

The Effects of Word Exposure Frequency on Incidental Learning of the Depth of Vocabulary Knowledge

Feng Teng

tengfeng@uni.canberra.edu.au

Hong Kong Baptist University, China

ABSTRACT

This study examined the rate at which depth of vocabulary knowledge was learned and retained from reading a 300-headword graded reader, *The Star Zoo*. A total of 30 tertiary-level students who learn English as a foreign language (EFL) in China volunteered to take part in a reading program. Incidental learning and retention of the depth of vocabulary knowledge was measured, complementing previous research along this line. The depth of vocabulary knowledge was examined by using 36 test items within six bands of frequency (more than 20 times to only once). The target words were substituted with pseudo-words. Two tests were employed to measure receptive and productive vocabulary knowledge. This study was completed in three sessions: Students read the book for the first time and finished tests; students read the book for the second time seven days later, and finished tests; finally, retention tests were completed three months later. The findings showed that incidental vocabulary learning through reading was very limited, and this learning was largely based on a high frequency of word occurrence, that is, the higher the frequency level was, the better the learners' word-learning was. The number of times that Chinese EFL learners needed to encounter a word to recognize vocabulary knowledge was 14 times, and at least 18 times were needed for productive vocabulary knowledge. This study shows that attention to building knowledge of known words instead of solely introducing new words should be paid, and that both word exposure frequency and elaborate word processing are important as part of successful vocabulary development.

Keywords: frequency; word occurrence; graded readers; depth of vocabulary knowledge; vocabulary learning

INTRODUCTION

Reading is an interactive process in which while recalling life's experiences that relate to the passage and further probing into the deeper meaning of a text, a reader's view shifts between what he or she knows and what the text conveys. Of all the reading materials available, graded readers have been the most useful for EFL vocabulary learning as the material was specifically written which included most-frequent words (Hill, 2013). In addition, the coverage, density and repetition of vocabulary in graded readers have been statistically measured (Nation & Wang, 1999). With regard to the relationship between reading and incidental vocabulary learning, there has been much research supporting the premise that vocabulary could be learned gradually through repeated exposure to target words (e.g., Day, Omura & Hiramatsu, 1991; Horst, 2005; Webb, 2008; Webb, Newton & Chang, 2013). Although Paribakht and Wesche (1997) pointed out that the process of incidental vocabulary learning was time-consuming and inefficient, they were simply stating that there was still room for probing into the unsolved issue of incidental vocabulary learning (Schmitt, 1998). Incidental vocabulary learning refers to the acquisition of a word without the conscious intention to commit the word to memory (Hulstijn, 2013). One of the particularly worthy points addressed in researching the relationship between incidental vocabulary learning and

reading is the optimal number of times that a word should occur in a text, which is believed to affect learners' learning retention abilities (Horst, 2005; Teng, 2015a; Waring & Takaki, 2003).

Frequency of word occurrence is a continuing aspect of incidental vocabulary acquisition and a facet of representations of human memory. Numerous findings have shown that frequently occurring words in a text were easier to learn and remember (e.g., Day, Omura, & Hiramatsu, 1991; Horst, Cobb & Meara, 1998; Waring & Takaki, 2003; Teng, 2014). Rott (1999) concluded that six encounters with a target word was an adequate number. Waring and Takaki (2003) highlighted the value of higher repetition rates and pointed out that none of the word meanings that were encountered fewer than eight times would be retained after three months. Zahar, Cobb and Spada (2001) discovered that learners with low English levels needed more encounters to learn a word than learners with high English levels. The previously mentioned research provided data regarding how to determine the likelihood that a word could be learned and the number of times encountered that is preferred for the learner to acquire a word. One limitation was that typical vocabulary gains were focused by using multiple-choice tests or translation tests, and the consideration of reading treatment was restricted. To ascertain the rate at which new words are learned from reading, the aspect of the depth of vocabulary knowledge needs to be assessed.

Depth of vocabulary knowledge refers to the quality of knowing a word, which indicates that learners should have more than a superficial understanding of a word's meaning. For example, learners should know the word's pronunciation, meaning, spelling, register, morphology, and syntactic and collocational properties (Qian, 2002). Although Horst, Cobb and Meara (1998) tried to solve this issue by adding a word association test, more research on the frequency of a target word's occurrence for mastering a word from the perspective of depth of vocabulary knowledge is needed because fully understanding the aspects of vocabulary knowledge is a sign that a learner has full command of words (Nation, 2001; Waring & Nation, 2004). The current study measured the frequency of word occurrence and analysed the incidental vocabulary outcomes based upon measuring the depth of vocabulary knowledge within the context of EFL learning in China.

