effect in L2 extensive reading. Merton (1968) first introduced the term *Matthew effect* in relation to cases in which greater learning results from superior initial knowledge. The results of the present study indicate that a similar phenomenon may occur among EFL extensive reading students.

Stanovich (1986) found that there is a reciprocal relationship between L1 vocabulary learning and reading ability, in which greater L1 vocabulary knowledge leads to greater amounts of reading, and this, in turn, leads to greater vocabulary learning. One difference between the present

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press

findings and Stanovich's (1986) discussion of Matthew effects with young L1 readers is that the amount of text read by the high- and low-level participants was the same in this study. In the present study, the high-level learners did not have greater amounts of repetitions of unknown and partially known words through larger amounts of print exposure than the low-level learners. Because the number of graded readers was controlled, the findings would suggest that the high-level learners had superior ability to learn the meanings of the target words.

In answer to the second research question, the results for the high-level learners revealed that the vocabulary learning gains made through reading Level 2 graded readers were consistent with those made through reading Level 1 graded readers at both retention intervals. The relative gains between the pretest and the posttest were 63% in both Term 1 and Term 2. Similarly, the relative gains between the pretest and the delayed posttest were 59% in Term 1 and 51% in Term 2. The relative gains between the pretest and the posttest for the intermediate-level group were 45% in Term 1 and 41% in Term 2. The relative gains between the pretest and the delayed posttest in Term 1 (37%) and in Term 2 (44%) were also similar. The findings for the low-level learners also showed little variation from the pretest to the posttest (28% and 20%, respectively) and from the pretest to the delayed posttest (21% and 19%, respectively) in the two terms.

The comparison between the Term 1 and the Term 2 relative gains indicates that EFL learners' L2 vocabulary growth may continue at a similar rate over different levels of text difficulty. Moreover, it suggests that reading 10 Level 1 graded readers may be sufficient for more advanced learners to move up and to effectively begin reading Level 2 graded readers. However, the difference between the higher and lower level participants' results suggests that individual differences are likely to affect the amount of vocabulary learning progress at subsequent graded reading levels. Learners who are more efficient in their vocabulary learning (as indicated by relative learning gains) will learn more words, whereas those who are less efficient may learn fewer words and fall farther behind in their lexical development. This is indicated by absolute gains. In Term 1, the higher and lower level participants learned a similar number of words (16 and 14, respectively, on the delayed posttest) because the more advanced learners already knew a lot of the words. However, in Term 2, when both of these groups did not know many of the target words, the higher level group learned almost twice as many items as the lower level group (32 and 16, respectively, on the delayed posttest). This suggests that, despite relatively large amounts of L2 reading, some learners may not be ready to progress to subsequent levels, may struggle to learn words at the next level, and may fall farther behind their peers in their lexical development.

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press

Although the high-level participants learned vocabulary at a statistically similar rate in the two terms, it is also important to note that the two sets of target items were not controlled for in difficulty, and one set may have been easier to learn than another. Factors that affect the ease with which we learn words—such as word length (Ellis & Beaton, 1993), imageability (Ellis & Beaton, 1993), and word frequency level (Lotto & De Groot, 1998)—may have had an effect. Similarly, text-specific factors such as interference (Ellis, 2006), redundancy (Ellis, 2006), overshadowing (Ellis, 2006), illustrations (Horst et al., 1998), background knowledge (Pulido, 2004), the amount of information that could be used to derive the meanings of the target words (Webb, 2008), and the frequency of occurrence of the words in the text (e.g., Horst et al., 1998; Pigada & Schmitt, 2006; Waring & Takaki, 2003; Webb, 2007) may also have had an impact on learning. It may be impossible to control all of these factors in a longitudinal study of extensive reading to gauge the efficiency of vocabulary learning. We suggest caution when interpreting the vocabulary learning efficiency of the learners over the two terms.

