Vocabulary growth in an English as a foreign language context

2 authors, including:

Hua Flora Zhong
Western Sydney University
9 PUBLICATIONS **19** CITATIONS

SEE PROFILE

Vocabulary growth in an English as a foreign language context

HUA ZHONG

DAVID HIRSH

University of Sydney

ABSTRACT

This study examines vocabulary growth in two dimensions, receptive and controlled productive, and the interrelationship between these dimensions. The study focuses on the 2000, 3000, 5000 and academic word list levels. Data were collected from 41 high school students in China. The participants completed a pre-test in the third week of the first semester, and a post-test ten weeks later. The findings indicate that: (1) both receptive and controlled productive vocabulary knowledge grow significantly at some vocabulary levels after ten weeks of study; (2) the growth is greater overall in controlled productive vocabulary knowledge than in receptive vocabulary knowledge; (3) receptive vocabulary size is larger than controlled productive vocabulary size at all vocabulary levels; and (4) the gap between receptive and controlled productive vocabulary size lessens after ten weeks of instruction. The study, in considering the findings, reflects on the relationship between classroom focus and vocabulary learning.

Address for correspondence: Hua Zhong, Faculty of Education and Social Work, The University of Sydney, NSW 2006 Australia; Email: hua.flora.zhong@gmail.com.

University of Sydney Papers in TESOL, 4, 85-113.

 ISSN: 1834-3198 (Print) & 1834-4712 (Online)

INTRODUCTION

Vocabulary size has been shown to be a strong predictor of success in reading and writing in general language proficiency as well as academic achievement (Laufer, 1997; Laufer & Goldstein, 2004; Saville-Troike, 1984). However, second language learner vocabulary size often falls short of that required to use the language effectively (see Nurweni & Read, 1999). This study examines vocabulary learning within an EFL high school classroom setting to quantify the gains that can be achieved, if any, over a period of 10 weeks.

There are an estimated 117,000 word families (a word and its common family members) in English (Goulden, Nation & Read, 1990). Most of these words occur infrequently in the language and are not identified as suitable for second language learning. Second language (L2) vocabulary research is principally interested in a core vocabulary comprising of 2000 high frequency words, academic words in the case of students preparing for tertiary studies, word levels up to 3000, 5000 and 10000 words, and technical vocabulary in the case of learning for specific purposes (Nation, 2001). The principle being followed is that words occurring frequently in the language will be of most use to language learners, providing good return for learning effort. Current estimates are that a learner would need knowledge of 6000-9000 words for effective language use (Nation, 2006). To put this in perspective, well-educated native speakers know close to 20,000 words (i.e., 1000 words learned on average for the first 20 years) (Goulden, Nation & Read, 1990), while estimates of second language learner vocabulary size often fall well short of this (as further discussed in Table 1).

There is a research interest in addressing the perceived gap between English as a Foreign Language (EFL) vocabulary size and vocabulary need. Laufer (1998) examined vocabulary development over time by comparing vocabulary sizes for EFL learners one school year apart, measuring receptive, controlled productive and free productive knowledge. She found significant receptive and controlled productive vocabulary development following 180 hours

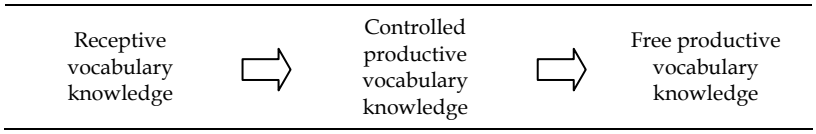
of English language class time over 1 year. More recently, Horst and Collins (2006) investigated vocabulary growth over time for one group of English as a second language (ESL) learners, in this case measuring free productive knowledge. They found no significant development in productive vocabulary use in writing following 400 hours of English language class time.

The current study builds on these two earlier studies, by investigating both receptive and productive vocabulary knowledge for one group of learners over 10 weeks. It seeks to provide teachers and researchers with an empirical account of vocabulary growth over 10 weeks representing 50 hours of English language class time in an EFL context.

Types of vocabulary knowledge

Vocabulary knowledge can be viewed as a continuum covering three dimensions, receptive, controlled productive and free productive (Laufer, 1998), as shown in Figure 1.

FIGURE 1
Dimensions of vocabulary knowledge represented as a continuum



Receptive knowledge reflects a "superficial familiarity with the word" (Laufer & Goldstein, 2004: 400), controlled productive knowledge concerns learners' use of the word when given clues or when required to use it (Laufer & Paribakht, 1998), and free productive knowledge refers to the ability to use the word correctly in free production such as a prompted writing task (Laufer, 1998).

Vocabulary knowledge of second language learners

Laufer (1992) suggests that students need to be familiar with 95% of the words in a text (i.e., 95% text coverage) for unassisted reading to

occur, and indicates that a vocabulary size of 5000 word families would be required to read texts written for native speakers. Hirsh and Nation (1992) suggest that learners need to be familiar with 97-98% of words in a text (i.e., 97-98% text coverage) for pleasurable reading to occur. Nation (2006), assuming that independent comprehension is based on knowing 98% of the words in a text, shows that knowledge of 8000 to 9000 word families is needed for comprehension of written texts such as newspapers and novels, and 6000 to 7000 for spoken texts such as lectures and movies. Webb and Rodgers (2009) suggested 5000 to 9000 word families provided 98% coverage of television programs in different genres. Depending on whether the goal is 95% or 98% text coverage, the threshold vocabulary size for using the language effectively in a broad range of contexts could range between 5000-8000 words. Previous studies regarding L2 learners' receptive vocabulary size are summarized in Table 1.