LITERATURE REVIEW

INCIDENTAL VOCABULARY LEARNING

Previous studies on incidental vocabulary acquisition through reading have shown how the number of times that a new word is encountered could affect learning and retention of word knowledge (Horst, 2005; Pigada & Schmitt, 2006; Rott, 2007; Teng, 2014; Waring & Takaki, 2003; Webb, 2007). For example, Waring and Takaki (2003) explored incidental vocabulary acquisition through reading the graded reader, *A Little Princess*. In their controlled study, 25 words belonging to different frequency groups were substituted with pseudo-words. Their study found small gains overall: EFL students recognized the meaning of 10.6 (42.4%), recalled the meaning of 4.6 (18.4%) and recognized the form of 15.3 (61.2%) of the 25 target words. Three months later, most of this knowledge was forgotten; learners generally remembered the meaning of only one word. They suggested that at least 20 encounters would be needed to gain full knowledge of a new word. Overall, greater gains were found for words with increased repetitions, reminiscent of earlier research in this area (e.g., Saragi, Nation & Meister, 1978; Rott, 1999).

Webb (2007) explored incidental learning of various dimensions of new word knowledge: knowledge of orthography, association, grammatical functions, syntax, meaning and form. His study revealed that all dimensions of word knowledge improved with repeated

encounters of words, confirming earlier studies which found that word exposure frequency provided significant results in incidental word learning. To gain substantial improvement in learning word knowledge, he proposed ten encounters (see also Webb & Chang, 2012, 2014). However, in previous studies (Pigada & Schmitt, 2006; Waring & Takaki, 2003), it had been suggested that learners needed to encounter the target word at least 20 times to gain full word knowledge.

Teng (2014) explored incidental learning of 30 words within five groups of frequency levels (occurring from once to more than 20 times). Three tests for measuring vocabulary gain were administered, these included word form recognition, meaning recall, and word usage. The results revealed that all word knowledge types were acquired in 68.3% of the words that occurred more than 14 times, which was a substantial improvement. His study concurred with a study by Pellicer-Sánchez and Schmitt (2010), which showed that meaning recognition required more exposures than form recognition, but contradicted Waring and Takaki's (2003) finding that meaning recognition required fewer exposures than form recognition. Teng's study also suggested that incidental word learning was affected by its part of speech. For example, verbs were easier to learn than nouns, and nouns were easier to learn than adjectives. This was in line with Zimmerman's (2009) findings.

In an attempt at gaining knowledge in the field of incidental vocabulary learning, Teng (2016) selected fifteen words that were unknown to the participants and allocated the words within three frequency groups (1 time, 5 times, and 10 times). Four types of the test (recall of form, recognition of form, recall of meaning, and recognition of meaning) were administered to measure students' ability in building the form-meaning link for a new word. Immediate posttests were administered for assessing vocabulary development, and delayed posttests were administered two weeks later for assessing retention. Results revealed that new words could be learned incidentally when reading and newly-acquired vocabulary knowledge would have decayed significantly after two weeks. The most easily acquired vocabulary knowledge was recognition of form, followed by recognition of meaning, recall of meaning, and recall of form.

The results of the aforementioned studies revealed that the number of times that a new word was encountered when reading affected learning and retention of vocabulary knowledge. However, results varied as to the precise number of word encounters. This was not surprising because incidental word learning was contingent on a number of variables. Three of these were participants' word level (Tekmen & Daloglu, 2006), context that a word is in (Webb, 2008), and part of speech (Zimmerman, 2009). As stated by Chen and Truscott (2010), "The goal of research should not be to identify a definitive number of exposures needed but rather to understand a complex process involving multiple, interacting variables" (p. 694). Accordingly, the present study aimed to deepen understanding of the complex process of incidental word learning through measuring the depth of vocabulary knowledge, to which scarce up-to-date attention has been paid.

