

## ***Exploring Chinese EFL Learners' Receptive and Productive Vocabulary Knowledge: Implications for EFL Vocabulary Teaching***

**Yongyan Zheng**

*University of Hong Kong, Hong Kong SAR*

Vocabulary knowledge has attracted much attention from both EFL learners and teachers, as it consists of the building block of language learning. This study investigated 88 Chinese EFL learners' receptive and productive vocabulary knowledge at five word frequency levels, and the change of the receptive-productive relationship with the shift of word frequency and of the learners' receptive vocabulary size. The Vocabulary Levels Test and Productive Vocabulary Levels Test were administered to the participants, and test scores were collected and analyzed. The participants' knowledge of words at the 3,000, 5,000 and Academic Word Level was found to be closely related with one another and with the vocabulary knowledge as a whole, but the participants were particularly weak in word knowledge at the 5,000 level. The results also showed that the receptive-productive gap gradually narrowed with the enlargement of the participants' receptive vocabulary size, which is incongruent with what was documented in the literature. The findings suggest that the receptive-productive relationship may vary across different learning contexts, affected by the quantity and quality of the L2 input and the specific vocabulary teaching approach featuring the given context. Pedagogical implications for EFL vocabulary teaching are laid out in light of the findings.

**Key words: second language vocabulary acquisition, EFL vocabulary teaching, receptive-productive gap**

## **INTRODUCTION**

Second language (L2) learners' vocabulary knowledge constitutes the building blocks of language learning. Although in Bachman's model of communicative language ability vocabulary knowledge is only listed under the subheading of Grammatical Knowledge (Bachman & Palmer, 1996), it actually spans the whole gamut of language abilities. Textual, functional and even sociolinguistic knowledge cannot be fully established without an adequate knowledge of vocabulary. Given the prominent position of vocabulary learning in second language acquisition, many learners see learning an L2 as essentially a matter of learning new words, and thus devote a great deal of time to it (Read, 2000). However, for many learners especially in the EFL (English as a foreign language) context, despite their years of endeavor in the English classroom, insufficient vocabulary knowledge has often been named as one of the hurdles that they feel hard to overcome. This paper reports a study which investigated a group of Chinese EFL learners' receptive and productive vocabulary size, with a special focus on the relationship between their receptive and productive vocabulary knowledge across different word frequency levels. A better understanding of the issue under investigation is believed to shed informative light on L2 vocabulary uses, and more importantly, to generate pedagogical implications for L2 vocabulary teaching.

## **LITERATURE REVIEW**

Up until now, consensus is yet to be reached as to the precise definition of vocabulary knowledge. Various attempts have been made to conceptualize this construct (Chapelle, 1994, 1998; Henriksen, 1999; Nation, 1990, 2001; Schmitt & McCarthy, 1997). Nation (1990, 2001), for example, assumed a componential perspective to lexical knowledge by proposing that knowing a word means knowing its form, meaning, and use. Schmitt and McCarthy (1997), on the other hand, viewed L2 vocabulary knowledge as consisting of

the number of words known, the amount of knowledge present for each word, and how quickly this knowledge can be utilized. In a similar vein, Henriksen (1999) looked at the construct in terms of three continua: the partial-precise knowledge dimension, the depth of knowledge dimension, and the receptive-productive dimension.

Although researchers differ in their conceptualization of L2 vocabulary knowledge, they nonetheless all make a distinction between receptive and productive knowledge. Receptive vocabulary knowledge generally involves perceiving the form of a word while listening or reading and retrieving its meaning from the mental lexicon; productive vocabulary knowledge involves wanting to express a meaning through speaking or writing and retrieving and producing the appropriate spoken or written word form (Nation, 2001, p. 356).

But even though the receptive-productive distinction has long been acknowledged by researchers, the relationship between the two types of vocabulary “remains interesting but unexplored; statements about this relationship have been vague and unsubstantiated” (Laufer & Paribakht, 1998, p. 369). Two basic assumptions underlie the general discussion of the receptive-productive relationship: 1) the receptive vocabulary size is larger than the productive vocabulary size (Aitchison, 2003; Laufer, 2005; Laufer, Elder, Hill, & Congdon, 2004; Melka, 1997; Nation, 2001), and 2) in order to use a word productively, the learner needs to know at least part of the word receptively, but not vice versa. Therefore, to view receptive and productive knowledge as two dichotomous or fundamentally different kinds of associational knowledge, such as what Meara (1990) argued for, does not seem to stand to reason. Instead, the productive knowledge probably includes and extends receptive knowledge (Nation, 1990). Taking a step further, Melka (1997) proposed to treat receptive and productive vocabulary as two poles on a continuum in terms of degrees of familiarity. These degrees are numerous, even finite, and the passage from receptive to productive is anything but neat and clear.

Provided that the receptive-productive distance is perceived as degrees of

familiarity, Laufer and Nation (1999) further argued that within the bulk of productive vocabulary, degrees of familiarity extend the continuum with the controlled productive vocabulary (i.e., words available when prompted by a task as in the case of filling the blank with a certain word) on one end and the free productive vocabulary (i.e., words at one's free disposal as in the case of free composition) on the other. Ideally, productive vocabulary should refer to free productive vocabulary. But due to methodological limitations, no measure at the current stage can count the actual number of words one uses in free production, unless the person's vocabulary is really small (Laufer, 1998), so the current study takes controlled productive vocabulary to operationalize productive vocabulary knowledge.