TABLE 1
Selected studies of L2 vocabulary size

Study	L2 learners	Receptive vocabulary size	Instruction time
Laufer (1998)	Israeli high school students (year 11 to year 12)	1900-3500	180 hrs (one year)
Cui and Wang (2006)	Chinese university students (1st to 4th year English majors)	3391-7199	4 years
Nurweni and Read (1999)	Indonesian university students (1st year)	1226	(not applicable)
Barrow *et al*. (1999)	Japanese university students (1st year)	2300	(not applicable)

The studies shown in Table 1 indicate that L2 vocabulary size varies between contexts. Receptive vocabulary size, as measured through test instruments, has consistently been shown to be larger

than productive vocabulary size (Aitchison, 1994; Fan, 2000; Laufer, 1998; Laufer & Goldstein, 2004; Laufer & Paribakht, 1998; Webb, 2008). Laufer (1998) reports that the controlled productive vocabulary size of Israeli high school graduates is around 2550 word families. Free productive vocabulary size can be investigated by comparing the percentage of frequent and infrequent words used in writing. Lee and Muncie (2006), investigating writing from ESL learners from different first language (L1) backgrounds, found that words from the 1000-2000 word level predominate. A similar result was found in Horst and Collins (2006), where the vocabulary from the 1000-2000 word level represented a large proportion of words in the narrative texts generated by 210 11-12-year-old francophone learners of English. Cui and Wang (2006), analyzing the writing from English major students in four different years of study, found that the higher the grade, the less 2000 high frequency words and more lower frequency words were used.

Research examining the relationship between the receptive and controlled productive dimensions of vocabulary knowledge measures differences in vocabulary size over time to indicate speed of growth. Studies show that receptive vocabulary grows faster than productive vocabulary, and the gap between receptive and productive vocabulary size becomes smaller with time (see Laufer, 1998; Laufer & Paribakht, 1998).

Measuring vocabulary size

The Vocabulary Levels Test (Nation, 1983; Schmitt *et al*., 2001) can be used to measure receptive vocabulary knowledge. It is a matching test in which words are sampled from five levels – the 2000, 3000, 5000, 10000 and academic words lists – with 30 items in six clusters at each level. The test has produced a reliability based on the Cronbach alpha figures above 0.90 (Schmitt *et al*., 2001). In the Vocabulary Levels Test, for each item, three words are required to be selected from a group of six in order to match them with a corresponding paraphrase, as shown in the following example (Schmitt *et al*., 2001: 82):

1. business	____ part of a house
2. clock	____ animal with four legs
3. horse	____ something used for writing
4. pencil	
5. shoe	
6. wall	

Controlled productive vocabulary size is measured using the Controlled Productive Vocabulary Levels Test (Laufer & Nation, 1999) which is a cued recall test that involves participants completing a word in a sentence with initial letters of target words provided as a cue or prompt. This tests vocabulary knowledge at five levels: 2000, 3000, 5000, 10000 and academic words. There are 18 items in each of the five frequency level sections in this test. An example of eliciting *bicycle* is (Laufer & Nation, 1999: 46):

He was riding a bic_____.

The Vocabulary Levels Test and Controlled Productive Vocabulary Levels Test sample words from five frequency levels. The academic word sections in both the Vocabulary Levels Test and the Controlled Productive Vocabulary Levels Test are based on items in either the University Word List (UWL) (Xue & Nation, 1984) or the Academic Word List (AWL) (Coxhead, 2000). The AWL is regarded as "the most extensive investigation to date into core academic vocabulary" (Hyland & Tse, 2007: 238), and includes words which overall provide better text coverage of academic texts than do the words in the UWL (Coxhead, 2000).

Free productive vocabulary knowledge can be measured through two methods. The Lexical Frequency Profile (Laufer & Nation, 1995) can compare high and lower frequency word use in two 300-word essays to generate a stable estimate of a learner's vocabulary size. The Lex30 (Meara & Fitzpatrick, 2000) is a word association test which provides a stimulus word and requires test takers to write as many responses to that stimulus as they can. While the Lexical Frequency Profile is suited to more advanced learners, Lex30 is

suited for learners whose language proficiency is at a lower level. Both approaches involve processing the words generated by the test takers using Range (Heatley & Nation, 1998). This program automatically recognizes words and classifies them into different word families. It reports the number of word families in the texts according to word levels such as the first 1000 frequency level, the second 1000 frequency level and the academic word list. A measure of free productive vocabulary knowledge, as a third dimension of vocabulary knowledge alongside receptive and controlled productive, was not included in the study.

Implications for the current study

The discussion so far has highlighted the concerns associated with a vocabulary gap between EFL vocabulary size and the vocabulary knowledge required to use the language effectively. It has indicated a research interest in quantifying vocabulary learning over time, and has identified the key levels of vocabulary (2000, 3000, 5000, 10000, AWL), types of vocabulary knowledge (receptive, controlled productive, free productive), and test instruments (Vocabulary Levels Test, Controlled Productive Vocabulary Levels Test, Lexical Frequency Profile, Lex30) that are commonly associated with research in this area. In doing so, the discussion suggests a theoretical basis and methodological approach for the current study, from which a methodology is developed to suit the scope and goals of the research.

THE CURRENT STUDY

The current study seeks to measure the vocabulary size and vocabulary growth of a select group of EFL students in a high school setting over 10 weeks of normal high school studies representing 50 hours of English language instruction. The focus is on measuring vocabulary size and vocabulary growth longitudinally at different levels (2000, 3000, 5000 and academic word levels) in terms of both receptive and controlled productive knowledge. The study investigates three research questions:

(1) What receptive and controlled productive vocabulary knowledge growth occurs during ten weeks of study?