DEPTH OF VOCABULARY KNOWLEDGE

Depth of vocabulary knowledge is regarded as the comprehension level of various aspects of a given word. Put succinctly, depth of vocabulary knowledge measures how well a learner knows an individual word or how well words are organized in a learner's mental lexicon (Nation, 2001). The conceptual framework of depth of vocabulary knowledge could be traced back to a study presented by Richards (1976), in which it was suggested that knowing a word means knowing its frequency, register, syntactic behavior, derivations, association, semantic value, and polysemy (i.e., one single word having two or more different senses). Nation (1990) added receptive and productive knowledge to the framework, and suggested form,

position, function and meaning as the four components of lexical knowledge. Qian (1998) refined the theoretical frameworks of Richards (1976) and Nation (1990) by including pronunciation, spelling, morphological properties, and syntactic properties to the depth of vocabulary knowledge. Furthermore, Qian (1999) added collocational properties. The previously noted studies suggested significance in measuring the complexity and multidimensionality of depth of vocabulary knowledge. For instance, some learners might be good at the grammatical functions of particular words while others seemed to have a stronger knowledge of English word parts (Lessard-Clouston, 2013). Thus, previous research has examined the construct of depth of vocabulary knowledge in a number of ways.

In some studies, depth of vocabulary knowledge has been related to stages or degrees on a continuum rather than extremes or ends of the continuum. For instance, Henriksen (1999) proposed three continua: partial to precise knowledge; depth of knowledge; and receptive and productive knowledge. The first word-knowledge dimension described the different levels of word knowledge. Within this continuum, breadth of vocabulary knowledge might be located closer to the partial knowledge stage while depth of vocabulary knowledge would be found at the precise knowledge stage. Researchers who adopted this approach attempted to capture stages of semantic meaning through using a vocabulary knowledge scale (Wesche & Paribakht, 1996).

Another approach was to incorporate not only semantic knowledge but also a wide range of other aspects of knowing a word, for example, paradigmatic (antonymy, synonymy, hyponymy) and syntactic characteristics (collocational). This approach has been used to specify multiple components of word knowledge (Nation, 2001).

A third approach was to measure a learner's ability to link a word to other related words. Within this context, the learners' associative behaviors were related to depth of vocabulary knowledge, and researchers attempted to measure depth of vocabulary knowledge from word associations (Read, 2004).

At present, depth of vocabulary knowledge is a critical yet under-researched dimension in the field of incidental vocabulary learning. It is rather safe to assume that there are broad stages, levels or degrees of word knowledge (Waring & Takaki, 2003). These could range from recognition of the word-form without being able to recall the meaning, to a receptive understanding of the word meaning and its various nuances and productive collocational use. In previous studies related to incidental vocabulary research, more focus has been placed on how many lexical items were acquired, and little attention has been focused on quantifying how well items become known incidentally. The lack of sensitive assessment in measuring depth of vocabulary knowledge in the research of incidental vocabulary knowledge is surprising. As learning gains from incidental learning have been found to be small, even significantly smaller than those from intentional learning (Teng, 2015b), the measurement approach should serve particularly well in revealing the smallest increments in learning (van Zeeland & Schmitt, 2013).

RESEARCH QUESTIONS

This study attempted to provide a more complete picture of incidental vocabulary acquisition from reading through measuring the depth of vocabulary knowledge. The following research questions were explored:

1. How well are new words acquired from reading a graded reader and retained over time?
2. Are words that appear frequently in the selected text more likely to be known than words with relatively fewer frequency levels?
3. At what occurrence rate are the newly-acquired words forgotten more easily?

4. To what extent are the results yielded from second-time reading different from first-time reading results?
5. To what extent are different gain scores produced by different test formats?

METHOD

PARTICIPANTS

The participants were 30 English major university students, ranging in age from 19-21 years old. They were selected from two parallel classes at a university in China. The participants were Chinese-native speakers, and had studied English for an average of eight years. Of the participants, 15 were female and 15 were male, and none had studied or worked in a country where English was the official language.

According to an internal reference at the school, they had passed Level 4 of the College English Test (CET) during the first semester. CET is a reliable standardized test of English proficiency in China. In order to pass this test, a student needs to know at least 4,000 words to read and understand materials with academic contents (CET, 2005). Therefore, the participants of the study could be described as relatively proficient learners of English.