Some basic information of the vocabulary size of native speakers (NSs) of English may provide a backdrop for the ensuing discussion of L2 learners' vocabulary knowledge. Studies showed that a five-year old child whose first language (L1) is English is expected to know 4,000 to 5,000 word families (i.e., a stem plus all related affixed forms), and a 20-year-old university level NS probably masters a total number of, according to modest estimates, 18,000 to 20,000 word families (D'Anna, Zechmeister, & Hall, 1991; Goulden, Nation, & Read, 1990). Regarding productive vocabulary use, although there has been little available research directly addressing NSs' productive vocabulary size, some studies using the Vocabulary Knowledge Scale (Wesche & Paribakht, 1996) to measure the familiarity of a given word offered some indirect evidence: For NSs, so long as they can recognize a word, they can produce it when prompted by a task (Wolter, 2001; S.-j. Zhang, 2005). This suggests that NSs' controlled productive vocabulary is roughly equal to their receptive vocabulary.

With regard to L2 learners' vocabulary knowledge, a threshold of 3,000 word families has been postulated for minimal comprehension (Laufer, 1992), and a mastery of 5,000 word families was suggested for reading for pleasure (Hirsh & Nation, 1992). When engaging in English-medium academic studies, an L2 learner needs to know at least 3,000 to 5,000 word families (Nation & Waring, 1997), and a knowledge of 4,000 to 5,000 English words

is considered a prerequisite for understanding an undergraduate economic textbook (Sutarsyah, Nation, & Kennedy, 1994). More recently, by analyzing a number of major corpuses, Nation (2006) concluded that a 8,000 to 9,000 word family vocabulary is needed for dealing with written text, and 6,000 to 7,000 word families for dealing with spoken text, provided that 98% text coverage (1 unknown word in 50) is required for learners to gain adequate comprehension.

Some other studies have focused on the actual number of words that L2 learners master (Cobb & Horst, 2000; Fan, 2000, 2001; Laufer, 1998; Laufer & Paribakht, 1998). Laufer and her colleagues (Laufer, 1998; Laufer & Paribakht, 1998) found out that the receptive vocabulary size of 10<sup>th</sup>-grade EFL learners in Israel (with 6 years of English instruction) was around 1,900 word families, and that of 11<sup>th</sup>-grade EFL learners (with 7 years of English) was around 3,500 word families<sup>1</sup>. Fan (2000, 2001) investigated the receptive vocabulary size of senior secondary school graduates in Hong Kong and found out an average size between 3,700 and 4,200 word families. A study with college-level EFL learners in China demonstrated that the receptive vocabulary size of freshmen reached 3,254 word families, and that for sophomore students reached 3,450 word families (G.-p. Zhang, Han, & Chu, 2005).

Not only is L2 learners' receptive vocabulary incomparable to their NS peers (3,500 to 4,000 word families vs. 18,000 to 20,000 word families), but their productive vocabulary is substantially poorer as well. Laufer (1998) found that the 10<sup>th</sup>-grade EFL learners in Israel had a productive vocabulary of 1,700 words families, and the 11<sup>th</sup>-grade counterparts had around 2,550 word families. Zhang et al. (2005) found a 1,665 word family productive vocabulary for freshman EFL learners and 2,100 word families for sophomores enrolled in a Chinese university. Littlewood and Liu (1996) and Chui (2006), both targeting at 1<sup>st</sup>-year university students in Hong Kong,

---

<sup>1</sup> The figures are approximations of the actual vocabulary size. Please see Laufer (1998) for the detailed procedure of calculation. The figures for Fan's (2000, 2001) and Zhang et al.'s (2005) studies were also calculated following this procedure.

found that the students exhibited good productive knowledge of words at the 2,000 level, and that of academic words were significantly larger than nonacademic ones (the 3,000 and 5,000 level). Similarly, Fan (2001) also reported an obvious drop of the Hong Kong learners' receptive vocabulary knowledge from the 3,000 to 5,000 level, and a more drastic one at the 10,000 word level. Table 1 summarizes the different vocabulary sizes of NSs and L2 learners.

**TABLE 1**  
**A Comparison of NSs and L2 Learners' Vocabulary Knowledge**

	NSs of English (word families)	L2 learners of English (word families)
Receptive vocabulary	-5-yr-old: 4,000-5,000 -20-yr-old: 18,000-20,000	-10 <sup>th</sup> -graders in Israel: 1,900 -11 <sup>th</sup> -graders in Israel: 3,500 -senior secondary graduates in HK: 3,700-4,200 -1 <sup>st</sup> -year and 2 <sup>nd</sup> -year college-level EFL learners in China: 3,254-3,450
Productive vocabulary	similar to the receptive size	-10 <sup>th</sup> -graders in Israel: 1,700 -11 <sup>th</sup> -graders in Israel: 2,550 -college freshman in HK: 2,000 -1 <sup>st</sup> -year and 2 <sup>nd</sup> -year college-level EFL learners in China: 1,665-2,100

When it comes to the relationship between L2 learners' receptive and productive knowledge, the picture becomes more complicated. Findings from various studies conducted by Laufer and her colleagues demonstrated that L2 learners' receptive vocabulary grows faster than their productive vocabulary at different word frequency levels (Laufer, 1998; Laufer & Goldstein, 2004; Laufer & Paribakht, 1998). Laufer and Paribakht (1998), for example, found that when the EFL learners' receptive vocabulary size increased, the receptive-productive gap, as measured by the P/R ratio (i.e., the percentage of receptively known words that can be used productively), also widened: the P/R ratios dropped from 91.1% (low receptive vocabulary size group) to 74% (intermediate receptive vocabulary size group) and to 71.4% (advanced receptive vocabulary size group). They interpreted the increasingly widened

receptive-productive gap as indicating that the two vocabularies develop at different rates; with receptive vocabulary grows faster, the more words one learns, the larger the gap between the two vocabularies become. Another study conducted by Laufer and Goldstein (2004) reported P/R ratios ranging from 16% at the 5,000 frequency level, 24% at the Academic word level, 30% at the 3,000 level and 35% at the 2,000 level. That is to say, the receptive-productive gap is gradually closed up as word frequency drops. But Fan's (2000) study with Hong Kong learners yielded some inconsistent findings. By closely examining the P/R ratios of nine groups of participants from different majors of study, she found that the ratios varied vastly across the groups, ranging from 53% to 81.3%. This finding prompted her to conclude that the receptive-productive gap might not be constant, and thus it was inappropriate to generalize that L2 learners can recall X% of words that they can recognize.