In answer to the third research question, the correlations between preintervention vocabulary knowledge and the scores on the initial posttests were large in both Term 1 (r = .74) and Term 2 (r = .64). The correlations between prior vocabulary knowledge and the scores on the delayed posttests were also large in Terms 1 (r = .76) and 2 (r = .62). The correlations found in this study are higher than those reported in Horst et al.'s (1998) study ($.31 \leq r \leq .36$), indicating that prior vocabulary knowledge may have a greater effect on vocabulary learning through reading multiple texts than through reading single texts.

There are two reasons why prior vocabulary knowledge and subsequent vocabulary learning through reading are closely related. First, there is a large positive correlation between L2 vocabulary size and reading comprehension (Laufer, 1992; Laufer & Ravenhorst-Kalovski, 2010; Qian, 1999, 2002). This is because texts that are made up of more known words are typically easier to understand than those that contain a smaller proportion of known words. Research on lexical coverage has shown that as the proportion of known words increases above 90%, reading comprehension is also likely to increase (Hu & Nation, 2000; Laufer, 1989; Schmitt et al., 2011). The second reason why prior vocabulary knowledge and learning vocabulary through reading are closely associated is that greater knowledge of the words in a text helps readers to successfully infer unknown words in that text (Liu & Nation, 1985). Liu and Nation (1985) found that L2 learners who knew a higher proportion of words in a written text learned more words through reading the text than those who knew fewer words. This is because greater knowledge of the words that surround an unknown word encountered in a text helps readers to derive the meanings of unknown words; it is difficult to guess the meaning of a word if you do not know the words that are included in its context.

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press

Pedagogical Implications

The findings highlight three major challenges that teachers may face when running L2 extensive reading programs. The first challenge for teachers is determining which learners are making efficient vocabulary learning gains through reading and which learners are not. The VLT scores from this study would suggest that all of the participants were ready to learn with the Stage 1 graded readers in Term 1. The overall differences in prior vocabulary knowledge as indicated by the VLT scores were relatively small. The high-, intermediate-, and low-level groups knew approximately 1,900, 1,600, and 1,400 words, respectively, a difference of only 500 words between the high- and low-level groups and 200 words between the intermediate- and low-level groups. Because the graded readers in Term 1 had 400 headwords and those in Term 2 had 700 headwords, it might be expected that all groups of learners would make similar vocabulary learning progress. However, the results clearly show that a 200-word difference in prior vocabulary knowledge can make quite a difference in learning gains. It may be necessary to measure vocabulary learning progress in an extensive reading program to ensure that students with a smaller vocabulary do not fall far behind those with a larger vocabulary. Moreover, the development of a vocabulary knowledge profile at the first three 1,000-word levels—that is, a profile similar to that of the high-level group—may provide an indication of successful vocabulary learning potential. The degree to which knowledge of the different word levels affects performance is not clear from the results, however. Words vary in their relative value; knowledge of words at the 1,000-word level is more valuable than knowledge of words at the 2,000-word level, because words at the 1,000-word level are encountered and used much more often in English. Thus, it is likely that mastery of the 1,000-word level is more important for vocabulary learning (and comprehension) than knowledge of words at the 3,000-word level, and perhaps mastery of the 1,000 level is a larger factor than overall scores. This would be a useful avenue for further research.

The second challenge for teachers is how to effectively provide sufficient support to lower level students so that they can adequately comprehend the vocabulary in graded readers and successfully infer unknown words from context. Perhaps the best way to do this would be through audio-assisted repeated reading of each graded reader in the early stages of the extensive reading program. Research has shown that reading fluency and comprehension tend to improve through repeated readings of the same text (Chang & Millett, 2013; Dowhower, 1987; Gorsuch & Taguchi, 2008; Herman, 1985), and the presence of aural versions of a text may help learners to process text in chunks, rather than through word-for-word reading, and may improve comprehension

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press

(Dowhower, 1987). Repeated readings of a text may also aid vocabulary learning in two ways. First, as comprehension improves through subsequent readings, learners may be able to divert processing resources away from the meaning of the text and to focus more attention on any unknown words. Second, the number of encounters with unknown items in the text increases with each additional reading, which in turn increases the potential to successfully infer those words (Horst & Meara, 1999). Support may also be provided through prelearning key words within graded readers prior to the first reading. Although there is limited research in this area, one study indicated that prelearning vocabulary may have a positive effect on comprehension (Webb, 2009).