READING MATERIALS

Many previous studies have included authentic novels as reading materials (Currie, 1997; Pellicer-Sánchez & Schmitt, 2010), offering a genuine context for motivating students. Other studies have included simplified materials, for example, graded readers (Horst, 2005; Teng, 2014; Waring & Takaki, 2003). For the present study, simplified material was preferred because simplification is a major technique used in language instruction. Nation (2007) believes that without simplification, the four strands—meaning-focused input, meaning-focused output, learning language items, and fluency development—become impossible for most EFL learners. Moreover, the measurement on depth of vocabulary knowledge is a complex process, for which simplified reading materials are regarded more suitable (Read, 2004). Graded readers, which offer effective ways of encountering most high-frequency words are the most accessible simplified texts (Nation & Wang, 1999), because they can provide students with opportunities to read a book without having to contend with so many difficult words.

The problem was how to choose an appropriate graded reader that met learners' current level. One method was to select certain words to test whether learners had already acquired them; if the learners have not acquired the words, it meant that they had a low coverage-level of known words in the reading materials. The disadvantage of using this method was that it highlighted the target words, which might influence the results of the study. Even if the target words had already been acquired by the learners, it was very difficult to validate whether other, non-test words, had also been acquired by the learners, which indicated that this was an unscientific method for choosing reading materials.

RANGE—a computer program that can provide a lexical profile of texts—was used to analyse various graded readers (Heatley, Nation, & Coxhead, 2002). Ten graded readers were typed into this program; a list of various word types and tokens of these books was generated. After a careful examination of the data generated by *RANGE*, *The Star Zoo*, a level 3 graded reader from the Bookworms series published by Oxford University Press, was selected. This book was written with 300 headwords and is a fictional story about a girl living on a new planet without plants and animals. A fictional story was preferred because it would not require background knowledge in contrast with academic texts and might provide favourable

conditions for learning vocabulary through reading (Nation, 2001). According to the results of *RANGE*, the book had 9,107 running words (tokens), 1,001 word types, and 601 word families. 86.16 % of the words were in the first 1,000-word list, 7.48 % of the words were in the second 1,000-word list, and 1.28 % of the words were in the third 1,000-word list. 5.07 % of the words were not in any of the lists. Most of the words not in the lists were inflected forms, for example, 'kissed' and 'families'. These words actually belonged to the first 1,000-word list. Since the participants in the present study had passed CET 4, it could also be reasonably assumed that the surrounding context for the target words was known to the learners, and based on testing the unknown target words, the frequency rate needed for acquiring the words could be measured.

TARGET WORDS

The factor regarding coverage of known words needed to be considered because this would determine whether a learner could achieve successful comprehension or not. For reasonable reading comprehension, it has been suggested that 95% lexical coverage of known words was needed (Hu & Nation, 2000; Laufer, 1989). In addition, the frequency of word occurrence was a key to determining which target words should be selected. The frequency of the occurrence for each word in the book was analysed by using a frequency-measuring application on the *Compleat Lexical Tutor* website (Cobb, n.d.).

To reach the desired coverage of known words, unknown words would need to be less than 5% of the total words. Several words could then be selected from this book and tested after participants completed reading it. One challenge was how to know whether the selected words were indeed unknown prior to reading. Although a pre-treatment intervention test of the chosen words could be administered, conducting a test like this would highlight the selected words for the participants, which might compromise the study.

The preferred alternative approach was to use a graded reader which would present no great problems lexically (e.g., a book with 300 headwords in the present study), and replace some of the words in the text with pseudo-words. In this environment, pseudo-words represented the set of new words to be learned and were used to test incidental learning of new words. In this manner, based on participants' word level, it could be concluded that the content surrounding the test items would be known and the rate of acquisition could be observed based solely on the test words.

Although it has been suggested to replace selected words with synonyms, Waring and Takaki (2003) asserted that this approach was not feasible because not all words have synonyms, and therefore, an additional test would be needed to evaluate whether learners knew the synonyms, which would also highlight the target words. Therefore, they suggested changing the spelling of target words, e.g., *yes* was substituted with *yoot*, a pseudo-word. The current study also used this method to replace the target words in the text. A pseudo-word is invented based on the orthographic and morphological rules of the target language (Pulido, 2007). The pseudo-words should look like plausible English words. The advantage of using this method is that participants' knowledge of the target words could be controlled. For example, none of the participants would know these target words. Additionally, the target words would not be encountered in their normal studies after the treatment session, which ensured the reliability for testing retention. Details regarding the target words are presented in Table 1.