The above review shows the daunting task that L2 learners face if they aspire to achieve the requirement of 8,000 to 9,000 word families for adequate comprehension in reading English publications, especially when we consider the humble status of their receptive vocabulary size reported by various studies. It then becomes imperative to probe into their receptive and productive vocabulary knowledge in depth. By investigating their vocabulary knowledge at different word frequency levels, we may be able to diagnose the specific needs of the learners and provide corresponding help. Furthermore, in view of the inconsistent findings concerning the receptive-productive relationship, it is worthwhile to further inquire the issue at different word frequency levels with another group of L2 learners, such as the EFL learners in China, in order to generate better understanding of the issue under debate.

## **THE STUDY**

The purpose of the current study was two-fold: To investigate the receptive and productive vocabulary knowledge of a group of Chinese EFL learners,

and to examine the receptive-productive relationship across different word frequency levels. Three research questions were addressed:

1. What are the participants' test scores of the receptive and productive vocabulary tests at five word frequency levels and how are the scores of each frequency level related with one another?
2. How are the participants' receptive and productive vocabularies related at five word frequency levels?
3. With shifts in the participants' receptive vocabulary, what changes can be observed in their productive vocabulary and the receptive-productive relationship at five word frequency levels?

## **Participants**

The participants of the study were 88 post-secondary and tertiary-level students who were attending an English enhancement program in an English training centre in Shanghai, China during the summer of 2007. Their average age was 19.7 years old. Their L1 was unanimously Chinese Mandarin, and they all received formal English instruction in the classroom for an average of 8.2 years. They came from different schools and universities with various majors of study, and they also came from different parts of China to Shanghai for English training during the period of the study. In this sense, the participants were randomly selected and could be regarded as representative of a wide range of English vocabulary knowledge.

## **Instruments**

To measure the participants' receptive vocabulary knowledge, the study adopted the Vocabulary Levels Test (VLT) (Nation, 1990, 2001; Schmitt, 2000; Schmitt, Schmitt, & Clapham, 2001). The VLT sampled words from five frequency levels: 2,000, 3,000, 5,000, Academic Words Level (AWL), and 10,000 levels. It consisted of 10 clusters of words from each frequency



level, each cluster consisting of 6 items and 3 short definitions. The participants were required to select 3 items from groups of 6 to match the corresponding definition. Therefore, 300 items were tested with 150 short definitions to match. The maximum score for each frequency level was 30 points, and the maximum score for the VLT in total was 150 points (3 definitions each cluster \* 10 clusters each frequency level \* 5 frequency levels). Definitions were kept short in order to make sure that comprehension ability would not interfere with the test results. An example is given below:

1. accumulation	
2. edition	
3. guarantee	_1_ collecting things over time
4. media	_3_ promise to repair a broken product
5. motivation	_5_ feeling a strong reason or need to do something
6. phenomenon	

To measure the participants' productive vocabulary knowledge, the Productive Vocabulary Levels Test (PVLТ) (Laufer & Nation, 1999) was adopted. As the productive version of the VLT, the PVLТ served to tap the productive vocabulary use in a controlled manner. It took the form of sentence completing, where a meaningful sentence context was presented and the missing target word was prompted by the initial letters. Corresponding to the VLT, the PVLТ consisted of 5 frequency levels as well (2,000, 3,000, 5,000, AWL, and 10,000 levels). Each frequency level contained 18 items that were all sampled from the VLT word pool. The maximum score for each frequency level was 18 points, and the maximum score for the PVLТ in total was 90 points (18 items each level \* 5 frequency levels). The initial letters of the target items were provided to prevent the participants from filling in another word which would be semantically appropriate in the same given context but which came from a different frequency level. The following is an example eliciting the word "fragrant".

The garden was full of fra\_\_\_\_\_ flowers.

### **Procedure, Scoring and Data Analysis**

The two tests were administered within one 2.5-hour class period. The participants were given half an hour to complete the VLT first at the beginning of the class period. At the end of the class, they were given another half an hour to complete the PVL. The scoring was in terms of correct (1 point) or incorrect/blank (0 points). In particular, when scoring the participants' PVL answers, an item was considered correct when it was semantically correct. Grammatically incorrect form (e.g., past tense instead of present tense) or spelling errors that did not distort the word (e.g., *fragrent* instead of *fragrant*) were also considered correct and thus given 1 point. All the participants' scores for the two tests were collected and submitted to SPSS 14.0 for data analysis.