(2) What is the relationship between receptive and controlled productive vocabulary knowledge?

(3) How does the relationship between receptive and controlled productive vocabulary knowledge change during ten weeks of study?

Participants and setting

Recruitment was carried out among 83 students (48 female and 35 male) in two grade 11 equivalent English language classes in a state high school in China within a single semester of study. The students attended a 40-minute class each week under the instruction of an English native speaker and six 40-minute classes each week taught by a non-native English speaker. The age of the participants was between 16 and 17 years. They had received between 7-10 years of prior English language learning. Their proficiency level ranged from intermediate to upper-intermediate. All of the participants were preparing to take the University Entrance Exam at the end of Year 12 of high school, and were planning to continue studies at university. Of the 83 students invited to participate, 64 students consented to participate, representing a 77.1% participation rate. Of the 64, 41 participants completed all the required tests and their data were included in the study, signaling a subsequent 35.9% attrition rate. The participants completed two pre-tests, and then two post-tests[1] 10 weeks later. The ten-week-period represented approximately 50 hours of English language class time.

[1] In this study, 'pre-tests' is used to refer to the first two tests conducted before the 50-hour instruction period, and 'post-tests' is used to refer to the second two tests conducted after the instruction period.

Test instruments

The 41 participants completed two sets of the tests in both pre-tests and post-tests to enable measurement of vocabulary size and vocabulary growth at four levels: 2000, 3000, 5000 and academic vocabulary. The Vocabulary Levels Test (see Zhong, 2008) was employed to measure receptive vocabulary size. The Controlled Productive Vocabulary Levels Test (see Zhong, 2008) was used to measure controlled productive vocabulary size. Taking into consideration the English language proficiency of the participants, the 10000 level was deemed largely beyond the expected vocabulary size of the participants and thus the 10000 word sections were removed from the tests used in the study. In order to maintain consistency in the word lists for this study, the academic section for Controlled Productive Vocabulary Levels Test was revised, so that the target words for this section were selected from the Academic Word List (AWL) (Coxhead, 2000) to replace University Word List (UWL) (Xue & Nation, 1984) target words. No such revision was required for the Vocabulary Levels Test version used in the study, as academic words in this test are drawn from the AWL.

The tests were administered in test conditions and 40 minutes were provided to complete the two tests. The pre- and post-tests were identical to enhance comparability between results. Participants and their teachers were not advised of this similarity between pre- and post-tests. All test papers were collected after the pre-tests, and results from the pre-tests were not communicated with either the participants or their English language teachers until the completion of the study. Such measures were designed to minimize the impact of the pre-test instruments on the language learning program for the 10-week duration. A test reliability analysis indicated high reliabilities for the four test instruments used in the study as indicated in Table 2.

TABLE 2
Reliability of test instruments

Test instrument	Test reliability value (KR-21 coefficient)
Vocabulary Levels Test (pre-test)	0.94
Controlled Productive Vocabulary Levels Test (pre-test)	0.87
Vocabulary Levels Test (post-test)	0.96
Controlled Productive Vocabulary Levels Test (post-test)	0.90
AWL section of Controlled Productive Vocabulary Levels Test (pre-test)	0.74
AWL section of Controlled Productive Vocabulary Levels Test (post-test)	0.77

The Kuder-Richardson 21 (KR-21) coefficients for the Vocabulary Levels Test and the Controlled Productive Vocabulary Levels Test were, respectively, 0.94 and 0.87 for the pre-tests; and, respectively, 0.96 and 0.90 for the post-tests. The reliability for the AWL section in the Controlled Productive Vocabulary Levels Test was also calculated due to the revision on this section. The KR-21 coefficients were 0.74 and 0.77 for the pre-test and post-test, respectively.

Pilot study

A pilot study was conducted with 2 participants using the test instruments and scoring methods proposed for the main study to investigate the scope for making comparisons between receptive, controlled productive and free productive vocabulary size. The pilot study indicated that including measurement of free productive knowledge in the study using Lex30 (Meara & Fitzpatrick, 2000) would require development of word family lists at the 3000 and 5000 word levels following rules of word family formation adopted in the development of the 2000 and academic word lists (see Bauer &

Nation, 1993). Although this was methodologically possible, development of new word lists was deemed outside the scope of the current study, and the decision was made to remove analysis of free productive vocabulary knowledge from the study, restricting the focus to receptive and controlled productive vocabulary knowledge.

Data analysis

The instrumentation and data analysis used in this study to answer each of the three research questions are indicated in Table 3.

TABLE 3
Research questions, instrumentation and data analysis

Research questions	Instrumentation	Data analysis
1. What receptive and controlled productive vocabulary knowledge growth occurs during ten weeks of study?	• Vocabulary Levels Test scores from pre-test and post-test • Controlled Productive Vocabulary Levels Test scores from pre-test and post-test	• Descriptive statistics • Inferential statistics • Mathematical calculations transferring raw scores to vocabulary size
2. What is the relationship between receptive and controlled productive vocabulary knowledge?	• Vocabulary Levels Test in pre-test • Controlled Productive Vocabulary Levels Test in pre-test	• Pearson correlational analysis • Ratio description
3. How does the relationship between receptive and controlled productive vocabulary knowledge change during ten weeks of study?	• Vocabulary Levels Test in pre- and post-test • Controlled Productive Vocabulary Levels Test in pre- and post-test	• Pearson correlational analysis • Ratio description

Descriptive statistics were used to summarize the levels of students' vocabulary knowledge achieved in the pre-test and post-test as indicated in Table 4.