TABLE 1. Target Words with Different Bands of Frequency Level and their Substituted Pseudo-words

Original word	Substituted word	Frequency rate	Frequency band
warm	uooky	1	1 time
young	doonful	1	
happy	tuidy	1	
true	daccit	1	
fine	swital	1	
sudden	zettive	1	
strong	trettive	4	4-6 times
difficult	Torkult	5	
envious	buikous	6	
angry	darrous	4	
wicked	telam	6	
pregnant	tremal	4	
interested	tiffle	8	8-10 times
stupid	depid	10	
wrong	wilge	9	
slow	Shring	8	
beautiful	beildful	8	
bloody	treful	10	
kind	daberous	16	14-16 times
early	brecer	16	
dark	toildy	14	
dead	tuddy	16	
unhappy	perister	15	
little	tornet	16	
good	tartful	20	18-20 times
pleased	tielty	20	
easy	zarous	18	
quiet	teetous	19	
same	Mooming	20	
important	daltant	20	
small	Dawmal	21	21-22 times
clever	tawdy	22	
friendly	tufful	22	
lucky	perful	21	
lovely	brillous	22	
excited	blaous	21	

As shown in Table 1, a total of 36 words were chosen. The reasons for selecting these words were as follows: First, to answer the second research question, words with differing frequencies of occurrence needed to be selected. Second, the types of words needed to be taken into consideration. Adjectives were chosen for this study because they are generally easier to choose than adverbs. Nouns were not selected because there were enough nouns with differing frequencies of occurrence. Verbs were not chosen because they appear with their inflections, which might lead to difficulty in determining whether the word has occurred already and to determine how frequently the word has occurred.

The 36 words were placed into six bands, with each band including 6 words with frequencies of 1 time, 4–6 times, 8–10 times, 14–16 times, 18–20 times, and 21-22 times. The total number of occurrences of these 36 words accounts for 3.6 % coverage of the running words (9,107). It was assumed that the participants knew all of the other words in the book, as the participants had already passed CET 4. Thus, it was reasonable to presume that they had at least achieved a 95% lexical coverage of known words when reading this graded reader. This was the best supposition that could be made without asking all of the participants to underline every unknown word in a pre-testing session.

INSTRUMENTS

The measurement tools used in the present study were adapted from the Word Associates Test (WAT) developed by Read (1993; 2004). This test format was based on three relationships: paradigmatic (meaning), syntagmatic (collocation), and lexical progression (a process of building words). This test has been applied in a number of previous studies focusing on exploring the depth of vocabulary knowledge (Nassaji, 2004; Qian, 2002). Certain adjustments were made in the present study (see below, and more examples were provided in Appendix).

The first dependent measure was developed to test whether the learners could produce a word that is collocated or synonymous with the target word (productive test). This test measured depth of productive vocabulary knowledge.

Using the word *uooky* (warm) as an example:

Uooky _____ (collocation)

Uooky _____ (synonym)

The participants were at the 4,000-word level and were likely to have figured out the meaning of *uooky* based on its surrounding context in the text. The contextual cues would be enough to lead them to devise an acceptable collocate or synonym. Some of the answers that would be considered acceptable for the word *uooky* include: weather (collocation), person (collocation), balmy (synonym), and friendly (synonym).

The second dependent measure was a multiple-choice test. As a multiple-choice test format, this test measured depth of receptive vocabulary knowledge. The measurement tool tested whether the learners could identify the collocation, synonymous relationship between a stimulus word (adjective) and eight options. The eight options were put into two groups, with six being distracters, distributed in either the left or right box. For example:

Uooky (warm)

mild cool cruel cold

glass clothes desk boat

There were two noteworthy issues for the four tests. First, the two tests were printed and conducted separately. Second, the sequence for taking them was: productive tests, followed by receptive tests. This was to avoid the possibility of the receptive test providing hints to the productive test.

SCORING SYSTEM

Two experienced raters were invited to score the tests. The maximum possible score for either Test 1 or Test 2 was 72 points. For Test 1, participants received one point for writing an acceptable answer and zero points for an unacceptable answer. Before scoring, the two raters held a joint meeting to reach a consensus on possible acceptable answers. Following this, they rated the test independently. In case of differences between their opinions, a third, independently trained rater would be called upon. In this situation, the score of the controversial items would be determined by the majority opinion. In terms of the first test, there were 95 different opinions out of the total 2,160 responses. Thus, there was a 96% inter-rater agreement. The discrepancies were resolved with the help of a third rater. For the second test, participants dichotomously received one point for choosing a correct option and zero points for an incorrect answer. The inter-rater agreement was 100 percent.