Another possible approach to supporting lower level learners is to start at an easier graded reading level. The Level 1 Oxford Bookworm graded readers used in this study have 400 headwords. The Extensive Reading Foundation labels these texts as belonging to the *early elementary* stage, which is the 4th of 20 stages; the first stage is focused on learning the alphabet, and the 20th involves near-native text. Considering the lower level learners' vocabulary scores on the VLT and their reading progress across the Level 1 books, we believe that these texts were at the appropriate level for them. However, perhaps starting at an easier level could motivate later reading and, consequently, lead to greater vocabulary learning.

The third challenge for teachers is helping lower level learners effectively progress through subsequent graded reading levels. Progression through graded reading levels is particularly important for two reasons. First, the number of headwords in each graded reading level increases at subsequent levels. Thus, moving through graded reading levels provides a means for learners to repeatedly encounter and learn additional vocabulary at each new level. Second, the number of words per text tends to increase as the level of the reader increases. Because one of the key objectives of extensive reading programs is for learners to receive as much L2 input as possible, reading longer texts may help to increase the amount of L2 written input. However, it is not clear how many texts or how long a period learners need to progress through graded reading levels. Ideally, learners will move up to subsequent levels when they are ready. However, if this requires reading far more than 10 graded readers, institutions need to consider whether there will be sufficient time to progress through all levels. Nation and Wang (1999) suggested that learners read a minimum of five graded readers at each level to provide sufficient encounters with the vocabulary at that level. The results of this study indicate that the number of texts that should be read at one level may vary considerably among learners. The higher level learners continued to make impressive vocabulary learning gains when they moved from Level 1 to Level 2; however, although the lower level learners' vocabulary gains were consistent at both levels, the larger

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press

number of unknown words and the relatively small number of words learned suggests that they may not have been ready to progress. Thus, reading 10 Level 1 graded readers may be sufficient to move to Level 2 graded readers for those learners who have a head start in vocabulary learning but not for those who are behind in L2 vocabulary development. Perhaps the best solution to this problem is to begin extensive reading relatively early in the L2 learning process so that there is enough time for all learners to progress through the different levels of text difficulty. This may have the added bonus of allowing those at the upper levels to move beyond graded readers to the midfrequency readers that have recently become available (Nation, 2014).

The preceding discussion has focused on supporting reading. However, it could also be hypothesized that the problem for the lower level learners was not their ability to read but rather their ability to successfully infer words from context. Learners vary in their ability to successfully learn words in context (Sasao, 2013). Therefore, it may be useful for teachers to measure their students' skill at learning words through reading at the beginning of an extensive reading program. The Guessing from Context Test (Sasao, 2013; sample test available at http://ysasaojp.info/testen.html) is a validated and freely available test designed for this purpose. Using this test, teachers could train their students to more effectively use the guessing-from-context strategy (Clarke & Nation, 1980). Further research examining the progress of L2 learners in an extensive reading program with varying abilities to guess words in context would be a useful follow-up to the present study.

Limitations

It is important to note two limitations of this study. First, the vocabulary learning gains in this study should not be viewed as purely incidental, because there was the possibility that the participants gained knowledge of the words through postreading activities. However, the vocabulary learning gains may reflect those of extensive reading programs, which typically include some form of postreading activity.

Second, although the test used in this study provided a useful measure of knowledge of the form-meaning connections of words, it did not indicate how well these words were known. It would be useful to investigate how depth-of-vocabulary-knowledge scores progress through extensive reading by using multiple measures to assess learning. This may provide a more accurate assessment of how high-, intermediate-, and low-level learners' preintervention vocabulary knowledge affects subsequent learning.

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press

CONCLUSION

The present study sheds further light on vocabulary learning through extensive reading. The findings indicate that prior L2 vocabulary knowledge may have a large effect on the amount of vocabulary learned through extensive reading and the progress that is made at subsequent levels. Learners with a larger vocabulary size may make efficient progress in learning words through the different graded reading levels. However, the vocabulary learning gains of students with less word knowledge may be relatively small at initial graded reading levels and may fail to increase at subsequent levels. The findings indicate that it may be useful to measure students' vocabulary knowledge before they begin an extensive reading program and to either monitor the progress of those students with the smallest vocabulary sizes to ensure that they are able to cope with the lexical burden of extensive reading and/or provide support for these students.