TABLE 4
Descriptive Statistics (n=41)

Test instrument	Min	Max	Mean	SD	Skewness	Kurtosis
VLT 2K pre-test	17.00	30.00	26.46	4.14	-1.14	0.05
VLT 2K post-test	8.00	30.00	27.29	4.24	-2.84	10.25
VLT 3K pre-test	8.00	30.00	23.14	5.38	-0.62	-0.08
VLT 3K post-test	2.00	30.00	23.85	5.98	-1.53	3.01
VLT 5K pre-test	5.00	29.00	16.34	5.52	-0.02	-0.46
VLT 5K post-test	2.00	30.00	19.41	6.11	-0.58	0.55
VLT AWL pre-test	7.00	30.00	21.09	5.84	-0.74	-0.30
VLT AWL post-test	6.00	30.00	22.80	6.39	-0.90	-0.04
CPVLT 2K pre-test	7.00	18.00	15.29	2.83	-1.60	2.19
CPVLT 2K post-test	8.00	18.00	15.92	2.70	-1.61	1.85
CPVLT 3K pre-test	1.00	15.00	8.17	3.74	-0.11	-1.12
CPVLT 3K post-test	2.00	17.00	11.00	4.27	-0.52	-0.69
CPVLT 5K pre-test	1.00	12.00	5.36	2.64	0.56	0.24
CPVLT 5K post-test	1.00	14.00	7.48	3.05	0.13	-0.09
CPVLT AWL pre-test	1.00	15.00	8.21	3.99	-0.19	-0.64
CPVLT AWL post-test	2.00	17.00	10.24	4.09	-0.24	-0.88

Notes: VLT=Vocabulary Levels Test; CPVLT=Controlled Productive Vocabulary Levels Test; 2K=first 2000 word list; 3K=3rd 1000 word list; 5K=5th 1000 word list; AWL=Academic Word List.

The means achieved in the tests were used to compare groups and describe the amounts of a characteristic possessed by the group (Fielding & Gilbert, 2006). The standard deviation, which shows how scores are spread around the mean, is also calculated. These two values show the distribution of the clusterings of scores from the greatest number around the midpoint to the smaller number of scores towards the extremes (Gravetter & Wallnau, 2003). An estimated vocabulary size was calculated by multiplying the target

word level with the ratio between the raw score and maximum score at that level. This method is based on the assumption by Schmitt and Meara (1997) that a score out of a total score at each level indicates the proportion of words the test taker knows. For example, if one student scores 27 out of 30 at the 2000 level in the Vocabulary Levels Test, it statistically suggested that this student's vocabulary size was 1800 (27 × 2000/30). This vocabulary size would be a statistical estimation because the items in the tests were randomly selected and it is assumed that students would get the same scores whichever version of the tests they were given even though they might actually encounter more or less items they were familiar with in different versions of the test.

Owing to the small sample size and abnormal distribution of the data, non-parametric techniques were adopted for this study. Wilcoxon Signed Rank Test was used to determine whether the pre-test was significantly different from the post-test so as to answer Research Question One. It provides a z value and related significance level or *p* value whereby the former indicates the difference between the proportion of one variable in the sample and the hypothesized proportion in the population. The latter is the probability of the significant difference occurring by chance (Pallant, 2007). P value can be expressed as .05 or as .01 which are the accepted values for L2 research studies (Mackey & Gass, 2005). The current study set the significant value at 0.05 ($p < 0.05$). In other words, when *p* is less than .05, the difference between the two scores is statistically significant.

Cohen's *d* was calculated to indicate the effect size for each vocabulary level in pre- and post-test. Effect size tells the degree of the association strength between two groups of mean scores (Kirk, 1996). Cohen (1992) suggested values of 0.2= small effect size, 0.5= medium effect size, and 0.8= large effect size. The effect size indicates the percentage of overlapping positions of two groups with each other. The larger the effect size, the less overlap between the two groups (Salkind, 2008).

Correlational analysis with a Pearson correlation coefficient was used to explore the relationship between receptive and controlled productive vocabulary knowledge as proposed in Research Question Two and Research Question Three. This technique can be used to determine the strength of relations (Mackey & Gass, 2005). Pearson correlation coefficient (ranging from +1 to -1) gives information about the extent to which there is a linear relationship between two dimensions of vocabulary. A positive coefficient value indicates a positive relationship, in other words, a large-sized vocabulary in one dimension associated with a large-sized vocabulary in the other. A negative coefficient value suggests a negative relationship, i.e., a small vocabulary of one type with a large vocabulary of the other.

Ratios between receptive and controlled productive vocabulary were calculated for each level to show the gap in values for size between the two types of vocabulary knowledge.

FINDINGS

The study results are reported at different frequency levels (2000, 3000, 5000 and academic words), and as receptive or controlled productive vocabulary knowledge. While an overall figure for vocabulary size, for receptive and controlled productive, would be useful to compare with results from previous studies (e.g., Laufer, 1992, 1998), there is lack of a reliable means of making the calculation. Laufer (1992), in providing estimates for overall vocabulary size, presumably places each learner into a vocabulary size level based on the number and type of correctly used words. More recently, Laufer (1998) presents a more complicated calculation to reach an overall size, taking an average of 3rd and 5th 1000 word frequency levels as the score for 4000 level and inserting academic words into 5000 word pool. It was deemed problematic in terms of reliability, in the current study, to place learners into say the 4000 word level based on the performance at both the 3000 and 5000 levels because no account is made of the 4000 level knowledge in the test instruments. To ensure transparency in reporting findings and to present meaningful and replicable data, no calculations for estimated

overall vocabulary size have been generated from the data in the present study.