PROCEDURES

The author prepared and printed the book with 36 substituted pseudo-words. Participants were told that if they encountered some unknown words, they should try to read the book naturally. The present study was completed in three sessions. In session one, the students were required to read the book in one hour. The students were then required to take the tests immediately after reading the book. The tests were completed without having the book at their disposal. The average time for the two tests was 30 minutes. The time for reading and completing tests were suggested based on a pilot study with other students of similar language proficiency and background. In session two (seven days after the test administration), the students read the book again and likewise, took the tests again. The aim of this was to measure the effects of repeated reading. Three months after the reading, for the third session, the test was administered again. During the three months, the students were not provided any opportunities to read the book. It was also unlikely for the students to be exposed to the target words. The aim of this was to measure the learners' retention ability. The delayed tests were identical to the previous tests, except for the order in which the target words were presented. This was to preempt ordering effects. This could also avoid difficulties in developing equal-level tests for measuring retention. During the three sessions, the participants were not informed that there would be tests for them. They were only informed that they would need to attend a reading program. Consequently, the present study could be defined as incidental learning (Hulstijn, 2013).

RESULTS

The overall results are shown and presented in Table 2 and Figure 1.

TABLE 2. The Distribution of Mean Scores of the Two Tests

	First-time reading		Second-time reading		After three months	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Productive tests	19.90	1.39	28.32	1.61	1.30	1.81
Receptive tests	28.02	1.17	39.06	1.12	5.79	1.13

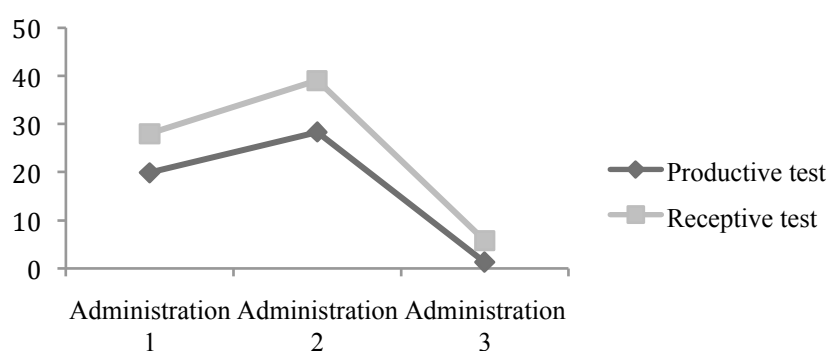


FIGURE 1. The Distribution of Mean Scores of the Two Tests

Table 2 and Figure 1 display the overall development of the mean scores that the learners achieved at different administered times (maximum score=72). For the productive test, the learners achieved a mean score of 19.90 after reading for the first time. The mean score increased to 28.32 after reading for the second time, and then decreased to 1.25 three months after the reading program. For the receptive test, the mean score increased from 28.02 after first-time reading to 39.06 after second-time reading, but decreased to 5.79 three months after

the study. In addition, the scores collected from the receptive test seemed to be higher than those collected from the productive test at each administered time.

Wilcoxon signed-rank tests were applied to measure whether there was a significant difference between results from the two tests because data was not distributed normally. Results are shown in Table 3.

TABLE 3. Wilcoxon signed-rank Results of the Two Tests Administered at Different Times

First-time reading	Second-time reading	After three months
Z = -3.93	Z = -4.51	Z = -3.67
P = .000	P = .000	P = .000

The Wilcoxon signed-rank tests revealed that the two administered tests produced significantly different results immediately after first-time reading ($P < .001$), second-time reading ($P < .001$), and three months after reading program ($P < .001$). This was evidence that learners acquired receptive knowledge before productive knowledge.

The results from the words with different frequencies of occurrence for the productive test are presented in Table 4 and Figure 2.