There are very few longitudinal studies of vocabulary learning through extensive reading, so further research in this area is clearly warranted. Because the results of this study indicate that variation in prior vocabulary knowledge among learners with a similar L2 learning background may have a large effect on vocabulary learning gains made through reading, it would be useful for future studies to examine the gains across learners to provide an accurate assessment of learning. Reading studies that simply provide an overall mean learning gain may not reveal a precise representation of the learning gains of the participants.

Received 12 August 2013
Accepted 23 June 2014
Final Version Received 10 August 2014

NOTES

1. The extent to which vocabulary is learned in a longitudinal study of extensive reading in relation to the frequency of occurrence of target vocabulary within and between graded readers will be examined in a follow-up article.

2. In an earlier version of this article, a control group that did not learn through extensive reading was included to show that outside learning and a learning effect from taking the pretest did not contribute to the learning gains of the experimental groups. However, at the suggestion of one reviewer, the control group was excluded from the study because it did not score as highly on the VLT as the experimental participants and was thus not a true control group.

3. The first 1,000-word level was developed by Paul Nation at Victoria University of Wellington, New Zealand, and the Chinese translation was done by Karen Wang, Nation's former student. It should be noted that this version of the VLT has not gone through the same rigorous validation procedure as the monolingual version on which it was based.

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press

REFERENCES

Brown, R., Waring, B., & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language*, *20*, 136–163.

Chang, C.-S., & Millett, S. (2013). Improving reading rates and comprehension through timed repeated reading. *Reading in a Foreign Language*, *25*, 126–148.

Chen, C., & Truscott, J. (2010). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition. *Applied Linguistics*, *31*, 693–713.

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.

Clarke, D. F., & Nation, I. S. P. (1980). Guessing the meanings of words from context: Strategy and techniques. *System*, *8*, 211–220.

Cobb, T. (2007). Computing the vocabulary demands of L2 reading. *Language Learning and Technology*, *11*, 38–63.

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

Dowhower, S. L. (1987). Effects of repeated reading on second-grade transitional readers' fluency and comprehension. *Reading Research Quarterly*, *22*, 389–406.

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. A. (1993). Factors affecting the learning of foreign vocabulary: Imagery keywords mediators and phonological short-term memory. *Language Learning*, *43*, 559–613.

Gorsuch, G., & Taguchi, E. (2008). Repeated reading for developing reading fluency and reading comprehension: The case of EFL learners in Vietnam. *System*, *32*, 253–278.

Herman, P. A. (1985). The effect of repeated readings on reading rate, speech pauses, and word recognition accuracy. *Reading Research Quarterly*, *20*, 553–564.

Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study. *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.

Horst, M., & Meara, P. M. (1999). Test of a model for predicting second language lexical growth through reading. *Canadian Modern Language Review*, *56*, 308–328.

Hu, M., & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, *13*, 403–430.

Jenkins, J. R., Stein, M. L., & Wysocki, K. (1984). Learning vocabulary through reading. *American Educational Research Journal*, *21*, 767–787.

Koskinen, P. S., Blum, I. H., Bisson, S. A., Phillips, S. M., Creamer, T. S., & Baker, T. K. (2000). Book access, shared reading, and audio models: The effects of supporting the literacy learning of linguistically diverse students in school and at home. *Journal of Educational Psychology*, *92*, 23–36.

Laufer, B. (1989). What percentage of text lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special language: From humans thinking to thinking machines* (pp. 316–323). Clevedon, UK: Multilingual Matters.

Laufer, B. (1992). How much lexis is necessary for reading comprehension? In H. Bejoint & P. Arnaud (Eds.), *Vocabulary and applied linguistics* (pp. 126–132). Basingstoke, UK: Palgrave Macmillan.

Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, *22*, 15–30.