Table 5 reports on data gathered using the Vocabulary Levels Test to indicate receptive vocabulary size and growth, while Table 6 reports on data gathered using the Controlled Productive Vocabulary Levels Test to indicate controlled productive vocabulary size and growth. Table 5 and Table 6 provide information about the means and standard deviations (SD) summarized from each section in the vocabulary tests used while also presenting the comparative results obtained from the Wilcoxon Signed Ranks Test and the effect size for each level.

TABLE 5
Receptive vocabulary scores from pre-tests and post-tests using Vocabulary Levels Test

Vocabulary Frequency Level	Pre-test (n=41)		Post-test (n=41)		Difference		Effect size *(d)*
	Mean	SD	Mean	SD	z-value	*p*	
2000	26.46	4.14	27.29	4.24	2.18	.029	.20
3000	23.15	5.38	23.85	5.99	1.11	.268	.12
5000	16.34	5.53	19.41	6.12	3.87	.000	.53
AWL	21.10	5.85	22.80	6.40	3.28	.001	.28

Table 5 and Table 6 show that in both pre- and post-tests using the Vocabulary Levels Test and Controlled Productive Vocabulary Levels Test, the means decrease as the frequency level decreases from the 2000 to 5000 word levels, while the mean of the AWL section falls between the 3000 and the 5000 levels.

In Table 5 for Vocabulary Levels Test, the SDs relative to the means in all cases suggest that students' vocabulary knowledge levels tend to be homogeneous across the participant group. However, in Table 6 for Controlled Productive Vocabulary Levels Test, except for the 2000 level, the SDs are comparably larger to the relative means than in Vocabulary Levels Test, suggesting greater variability in controlled productive vocabulary size among the participant group than is evident with receptive vocabulary size. The

slight increase in the mean test scores from pre-tests to post-tests for both Table 5 and Table 6 suggests growth of vocabulary knowledge at all levels.

TABLE 6
Controlled productive vocabulary scores from pre-tests and post-tests using Controlled Productive Vocabulary Levels Test

Vocabulary Frequency Level	Pre-test (n=41)		Post-test (n=41)		Difference		Effect size *(d)*
	Mean	SD	Mean	SD	z-value	*p*	
2000	15.29	2.84	15.93	2.71	1.93	.054	.23
3000	8.17	3.75	11.00	4.27	4.03	.000	.70
5000	5.37	2.64	7.49	3.06	4.17	.000	.74
AWL	8.22	3.99	10.24	4.09	4.24	.000	.50

The Wilcoxon Signed Ranks Test shows statistically significant vocabulary growth at the 2000 level ($z = 2.18$, $p < .05$, $d = .20$), 5000 level ($z = 3.87$, $p < .05$, $d = .53$) and AWL level ($z = 3.28$, $p < .05$, $d = .28$) in receptive vocabulary knowledge. The test also shows statistically significant vocabulary growth at the 3000 level ($z = 4.03$, $p < .05$, $d = .70$), 5000 level ($z = 4.17$, $p < .05$, $d = .74$) and AWL level ($z = 4.24$, $p < .05$, $d = .50$) in controlled productive knowledge. The increase in scores is not statistically significant at the 3000 level ($z = 1.11$, $p > .05$, $d = .12$) in the Vocabulary Levels Test or in the 2000 level ($z = 1.93$, $p > .05$, $d = .23$) in the Controlled Productive Vocabulary Levels Test.

The two cases of non-significant differences may result from the performance in the pre-test. Students achieved high scores in the 2000 and 3000 levels in the pre-test, which did not leave a measurably large enough margin for improvement in the corresponding post-test levels. Therefore growth at these levels is not as obvious as at other levels. The second reason why growth in size at the 2000 level in receptive vocabulary and at 3000 level in controlled productive vocabulary is not significant may be due to the small sample size. With large samples, even very small differences between groups can become statistically significant. The sample size

(n=41) for this study is small and the increase in vocabulary size at the 3000 level in Vocabulary Levels Test and at the 2000 level in Controlled Productive Vocabulary Levels Test is minimal. As a result, the possibility of achieving statistically significant results at these two levels may be slight.

Effect size is reported based on Cohen's (1992) scale of small effect size (d = .20), moderate effect size (d = .50) and large effect size (d = .80). For the receptive test, the effect sizes were small for 3000 (d =.12), 2000 (d = .20) and AWL (d = .28) levels, and were moderate for 5000 level (d = .53). For the productive test, the effect sizes were small for 2000 (d = .23), moderate for AWL (d = .50) and large for 3000 (d = .70) and 5000 (d = .74) levels. The results indicate that there is a greater effect overall for productive than for receptive tests, there is greater effect overall within receptive vocabulary knowledge testing for the 5000 level compared to the other three levels, and that there is greater effect overall within productive vocabulary knowledge testing at the 5000 and 3000 levels, less at the AWL level, and the least at the 2000 level.

For ease of comparison, Table 7 presents students' vocabulary size according to the pre- and post-tests to show the growth in size of vocabulary over 10 weeks. Gains are presented in size and percentage whereby gains in size are the difference between the sizes measured in the pre-test and those measured in the post-test. Differences are represented as a percentage of the pre-test results.