## **RESULTS**

### **Receptive and Productive Vocabulary Size at Different Word Frequency Levels**

The first research question intended to address the participants' receptive and productive vocabulary knowledge. The descriptive statistics (means, standard deviations, and percentages of words known) of their scores of the VLT and PVL are presented in Table 2. Following Laufer's (1998, p. 269) calculation formula, the participants have an average receptive vocabulary size of over 4,327 word families and a productive vocabulary size of over 3,433 word families<sup>2</sup>. Although the participants' performance on the VLT and PVL in the present study is slightly better compared to other EFL learners reported in published research (see Table 1), the number of words

---

<sup>2</sup> Because the words the participants knew receptively and productively at the 10,000 level were very limited in view of their test scores, the results at this level were excluded in the present calculation.

they know is still a far cry from the 8,000 to 9,000 word family requirement as suggested by Nation (2006). In order to get a clearer picture, the scores were broken down by the five word frequency levels. The results demonstrate a clear pattern of implicational scaling. That is, the more frequent the word is, the more likely that the learner knows it receptively and uses it productively. Concerning the receptive knowledge across different frequency levels, there is a moderate drop from the 3,000 to the 5,000 level (87.00%-76.33%), followed by a drastic one to the 10,000 level (76.33%-28.17%). The gap between different word frequency levels in the productive vocabulary is more obvious: The productive vocabulary size at the 3,000 level is around two thirds of the 2,000 level (52.84% vs. 86.74%), and the size at the 5,000 level is only one third of the 2,000 level (29.30% vs. 86.74%). Notably, the participants' performance at the AWL (84.85% receptively and 45.72% productively) is better than their performance at the 5,000 level respectively, although academic words are no more frequent than, if not as frequent as, words at the 5,000 level. The results are summarized in Table 2.

**TABLE 2**  
**Mean Scores and Percentages of Words Known in the VLT and PVL (N=88)**

Word levels	VLT (max=150)		PVL (max=90)	
	Mean (SD)	Percentage	Mean (SD)	Percentage
2,000	29.07 (1.15)	96.89%	15.61 (1.91)	86.74%
3,000	26.10 (3.36)	87.01%	9.51 (3.46)	52.84%
5,000	22.90 (4.40)	76.33%	5.27 (2.44)	29.30%
AWL	25.45 (3.94)	84.85%	8.41 (3.28)	46.72%
10,000	8.45 (4.78)	28.18%	1.81 (1.81)	10.04%
Total	111.98 (13.99)	74.65%	40.61 (10.58)	45.13%

*Note:* The maximum for the VLT at the 2,000, 3,000, 5,000, AWL and 10,000 levels is 30 points respectively. The maximum for the PVL at the 2,000, 3,000, 5,000, AWL and 10,000 levels is 18 points respectively.

We were also interested in the relationship of test scores between and among the five frequency levels, so Pearson correlation analysis was conducted. The results show that the participants' VLT and PVL scores at the 3,000, 5,000 and AWL levels are more closely related with one another

and with the total scores, whereas the scores at the 2,000 and 10,000 levels only bear loose relationships with the other frequency levels and with the total scores. Table 3 presents the correlations of the VLT test scores at different word frequency levels, and Table 4 presents the correlations of the PVLt test scores at different word frequency levels.

**TABLE 3**  
Correlations Between Test Scores of the VLT at Five Different Word Frequency Levels

	VLT 3,000	VLT 5,000	VLT AWL	VLT 10,000	VLT total
VLT 2,000	.369**	.412**	.165	.149	.398**
VLT 3,000		.765**	.618**	.497**	.855**
VLT 5,000			.656**	.494**	.885**
VLT AWL				.447**	.803**
VLT 10,000					.754**

Note: \*\*p<.01

**TABLE 4**  
Correlations Between Test Scores of the PVLt at Five Different Word Frequency Levels

	PVLt 3,000	PVLt 5,000	PVLt AWL	PVLt 10,000	PVLt total
PVLt 2,000	.563**	.508**	.492**	.320**	.689**
PVLt 3,000		.704**	.542**	.597**	.862**
PVLt 5,000			.764**	.583**	.890**
PVLt AWL				.555**	.848**
PVLt 10,000					.731**

Note: \*\*p<.01

### Receptive-productive Relationship

The second research question aimed at the relationship between the receptive and productive vocabulary. Since the maximum scores for the VLT (max=150) and for the PVLt (max=90) are different, the percentages of words known for each test were used to calculate the P/R ratios<sup>3</sup>. Overall,

<sup>3</sup> The P/R ratio (%) is calculated as follows: (the percentage of words known productively/the percentage of words known receptively)\*100.

only 59.83% of the participants' receptive lexicon can be accessed productively. Similar to the test scores of the VLT and PVLТ at the five word frequency levels, an implicational scaling can also be identified in the P/R ratios. Words receptively known at the 2,000 level can be used productively with relative ease (89.56%) compared to words at lower frequency levels. There is a sharp drop of P/R ratios from the 3,000 to the 5,000 level (60.08%-37.72%), but for academic words, the participants seem to be able use the receptively-known words slightly better (53.67%). In particular, it appears that the P/R ratio of words at the 10,000 level (35.71%) is almost comparable to that at the 5,000 level (37.72%), but it should not imply that the participants' ability to make use of the most infrequent words equals the ability to use words at the 5,000 level; rather, it may be because the participants' receptive vocabulary at the 10,000 level is very small in the first place.

**TABLE 5**  
**P/R Ratios at Five Word Frequency Levels**

Word levels	VLT (percentage)	PVLТ (percentage)	P/R ratio
2,000	96.89%	86.74%	89.56%
3,000	87.01%	52.84%	60.08%
5,000	76.33%	29.30%	37.72%
AWL	84.85%	46.72%	53.67%
10,000	28.18%	10.04%	36.41%
Total	74.65%	45.13%	59.83%

### **Changes with Shifts of Receptive Vocabulary Size**

The third research question was designed to see if the participants' productive vocabulary and the receptive-productive relationship change when their receptive vocabulary size increases. The participants were divided into three groups—the “low” “intermediate”, and “advanced”—on the basis of their VLT test scores.<sup>4</sup> One-way ANOVA was used to test if the differences

<sup>4</sup> All the participants were ranked according to their VLT scores in an ascending order. A rank of 1 to cases below the 33<sup>rd</sup> percentile were designated as the “low” group in terms of the receptive knowledge; cases between the 33<sup>rd</sup> and 67<sup>th</sup> percentiles were

across the three groups were statistically significant. The results of the three groups' VLT scores, PVLTS scores at different frequency levels are displayed in Table 6. In general, with the increase in receptive vocabulary knowledge, the productive knowledge grows accordingly. The three groups are significantly different in their receptive knowledge at the 3,000, 5,000 and AWL levels, but not on the 2,000 and 10,000 levels, whereas the differences among the three groups' productive knowledge are all statistically significant at each frequency level.