TABLE 4. The Mean Scores of Words with Different Frequency of Occurrence for the Productive Test

	1 time	4-6 times	8-10 times	14-16 times	18-20 times	21-22 times
First-time reading	0.4	1.6	2.4	3.8	5.6	7.4
Second-time reading	1.9	3.0	4	5.6	7.1	8.8
After three months	0	0	0.1	0.2	0.4	0.6

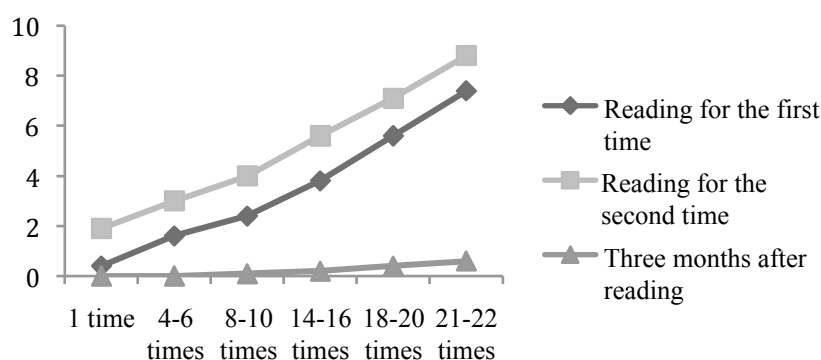


FIGURE 2. The Distribution of Mean Scores of Words with Different Frequency of Occurrence for the Productive Test

Table 4 and figure 2 show that the overall mean scores for the words with different occurrence rates after taking the productive test. The maximum score for each frequency band was 12 points. It was readily observed that the more frequently the words occurred, the better the performance in vocabulary learning demonstrated by the learners. Regarding the comparisons of the three administrations, the second-time reading was found to yield better results than the first-time reading, and a significantly decreased result was shown three months after the reading program regardless of the words' frequency level. In other words, if the learners did not receive any opportunities for reading the book for three months, the vocabulary learning outcome was lower than the results after reading for the first time.

The results of the words with different frequencies of occurrence for the receptive test are presented in Table 5 and Figure 3.

TABLE 5. The Mean Scores of Words with Different Frequency of Occurrence for the Receptive Test

	1 time	4-6 times	8-10 times	14-16 times	18-20 times	21-22 times
First-time reading	1.9	3.1	4.6	5.8	7.2	8.9
Second-time reading	3.9	4.8	7.2	8.6	9.1	10.2
After three months	0	0	0.1	0.8	1.9	2.9

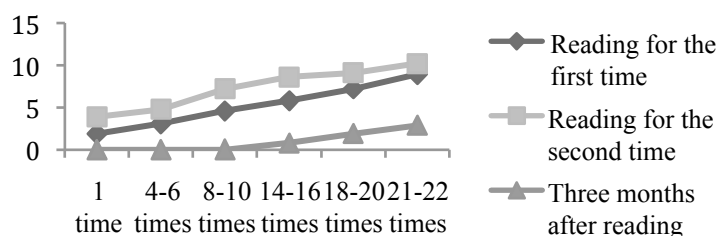


FIGURE 3. The Mean Scores of Words with Different Frequency of Occurrence for the Receptive Test

Table 5 and Figure 3 show the overall mean scores of the words with different occurrence rates after taking the receptive test. The maximum score for each frequency band was also 12 points. This test was for judging whether the learners had acquired deep knowledge of receptive word knowledge. The results demonstrated that the more frequently the words occurred, the better performance the learners attained. A maximum score of 10.2 was achieved when reading for the second time. Although this did not demonstrate a full mastery of depth of vocabulary knowledge, these results were satisfactory considering that the present study was based on incidental learning. The learners were found to have lower results for the test administered after three months regardless of frequency level.

Friedman tests were applied to compare mean gains achieved in different frequency bands. Results revealed significant differences across the six frequency levels in the productive test ($df=2, \chi^2=73.57, P<.001$) and receptive test ($df=2, \chi^2=62.35, p<.001$). This was evidence that the effect of word exposure frequency was consistent in the two tests. This also suggested that the higher the frequency level was, the better the learners' vocabulary gains were.

DISCUSSION

In this section, detailed discussions will be presented following the overall results.