Table 7 shows the least vocabulary growth over 10 weeks occurred at the 3000 level for receptive vocabulary knowledge (measured by Vocabulary Levels Test) and the 2000 level for controlled productive knowledge (measured by Controlled Productive Vocabulary Levels Test), while the largest growth in vocabulary size over 10 weeks occurred at the 5000 level in both receptive (18.72%) and controlled productive vocabulary (39.60%). The gain in percentage shows this growth in the Controlled Productive Vocabulary Levels Test to be twice as much as in the Vocabulary Levels Test. In Laufer's (1998) study of two groups of

learners, grade 10 and grade 11 in a high school in Israel, the growth in size after one year of instruction, as suggested by comparing data for the two groups of students, was 84% in receptive vocabulary and 50% in controlled productive vocabulary. Figures from the present study measured longitudinally for one group of learners suggest that many L2 words can be learned within a relatively short period of ten weeks corresponding to 50 hours of instruction.

TABLE 7
Receptive and controlled productive vocabulary growth at each level

	Receptive Vocabulary Size				Controlled Productive Vocabulary Size			
	2000 level	3000 level	5000 level	AWL	2000 level	3000 level	5000 level	AWL
Pre-test	1764	772	545	401	1699	454	298	260
Post-test	1819	795	647	433	1770	611	416	324
Gain in Size	55	23	102	32	71	157	118	64
Gain in %	3.12	2.98	18.72	7.98	4.18	34.58	39.60	24.62

The findings indicate that controlled productive vocabulary size at each level is smaller than that of the receptive vocabulary. This result is in line with former studies (see Fan, 2000; Laufer, 1998; Webb, 2008). It can also be concluded from Table 7 that growth of controlled productive vocabulary, in general, is larger than that of receptive vocabulary. The growth in percentage of the 3000 level and AWL in Controlled Productive Vocabulary Levels Test is over three times larger than for the Vocabulary Levels Test. This finding differs from those of previous studies. Laufer (1998), for example, comparing three dimensions of vocabulary knowledge, found that development of productive vocabulary was slower and less predictable than the development of receptive vocabulary for her participant groups. Laufer and Paribakht (1998) reached similar

findings in their study of EFL and ESL university students, offering evidence for Melka's (1997) claim that productive knowledge is at a more advanced level of the knowledge continuum, therefore often acquired later than receptive knowledge.

The findings from the current study can be explained in terms of types of learning. Receptive learning is claimed to contribute more to receptive knowledge, whereas productive learning more likely leads to increases in productive knowledge (Griffin & Harley, 1996; Waring, 1997). Webb (2005) designed a reading task (reading three glossed sentences) and a writing task (sentence production) for two experiments. The first experiment to compare the effect of receptive and productive tasks suggested that when given equal time to both tasks, receptive learning tasks may not only increase receptive knowledge but also lead to significant increases in productive knowledge. His second experiment, investigating the gains for each task from the first experiment, supports previous studies (see Griffin & Harley, 1996; Waring, 1997) in finding that productive learning was effective in promoting productive knowledge. In this regard, if participants in the current study received more productive tasks than receptive tasks during their ten weeks of instruction, this may account for the evidence of greater gains in controlled productive vocabulary size than in receptive vocabulary size.

A further finding was that a large increase was reported for the AWL for both receptive and controlled productive vocabulary size. Two reasons may account for this. One is that the main classroom supplementary textbook used in the classroom, *New Concept English* (Alexander & He, 2001), provides a good source of academic vocabulary as input for EFL high school students in China. The second reason may be that over 50% of the items tested in the academic sections of the Vocabulary Levels Test and the Controlled Productive Vocabulary Levels Test appear in the glossary of the National English Language Standard (NELS) (Ministry of Education of the People's Republic of China, 2001) which sets guidelines for selecting authorized textbooks as preparation for the national matriculation examination. Littlewood (2007) comments that test-

driven teaching is widespread in China and elsewhere in East Asia. Even though the authorized textbook used by the participants is considered self-teaching material and teachers do not use it in class as a teaching resource, the teachers of the participants in the current study may have been teaching the core vocabulary contained within the NELS glossary in an effort to better equip their learners with a core vocabulary knowledge deemed central to a strong performance in the matriculation examinations.

The Pearson correlation coefficient for receptive and controlled productive vocabulary at each level is presented in the following section. The ratio between both types of vocabulary was calculated for each level in the pre-test and post-test (productive size × 100 / receptive size). From these calculations, the larger the ratio, the smaller the gap between receptive and controlled productive vocabulary size.

TABLE 8

Correlation between receptive and controlled productive vocabulary size in the pre-test and post-test

		2000 level	3000 level	5000 level	AWL
Pre-test	*r*	0.545(**)	0.717(**)	0.561(**)	0.740(**)
	Sig. (two-tailed)	.000	.000	.000	.000
Post-test	*r*	0.640(**)	0.711(**)	0.626(**)	0.771(**)
	Sig. (two-tailed)	. 000	.000	.000	.000
N		41	41	41	41

** Correlation is significant at the 0.01 level (2-tailed).

Table 8 shows that there are moderate but significant correlations between receptive and controlled productive vocabulary at each frequency level. The strength of the relationship between these two types of vocabulary knowledge is significant at a level of 0.01 (2-tailed). Two-tailed posits a difference but in no particular direction (Salkind, 2008). The strength of the relationship at 2000 level (r = 0.545 in the pre-test and r = 0.640 in the post-test) and at 5000 level (r = 0.561 in the pre-test and r = 0.626 in the post-test) is weaker than

the relationship at 3000 level (r = 0.717 in the pre-test and r = 0.711 in the post-test) and at AWL level (r = 0.740 in the pre-test and r = 0.771 in the post-test). The correlation coefficients are all positive which means that the learners who have a large receptive vocabulary size also have a larger controlled productive vocabulary. While this result mirrored the findings in Laufer (1998) and Laufer and Paribakht (1998), the relationship between the two types of vocabulary knowledge cannot be simplistically expressed in two single scores (ratio and correlation, respectively) to predict outcomes for the entire study population because these relationships are neither uniform nor stable.