On the other hand, we also wanted to know if the receptive-productive gap would widen or lessen with shifts in the participants' receptive knowledge. To this end, the P/R ratios of the three groups at five word frequency levels were calculated, and one-way ANOVA was conducted to test if the differences were statistically significant.

The results are displayed in Table 7. It is shown that the P/R ratios continue to increase with the enlargement of the participants' receptive vocabulary size, which means that the receptive-productive gap gradually narrows down as the learners know more words. This result indicates that the learner's ability to use words in production develops in accordance with the growth of the number of words he/she can recognize receptively. For instance, the participants with the advanced receptive knowledge can make use of 67.30% of the receptively known words in their production, which is significantly different from those of the groups with low and intermediate level of receptive knowledge. Although the tendency of increased P/R ratios is interrupted at the 10,000 word level, the difference is not statistically significant.

---

designated as the "intermediate" group; cases above the 67<sup>th</sup> percentile was designated as the "advanced" group.

**TABLE 6**  
**One-way ANOVA of Test Scores of the Participants at the Low, Intermediate and Advanced Levels of Receptive Vocabulary Knowledge**

Word levels	VLT (percentage of words known %)				PVL (percentage of words known %)			
	Low (n=31)	Intermediate (n=28)	Advanced (n=29)	F Sig.	Low (n=31)	Intermediate (n=28)	Advanced (n=29)	F Sig.
2,000	29.03 (96.77)	28.96 (96.55)	29.21 (97.36)	.33 p= n.s.	14.71 (81.72)	15.43 (85.71)	16.76 (93.10)	10.82 p=.000
3,000	24.71 (82.58)	26.07 (86.90)	27.55 (91.84)	5.67 p=.005	7.90 (43.91)	9.32 (51.79)	11.41 (63.41)	9.24 p=.000
5,000	21.32 (71.08)	22.14 (73.81)	25.31 (84.37)	7.82 p=.001	4.13 (22.94)	4.82 (26.79)	6.93 (38.51)	13.61 p=.000
AWL	24.74 (82.47)	24.18 (80.60)	27.45 (91.49)	6.39 p=.003	6.77 (37.63)	7.82 (43.45)	10.72 (59.58)	15.29 p=.000
10,000	7.55 (25.16)	8.21 (27.38)	9.66 (32.18)	1.53 p= n.s.	1.26 (6.99)	1.75 (9.72)	2.45 (10.04)	3.43 p=.037
Total	107.42 (71.61)	109.57 (73.05)	119.17 (79.45)	6.66 p=.002	34.77 (38.64)	39.14 (43.49)	48.28 (53.64)	17.33 p=.000

Note: n.s.=not significant

**TABLE 7**  
**One-way ANOVA of the P/R Ratios of the Participants at the Low, Intermediate and Advanced Levels of Receptive Vocabulary Knowledge**

Word levels	Low (n=31)	Intermediate (n=28)	Advanced (n=29)	F	Sig.
2,000	84.63%	88.71%	95.65%	10.045	p=.000
3,000	53.07%	58.42%	69.19%	6.397	p=.003
5,000	32.40%	35.61%	45.46%	8.193	p=.001
AWL	45.54%	51.02%	64.94%	10.351	p=.000
10,000	25.42%	44.72%	40.13%	1.562	p= n.s.
Total	53.79%	58.77%	67.30%	16.912	p=.000

Note: n.s.=not significant

## DISCUSSION

The current study investigated a group of Chinese EFL learners' receptive and productive vocabulary knowledge. In particular, we sought to find more about their knowledge of and the relationship of the two vocabularies at

different word frequency levels. The clear pattern of implicational scaling in the participants' test scores for the VLT and the PVLТ as well as the P/R ratios is congruent with the results generated by other studies with different groups of EFL learners (e.g., Fan, 2000, 2001; Laufer, 1998; Laufer & Paribakht, 1998), all indicating the word frequency effects on L2 learners' mental lexicon. It is suggested that the more frequent the words are, the more likely that the words are known receptively and able to pass to the productive lexicon. According to Nation's (2006) recent analysis of several major corpora of written texts, words at the 2,000 level generally cover 86-90% of written texts. The plethora of these most frequent words in the language input greatly increases the chance of learner encounters. As the frequent words are indispensable for basic communication, and must therefore be activated repeatedly, the learners' long-term retention is accordingly reinforced. This probably gives rise to the strong receptive and productive knowledge of the most frequent words as exhibited by the participants in the study. On the other hand, as the learners have fewer chances to meet less frequent words, it is difficult for them to know these words, and even more difficult to practice them in real communication. Hence the increasingly larger receptive-productive gap as the frequency levels drop.