Liu, N., & Nation, I. S. P. (1985). Factors affecting guessing vocabulary in context. *RELC Journal*, *16*, 33–42.

Lotto, L., & De Groot, A. M. B. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Language Learning*, *48*, 31–69.

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press

Merton, R. K. (1968). The Matthew effect in science. *Science*, *159*(3810), 56–63.
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.
Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly*, *20*, 233–253.
Nation, I. S. P. (1983). Testing and teaching vocabulary. *Guidelines*, *5*, 12–25.
Nation, I.S.P. (2012). *The BNC/COCA word family lists*. Unpublished manuscript. Retrieved from www.victoria.ac.nz/lals/about/staff/paul-nation
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*(2), 1–16.
Nation, I. S. P., & Heatley, A. (2002). Range: A program for the analysis of vocabulary in texts [Software]. Retrieved from http://www.victoria.ac.nz/lals/staff/paul-nation/nation.aspx
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.
Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do *Things Fall Apart*? *Reading in a Foreign Language, 22*, 31–55.
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.
Qian, D. D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review*, *56*, 282–308.
Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*, *52*, 513–536.
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.
Sasao, Y. (2013). *Diagnostic tests of English vocabulary learning proficiency: Guessing from context and knowledge of word parts* (Unpublished doctoral thesis). Victoria University of Wellington, New Zealand.
Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Basingstoke, UK: Palgrave Macmillan.
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 Behaviour*, *22*, 71–97.
Stanovich, K. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*, 360–407.
Tekmen, E. A. F., & Daloglu, A. (2006). An investigation of incidental vocabulary acquisition in relation to learner proficiency level and word frequency. *Foreign Language Annals*, *39*, 220–243.
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. (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. A. (2009). The effects of pre-learning vocabulary on reading comprehension and writing. *Canadian Modern Language Review*, *65*, 441–470.

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press

Webb, S. A., & Chang, A. C.-S. (2012). Vocabulary learning through assisted and unassisted repeated reading. *Canadian Modern Language Review*, *68*, 267–290.
Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring vocabulary through reading: Effects of frequency and contextual richness. *Canadian Modern Language Review*, *57*, 541–572.

APPENDIX A

TERM 1 TARGET WORDS

arson, arthritis, attic, autumn, bar, beach, Bible, blood, brave, bridge, candle, castle, cemetery, centimetre, chairman, chandelier, chess, clay, clever, club, cow, crash, creature, crowd, curse, detective, diamond, director, empty, enormous, factory, ferry, field, ground, hanging, horrible, inspector, judge, kitchen, knee, lawyer, lie, lift, lighthouse, lonely, machinery, magic, map, married, mask, milkmaid, millionaire, mines, mystery, opera, pale, passage, paw, peddler, plateau, pocket, poor, pretty, princess, prison, reach, remember, rope, scream, secret, sergeant, servant, shadow, shiver, shout, skier, skin, sledge, snowstorm, soldier, stage, stairs, stick, stream, stupid, suit, tablet, tears, temperature, tent, torture, trial, uncover, veil, whisky, whisper, wine, witch, wooden

APPENDIX B

TERM 2 TARGET WORDS

adopt, alarm, amnesia, arrest, beard, behead, blade, branch, bucket, calf, canoe, carnival, cart, Catholic, chain, choice, cliff, coast, coffin, colonel, comb, committee, crazy, criminal, cross, cupboard, curl, divorce, doorman, dressing room, druggist, edge, faint, fake, firmly, flame, flat, forester, fortnight, freckles, fuel, garter, goblet, identity, imagine, insult, journalist, lightning, master, memento, miscarriage, mistress, moustache, mutineer, mutiny, mystery, negatives, nuclear, oatmeal, official, opposite, order, orphan, oval, palace, parliament, passage, pendulum, pirate, pit, plait, plant, poison, portrait, pot, pray, premature, protect, reply, rubbish, sack, safety, scanner, shore, sitting room, skeleton, slip, sob, spoil, stable, superintendent, surgery, swing, tap, tin, tomb, uranium, vicar, wheel, wreck

https://doi.org/10.1017/S0272263114000606 Published online by Cambridge University Press