TABLE 9
Relationship between receptive and controlled productive vocabulary size in the pre-test and post-test

Vocabulary Frequency level	CP/R Ratio (%)	
	Pre-test	Post-test
2000	96.31	97.26
3000	58.81	76.86
5000	54.68	64.30
AWL	64.93	74.87

Table 9 shows shows the ratio between receptive and controlled productive vocabulary knowledge, and indicates that, apart from AWL, the ratio between receptive and controlled productive vocabulary knowledge scores decreases with the decrease in the frequency level, from the 2000 to the 5000 level. This means that the gaps between the two types of vocabulary knowledge widens with lower frequency level. Thus, if a word from the highest frequency level is known receptively, there is a high probability (96.31%) that it will also be known productively. However, there is a smaller chance (54.68%) that if a lower frequency word from the 5000 word list is known receptively, it will also be known productively. From another perspective, these results support the earlier reported finding of the study that the growth in controlled productive vocabulary is larger than in receptive vocabulary. The gap between receptive and

controlled productive vocabulary at the AWL level lies between the 2000-3000 and 5000 word levels.

CONCLUSIONS

Although limited in scope and participant size, and in the absence of an independent validation of the findings, the current study revealed that controlled productive vocabulary knowledge grew faster than receptive vocabulary knowledge during the investigation period. This finding of faster development in productive vocabulary knowledge than receptive vocabulary knowledge differs from the commonly reported pattern of receptive vocabulary knowledge growing faster than productive vocabulary knowledge (see Cui & Wang, 2006; Laufer, 1998; Laufer & Paribakht, 1998). It can be assumed that productive vocabulary size cannot always grow faster than receptive knowledge, because if so, the size of productive vocabulary would be larger than receptive knowledge, which would seem unlikely. Two questions arise from this:

(1) What is the duration of productive vocabulary knowledge growing faster than receptive vocabulary knowledge?

(2) If the focus of classroom activities can alter the pattern of vocabulary learning favouring either receptive or productive vocabulary learning, what is the extent of this effect?

The study did not examine these questions, but it is assumed that at some stage receptive vocabulary knowledge would begin to increase faster than productive vocabulary knowledge even if teachers continued to focus more on productive vocabulary knowledge in class time.

The current study also suggests that significant levels of vocabulary growth among learners can be observed in language rich EFL contexts within as little as 50 hours of English classroom instruction, which is considerably less than the 180 hours of class time accounted for in Laufer's (1998) study.

Limitations

The study employed a relatively small sample size and exhibited abnormal distribution of the data. For these two reasons, non-parametric statistical techniques were adopted in the data analysis although it is recognized that a larger study could have employed parametric techniques to yield more sensitive results. In particular, owing to the small sample size, it was not possible to apply the Kruskal-Wallis test and Mann Whitney U-test to detect the difference in vocabulary development among groups, such as for three groups of students with different vocabulary sizes. Further, this also restricted the use of multiple regression techniques to explore how much the variance in Academic Word List performance could be explained by scores on other word frequency levels such as the 2000 or 3000 word lists.

Further research

Future studies may further investigate the relative pace of productive and receptive vocabulary acquisition, and the impact on this from teacher and classroom activities. Further studies of this type may seek to include data for free productive vocabulary knowledge to explore whether growth is evident after 50 hours of English classroom instruction. Horst and Collins (2006) found no noticeable increase in free productive vocabulary knowledge in writing over 400 hours of class time. In addition, a free productive vocabulary test would allow a research team to explore vocabulary knowledge beyond the constraints of vocabulary level lists, in areas such as technical vocabulary, idiomatic expressions, and vocabulary primed by the user's knowledge of their first language.

The study suggested a possible direct link between classroom focus and type of vocabulary learning. Although data were not collected on teaching or learning practice for the study population, the researchers are aware of a central focus on productive language tasks such as writing in the classrooms which the participants attended. This may account in part for the study findings of an

overall stronger development of productive vocabulary knowledge than receptive knowledge over 10 weeks. The researchers are also aware of a strong academic focus within the syllabus, and this may account in part for the study findings of significant vocabulary growth, both receptive and productive, for academic vocabulary. The suggestion is made, thus, that the nature of vocabulary learning, both in terms of types of words learnt and whether the focus is more on receptive or productive vocabulary knowledge, is shaped in part by the focus of classroom language learning and activities. Future studies may validate the link between the classroom focus and vocabulary learning outcomes. This may help to determine whether acquisition of productive vocabulary knowledge could be better promoted with the focus primarily on productive vocabulary knowledge rather than receptive vocabulary knowledge (see Melka, 1997). A study of this type may empirically investigate the effect of classroom vocabulary focus on receptive versus productive vocabulary growth.

THE AUTHORS

Hua Zhong completed her Master of Education (TESOL) degree at the University of Sydney in 2007. She has previously taught English as a Foreign Language at high school level in China. Her current research interests include vocabulary development and acquisition. The empirical data used in this paper is based on her Master's dissertation titled *Vocabulary Size Development: A Study of Chinese High School Students*. She is planning to commence her PhD studies in 2010 focusing on vocabulary learning in an EFL setting.

David Hirsh is a lecturer in TESOL at the University of Sydney. He has taught on postgraduate TESOL and EAP programs in Thailand, New Zealand and Australia. His research focuses on vocabulary studies, language assessment, and academic acculturation. He is currently investigating scientific vocabulary use, and academic pathways. Recent publications include *Teaching Academic Writing* (University of Michigan Press, 2009, with Paltridge *et al*.) and articles in *French Review of Applied Linguistics* and *EA Journal*.