One point worth further discussion is the participants' weaker vocabulary knowledge at the 5,000 level and stronger knowledge at the AWL level. Supposedly, words at the AWL level are no more frequent than, if not as frequent as, words at the 5,000 level. Although a good knowledge of academic words does not ensure an equally sufficient knowledge of words at the 5,000 level, such a disparity between the test scores and receptive-productive gap at the two levels (e.g., P/R ratio at AWL: 53.67% whereas P/R ratio at 5,000: 37.72%) was not anticipated. Nevertheless, the finding can be explained from the perspective of the L2 input the participants receive in the Chinese EFL classroom. Jiang (2000) observed that the EFL context is marked by the poverty of L2 input in terms of both quantity and quality. To illustrate this point, it was roughly estimated that "more than eighteen years of classroom L2 exposure would be needed to supply the same amount of L2



input as just one year of naturalistic exposure” (Singleton, 1989, p. 236). The limited amount of exposure is compounded by the exclusive source of input. In the Chinese EFL context, the learners’ exposure to English input is restricted to classroom, and the primary source of reading and listening comes from their textbooks, the genre of which primarily belongs to academic English. However, academic words are very specialized in nature, associated particularly with academic writing: The words from the AWL only account for 1.4% of the tokens in fiction collection (Coxhead, 2000) and 4% of the words used in popular magazines or newspapers, even though they can cover 10% of the running words in an academic text book (Hwang, 1989). It follows that a good knowledge of academic words does not guarantee sufficient understanding of general texts in English, and the weak vocabulary knowledge at the 5,000 level may hinder the learners from communicating effectively in English to fulfill daily functions. This is a *de facto* obstacle frequently reported by Chinese EFL learners. They feel frustrated because they cannot adequately understand English TV programs, newspapers or novels, even if they can handle academic texts fairly well. From this perspective, the obvious disparity between the participants’ vocabulary knowledge at the 5,000 and the AWL levels can probably be attributed to the limited exposure and exclusive source of L2 input the participants receive.

Correlation analysis of the participants test scores at different word frequency levels shows that words at the 3,000, 5,000 and AWL levels are more closely connected with one another and with the total scores in both the VLT and PVLТ than are words at the 2,000 and 10,000 level. In addition, when the participants are divided into three groups on the basis of their VLT scores, group differences are only significant at the 3,000, 5,000, and AWL levels of the receptive vocabulary size. These findings suggest that a solid knowledge of words at these three frequency levels is probably more useful to learners than the knowledge of either the most frequent (i.e., the 2,000 level) or the most infrequent words (i.e., the 10,000 level). The low correlation between the 10,000 and other word levels is probably due to a sampling problem beyond the 5,000 level (Fan, 2000) and the vast divergence of

learners' knowledge of less frequent words (Nation, 2001). The relatively low correlations between knowledge of words at the 2,000 level and that at the 3,000 and 5,000 level suggest that knowing the most frequent 2,000 words in English is more like a prerequisite but does not guarantee further growth of the knowledge into less frequent words.

From a cost-benefit point of view, Laufer and Nation (1999) argued that only the words in the most frequent 2,000 level deserve individual attention, whereas teaching attention should be directed more towards vocabulary learning strategies in coping words beyond this level. However, the results of the present study indicate that the learners may need more specific help on words at the 3,000, 5,000 and AWL, particularly the 5,000 level considering their poor performance at this level. The limited amount and exclusive source of L2 input in the EFL context largely diminish the opportunity for learners to be extensively exposed to words at the 5,000 level, not to mention expecting them to pick them up via incidental learning. As a possible remedy, teachers can divert more explicit and individual attention to teaching words at the abovementioned frequency levels, among many other things.

With regard to the receptive-productive gap, the findings reflect Chinese EFL learners' weakness in production in view of the mean P/R ratio of 59.83% they achieved, especially if we compare it with the ratio of Laufer and Paribakht's (1998) study with secondary-school EFL learners in Israel (70%) and Fan's (2000) study with post-secondary Hong Kong learners (69.2%). With shifts in the participants' receptive vocabulary size, it has been found that the P/R ratios also increase (53.79%-58.77%-67.30%), suggesting a gradual narrowing of the receptive-productive gap. Put another way, the participants' productive vocabulary ability seems to develop as fast as or even faster than their receptive vocabulary, a finding inconsistent with those generated by Laufer (1998) and Laufer and Paribakht (1998). In both studies, Laufer and her colleague found that as the participants (EFL learners in Israel) enlarged their receptive vocabulary, the receptive-productive gap gradually widened. They interpreted the finding as suggesting that receptive vocabulary develops faster than productive vocabulary. However, the finding

of the current study is in accordance with that of Zhang et al.'s (2005), who also found the increase of P/R ratios along with the growth of their participants' (1<sup>st</sup>-year and 2<sup>nd</sup>-year college EFL learners in China) receptive vocabulary knowledge.

Apparently, the participants' gradually narrowing receptive-productive gap revealed in the study seems to imply that the Chinese learners' productive vocabulary progresses as fast as or even faster than their receptive one. It would further suggest that the Chinese EFL learning context is favorable for the development of productive vocabulary, which actually runs contrary to the abovementioned finding of the Chinese learners' obvious weakness in their productive vocabulary ability. Nonetheless, the apparent paradox can be explained by the characteristics of the Chinese EFL learning context *per se*. In China, the direct or explicit vocabulary teaching approach generally prevails in the classroom. For instance, Tang and Nesi (2003) found that in a secondary-school EFL classroom in Guangzhou, a southern city in China, more words were explicitly taught, but learners were exposed to far fewer word types for incidental acquisition. In this sense, Chinese EFL learners probably invest much time on explicit vocabulary learning in the classroom, which may have enriched their understanding of individual words. Since the productive use of a word requires more comprehensive and precise knowledge of the word form (Laufer, 2005; Nation, 2001), the direct learning approach is probably conducive for a single word to pass from the receptive lexicon to the productive lexicon, although the total number of words learned explicitly remains rather meager. On the other hand, due to the restricted and exclusive source of L2 input the learners are exposed to, they are deprived of, to some degree, the opportunity of encountering a wide range of new words in their milieu. As a result, the growth of their receptive vocabulary may have been slowed down. The relatively slow development and more activation of the limited amount of receptive vocabulary together lead to the finding that the receptive-productive gap seems to gradually close up as the learners' receptive lexicon expands. The above analysis highlights that the receptive-productive relationship is probably affected by the vocabulary learning/teaching

approach, as explicit versus implicit, as well as the quantity and quality of L2 input featuring the learning context.