REFERENCES

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

Alexander, L.G. & He, Q.X. (2001). *New Concept English Book 2*. Beijing: Foreign Language Teaching and Research Press.

Barrow, J., Nakashimi, Y. & Ishino, H. (1999). Assessing Japanese College students' vocabulary knowledge with a self-checking familiarity survey. *System,* 27, 223-247.

Bauer, L. & Nation, P. (1993). Word families. *International Journal of Lexicogaphy,* 6, 253-279.

Cohen, J.W. (1992). A power primer. *Psychological Bulletin,* 112, 155-159.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly,* 34, 213-238.

Cui, Y.Y. & Wang, T.S. (2006). Study on development of receptive, productive vocabulary size and vocabulary knowledge depth and their relationship. *Modern Foreign Languages (Quarterly),* 29, 392-400.

Fan, M. (2000). How big is the gap and how to narrow it? An investigation into the active and passive vocabulary knowledge of L2 learners. *RELC Journal,* 31, 105-119.

Fielding, J.L. & Gilbert, G.N. (2006). *Understanding social statistics.* London: Sage.

Goulden, R., Nation, P. & Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics*, 11, 341-363.

Gravetter, F.J. & Wallnau, L.B. (2003). *Statistics for the behavioral sciences (6th Ed)*. Belmont, CA: Wadsworth.

Griffin, G.F. & Harley, T.A. (1996). List learning of second language vocabulary. *Applied Psycholinguistics,* 17, 433-460.

Heatley, A. & Nation, P. (1998). Range [Computer software]. Wellington, New Zealand: Victoria University of Wellington.

Hirsh, D. & Nation, P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language,* 8, 689-696.

Horst, M. & Collins, L. (2006). From faible to strong: How does their vocabulary grow? *The Canadian Modern Language Review,* 63, 83-106.

Howell, D.C. (2004). *Fundamental statistics for the behavioral sciences (5th ed)*. Belmont, CA: Thomson-Brooks/Cole.

Hyland, K. & Tse, P. (2007). Is there an "academic vocabulary"? *TESOL Quarterly,* 41, 235-253.

Kirk, R.E. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement,* 56, 746-759.

Laufer, B. (1992). Reading in a foreign language: How does L2 lexical knowledge interact with the reader's general academic ability? *Journal of Research in Reading,* 15, 95-103.

Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know and words you can't guess. In Coady, J. and Huckin, T. (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 20-34). Cambridge: Cambridge University Press.

Laufer, B. (1998). The development of passive and active vocabulary in a second language: Same or different? *Applied Linguistics,* 12, 255-271.

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

Laufer, B. & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics,* 16, 307-322.

Laufer, B. & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing,* 16, 33-51.

Laufer, B. & Paribakht, T.S. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning,* 48, 365-391.

Lee, S.H. & Muncie, J. (2006). From receptive to productive: Improving ESL learners' use of vocabulary in a postreading composition task. *TESOL Quarterly,* 40, 295-320.

Littlewood, W. (2007). Communicative and task-based language teaching in East Asian classrooms. *Language Teaching,* 40, 243-249.

Mackey, A. & Gass, S.M. (2005). *Second language research: Methodology and design*. Mahwah, NJ: Lawrence Erlbaum.

Meara, P. & Fitzpatrick, T. (2000). Lex30: An improved method of assessing productive vocabulary in an L2. *System,* 28, 19-30.

Melka, F. (1997). Receptive versus productive aspects of vocabulary. In Schmitt, N., and McCarthy, M. (Eds.), *Vocabulary: Description, acquisition, and pedagogy* (pp. 84-102). New York: Cambridge University Press.

Ministry of Education of the People's Republic of China (2001). *National English Language Standard*. Beijing: Beijing Normal University Press.

Nation, P. (1983). Testing and teaching vocabulary. *Guidelines,* 5, 12-25.

Nation, P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nation, P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review,* 63, 59-82.

Nurweni, A. & Read, J. (1999). The English vocabulary knowledge of Indonesian university students. *English for Specificl Purposes,* 18, 161-175.

Pallant, J. (2007). *SPSS survival manual (3rd ed)*. Maidenhead: Open University Press.

Salkind, N.J. (2008). *Statistics for people who (think they) hate statistics*. Thousand Oaks, Calif.: Sage Publications.

Saville-Troike, M. (1984). What really matters in second language learning for academic achievement? *TESOL Quarterly,* 18, 199-219.

Schmitt, N. & Meara, P. (1997). Researching vocabulary through a word knowledge framework: Word associations and verbal suffixes. *Studies in Second Language Acquisition, 19*(1), 17-36.

Schmitt, N., Schmitt, D. & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing,* 18, 55-88.

Waring, R. (1997). A comparison of the receptive and productive vocabulary sizes of some second language learners. *Immaculata (Notre Dame Seishin University, Okayama),* 1, 53-68.

Webb, S. (2005). Receptive and productive vocabulary learning: The effect of reading and writing on word knowledge. *Studies in Second Language Acquisition,* 27, 33-52.

Webb, S. (2008). Receptive and productive vocabulary sizes of L2 learners. *SSLA,* 30, 79-95.

Webb, S. & Rodgers, M.P.H. (2009). Vocabulary demands of television programs. *Language Learning*, 59(2), 335-366.

Xue, G. & Nation, P. (1984). A university word list. *Language Learning and Communication,* 3, 215-229.

Zhong, H. (2008). *Vocabulary size development: A study on Chinese high school students*. Unpublished Master of Education (TESOL) dissertation submitted to The University of Sydney, Australia.