## **IMPLICATIONS FOR EFL VOCABULARY TEACHING**

The study provides some detailed information on the vocabulary knowledge of a group of Chinese EFL learners and the receptive-productive relationship at different word frequency levels. The results show that the participants' vocabulary knowledge is rather inadequate. What is particularly weak is their knowledge of words at the 5,000 level, which hints at the likelihood that they would encounter a large proportion of unknown words when reading fiction or newspapers other than academic textbooks. More importantly, the results highlight that the relationship between receptive and productive vocabulary may vary across different learning contexts, affected by the vocabulary teaching/learning approach and the quantity and quality of L2 input featuring a given learning context. This is probably one reason why the P/R ratios found in different studies (Fan, 2000; Laufer, 1998; Laufer & Paribakht, 1998; G.-p. Zhang et al., 2005), including the study reported here, all display varying patterns.

The author has no intention to generalize the results of the present study on a group of Chinese EFL learners to learners in other Asian countries. Nonetheless, teachers engaged in EFL vocabulary teaching may find similarities between their students and the participants described in the study, and therefore find the following suggestions for EFL vocabulary teaching useful.

The first suggestion is that EFL teachers are strongly recommended to make full use of frequency lists so that they can properly and strategically select the words that are most important to enhance the learners' entire vocabulary knowledge. Horst and Cobb (2006) observed that while instructional materials regularly pay lip service to vocabulary learning, they do not bother much to dwell on the details on how vocabulary is selected or on references

to frequency lists. But findings of the study indicate that once the learners have achieved adequate knowledge of the most frequent words, particular attention should be focused onto the bulk of knowledge of words at the 3,000, 5,000 and AWL, because it is the most important words for learners to conduct academic studies in English and to fulfill daily communicative functions. As to the most infrequent words, such as the words at the 10,000 level, no special attention is suggested as they are too infrequent and bear loose relationship to the vocabulary knowledge as a whole. However, despite the understanding of what should be done in teaching vocabulary, the reality is that more unplanned than planned words were explicitly taught in the EFL classroom in China (Tang & Nesi, 2003). It is conceivable then that the learners would not be systematically introduced to the words worthy of most attention and effort. To address this situation, it is suggested that teachers work closely with researchers and make full use of the latest and reliable information about vocabulary itself, such as the frequency lists suggested by Nation (2006), so that they can properly decide the priority of learning what words and make detailed plans accordingly.

The study has revealed the participants' relative weakness in productive vocabulary in spite of the explicit learning approach prevalent in their classroom. To address this problem, EFL teachers are recommended to conduct a spectrum of activities that can maximize learners' engagement with the lexical terms, such as productive learning activities. Schmitt (2008) made the observation that merely learning words from receptive input does not seem to guarantee productive conversion. A number of recent classroom-based studies (Lee, 2003; Lee & Muncie, 2006; Mondria & Wiersma, 2004; Webb, 2005) all attested to the efficiency of productive learning of vocabulary. The particular type of learning has been found to lead to a considerable amount of receptive retention as well as to the growth of productive knowledge.

The present study has also shown that the group of Chinese EFL learners' vocabulary problem may be caused by the restricted and exclusive source of L2 input in their milieu. In this sense, it is imperative that EFL teachers,

probably not only in China but in other Asian countries, take effective measures to increase the chances for learners to encounter a wider range of words besides academic vocabulary in textbooks. Graded readers can be a good choice. For beginning- and intermediate-level EFL learners whose vocabulary knowledge is still a far cry from that of the NSs, reading authentic texts may pose too much a challenge and would even result in loss of confidence. But the vocabulary load of graded readers is fine tuned for the learner's level. For example, it was demonstrated that knowing 3,000 word families plus proper nouns was enough for one to gain adequate comprehension of the graded reader *The Picture of Dorian Gray*, the Oxford Bookworms Series at Level 3 (Nation, 2006). Empirically, a number of studies have proved graded readers' efficacy to lead to substantial vocabulary learning. For instance, Horst (2005) found that her participants gained over 50% of the unfamiliar words they encountered in the graded readers they read. Al-Homoud and Schmitt (in press), working with EFL learners in Saudi Arabia, found that the vocabulary at the 2,000, 3,000, and 5,000 word levels were greatly increased during a 10-week training period incorporating extensive reading and graded readers. As graded readers supply a host of words of similar word frequency and thus maximize the chances of encountering and recycling these words in the reading process, learners' vocabulary knowledge at the given frequency level is expected to be efficiently reinforced and consolidated. Indeed, when extensive reading of carefully selected materials is combined with explicit focus on the most important words, the EFL learners can hopefully have the best of both worlds.

## THE AUTHOR

*Yongyan Zheng* is currently a PhD candidate at Centre for Applied English Studies, University of Hong Kong. Her research interest includes second language vocabulary acquisition and vocabulary development. She is currently working on a longitudinal project tracking a group of Chinese EFL learners'

lexical progression over a year.  
Email: yyzheng@hkusua.hku.hk

## REFERENCES

- Aitchison, J. (2003). *Words in the mind: An introduction to the mental lexicon* (3rd ed.). Oxford and New York: Basil Blackwell.
- Al-Homoud, F., & Schmitt, N. (in press). *Extensive reading in a challenging environment: A comparison of extensive and intensive reading approaches in Saudi Arabia*.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.
- Chapelle, C. (1994). Are C-tests valid measures for L2 vocabulary research? *Language Testing*, 10(2), 157-187.
- Chapelle, C. (1998). Construct definition and validity inquiry in SLA research. In L. F. Bachman & A. D. Cohen (Eds.), *Interfaces between second language acquisition and language testing research* (pp. 32-70). Cambridge: Cambridge University Press.
- Chui, A. S. Y. (2006). A study of the English vocabulary knowledge of university students in Hong Kong. *Asian Journal of English Language Teaching*, 16, 1-23.
- Cobb, T., & Horst, M. (2000). Vocabulary size of some City University students [Electronic Version] from <http://www.er.uqam.ca/nobel/r21270/cv/Citysize.html>.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213-238.
- D'Anna, C. A., Zechmeister, E. B., & Hall, J. W. (1991). Toward a meaningful definition of vocabulary size. *Journal of Reading Behavior*, 23, 109-122.
- Fan, M. Y. (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(2), 105-119.
- Fan, M. Y. (2001). An investigation into the vocabulary needs of university students in Hong Kong. *Asian Journal of English Language Teaching*, 11, 69-85.
- Goulden, R., Nation, I. S. P., & Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics*, 11, 341-363.
- Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition*, 21, 303-317.
- Hirsh, D., & Nation, I. S. P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language*, 8(2), 689-696.
- Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement

- study. *Canadian Modern Language Review*, 61(3), 355-382.
- Horst, M., & Cobb, T. (2006). Editorial: Second language vocabulary acquisition. *Canadian Modern Language Review*, 63(1), 1-12.
- Hwang, K. (1989). Reading newspapers for the improvement of vocabulary and reading skills. Unpublished MA thesis, Victoria University of Wellington, New Zealand.
- Jiang, N. (2000). Lexical representation and development in a second language. *Applied Linguistics*, 21, 47-77.
- Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P. J. L. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp. 126-132). London: Macmillan.
- Laufer, B. (1998). The development of passive and active vocabulary in a second language: Same or different? *Applied Linguistics*, 19(2), 255-271.
- Laufer, B. (2005). Focus on form in second language vocabulary learning. *EUROSLA Yearbook*, 5, 223-250.
- Laufer, B., Elder, C., Hill, K., & Congdon, P. (2004). Size and strength: do we need both to measure vocabulary knowledge? *Language Testing*, 21(2), 202-226.
- Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning*, 54(3), 399-436.
- Laufer, B., & Nation, I. S. P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, 16(1), 33-51.
- Laufer, B., & Paribakht, T. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning*, 48(3), 365-391.
- Lee, S. H. (2003). ESL learners' vocabulary use in writing and the effects of explicit vocabulary instruction. *System*, 31, 537-561.
- Lee, S. H., & Muncie, J. (2006). From receptive to productive: Improving ESL learners' use of vocabulary in a postreading composition task. *TESOL Quarterly*, 40(2), 295-320.
- Littlewood, W., & Liu, N. F. (1996). *Hong Kong students and their English*. Hong Kong: Macmillan Publishers.
- Meara, P. (1990). A note on passive vocabulary. *Second Language Research*, 6, 150-154.
- Melka, F. (1997). Receptive vs. productive aspects of vocabulary. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 84-102). Cambridge: Cambridge University Press.
- Mondria, J., & Wiersma, B. (2004). Receptive, productive, and receptive + productive L2 vocabulary learning: What difference does it make? In P. Bogaards & B.



- Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition and testing* (pp. 79-100). Philadelphia, PA: John Benjamins Publishing.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. New York, NY: Heinle and Heinle.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63(1), 59-82.
- Nation, I. S. P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6-19). Cambridge: Cambridge University Press.
- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
- Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research*, 12(3), 329-363.
- Schmitt, N., & McCarthy, M. (1997). Editors' comments: Pedagogy section. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 321-326). Cambridge: Cambridge University Press.
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behavior of two new versions of the Vocabulary Levels Test. *Language Testing*, 18(1), 55-88.
- Singleton, D. (1989). *Language acquisition: The age factor*. Clevedon, Avon, England: Multilingual Matters.
- Sutarsyah, C., Nation, I. S. P., & Kennedy, G. (1994). How useful is EAP vocabulary for ESP? A corpus based case study. *RELC Journal*, 25(2), 34-50.
- Tang, E., & Nesi, H. (2003). Teaching vocabulary in two Chinese classrooms: Schoolchildren's exposure to English words in Hong Kong and Guangzhou. *Language Teaching Research*, 7(1), 65-97.
- Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in Second Language Acquisition*, 27, 33-52.
- Wesche, M., & Paribakht, T. S. (1996). Assessing second language vocabulary knowledge: Depth versus breadth. *Canadian Modern Language Review*, 53, 13-40.
- Wolter, B. (2001). Comparing the L1 and L2 mental lexicon: A depth of individual word knowledge model. *Studies in Second Language Acquisition*, 23, 41-69.
- Zhang, G.-p., Han, S.-q., & Chu, M.-l. (2005). Development of L2 passive and active

Exploring Chinese EFL Learners' Receptive and Productive Vocabulary Knowledge

vocabularies: A comparative study of English and non-English majors. *Modern Foreign Languages (Quarterly)*, 28(4), 374-382.

Zhang, S.-j. (2005). Investigating connections in L2 mental lexicon with word association tests. *Journal of PLA University of Foreign Languages*, 28(2), 52-57.