OVERALL RESULTS

In answer to research question 1 regarding whether learners' vocabulary performances had increased after this reading program, the results disclosed that the learners advanced in learning vocabulary from taking this program. The results also revealed that, although learners did learn new words, the number of words that they learned was very limited and the learning of most of the new words was the result of repeated exposure to target words or higher frequencies of occurrence. The exact number of words that were incidentally learned was not calculated in the current study. To answer research question 2 regarding whether learners learned words with higher occurrence rates better, the results showed that the higher the frequency of the word occurrence rate, the better results that the learners achieved. In answer to research question 3 regarding at which occurrence rate the words would be forgotten more easily, the overall conclusion was that more than half of the words in the receptive test that occurred fewer than 14 times were easily forgotten, and that more than half of the total number of words that occurred fewer than 18 times were forgotten for productive word knowledge. In response to research question 4 regarding whether second-time reading

yielded better results than first-time reading, the analysis revealed that the second time reading improved learners' scores for both tests. In response to research question 5 regarding whether the two tests yielded different results, the findings suggested that the two tests showed significantly different results at differently administered times.

RESULTS OF VOCABULARY TEST

The first test required learners to produce a synonym or a collocation for each target word they had encountered while reading. The 36 words chosen for this test were the items with different frequencies of word occurrence in the book. After the second-time reading, the learners' vocabulary performance increased significantly from 19.90 to 28.32, but the result decreased to 1.30 when they did not read the book for three months (Table 2). This is evidence that reading a graded reader with 95% known words did not greatly facilitate learners in maintaining an initial learning of productive word knowledge. This highlighted the difficulties and confusion involved in acquiring productive word knowledge. This also suggested that receptively-known items have no incoming links from the lexicon, and cannot be recalled unless activated by some outside stimulus. The relatively slow development of productive vocabulary together led to the finding that there was a need to arrange a spectrum of activities that could maximize learners' engagement with the lexical terms, such as productive learning activities. Mere exposure to words from receptive input does not seem to guarantee productive conversion (Schmitt, 2008). The efficiency of productive learning of vocabulary has also been suggested in other studies (e.g., Lee, 2003; Mondria & Wiersma, 2004).

The second test required the learners to identify the collocation, synonymous relationship between a stimulus word (adjective) and eight options. The learners' vocabulary knowledge improved from 28.02 (38.9% of the total test items) after first-time reading to 39.06 (54.2% of the total test items) after second-time reading. The learners' performance dropped to 5.79 (8% of the total test items), which was significantly lower than the mean score of 28.02, the result after first-time reading (Table 2). The results of the second test also showed that learners' improvement in recognizing the depth of vocabulary knowledge largely depended on repeated reading and exposures to target words.

The two test results at differently administered times were significantly different. The data from the test results also indicated that when they completed second-time reading, they achieved better performance in learning the depth of vocabulary knowledge. The learners showed more ease in recognizing word knowledge than producing word knowledge. In addition to this, learners faced a critical threshold of word occurrence in learning depth of vocabulary knowledge. In other words, if learners did not encounter a word enough times, they would forget the word easily, especially when they had no chance for further reading of the target materials. This section of the results falls in line with Baddeley's (1990) research in human memory. The results also conformed to a previous study conducted by Waring and Takaki (2003) in incidental vocabulary acquisition. For example, in Waring and Takaki's (2003) research, approximately half of the word knowledge gained in the reading was lost after three months.

FREQUENCY OF WORD OCCURRENCE AND THE NUMBER OF TIMES FOR READING

The results in Tables 4 and 5 showed that learners' vocabulary performance was greatly affected by the frequencies of word occurrence in this study. The overall conclusion was that the higher the frequency of word occurrence is; the better results the learners achieve.

Where the estimates of the number of times that learners need to learn a word is concerned, the results in Tables 4 and 5 exhibited that with words encountered only one time,

it was unlikely for learners to recognize or produce depth of vocabulary knowledge. It was also assumed that to have a 50% chance of recognizing the depth of vocabulary knowledge for a first-time reading, 14 encounters of the word was the prerequisite, and 18 times were needed for productive vocabulary knowledge. After the second-time reading, a frequency rate of only 8 times was needed for a 50% chance of vocabulary knowledge recognition, and a frequency rate of only 14 times was needed for a 50% chance of productive vocabulary knowledge.

Overall findings in the effects of word exposure frequency reasonably matched previous studies (e.g., Chen & Truscott, 2010; Webb, 2007). The results demonstrated an overall pattern for each test of rising scores occurring as the number of exposures increased. Acquisition of productive knowledge required more exposures than were needed for receptive knowledge. Differences appeared on some specific points. For example, Webb (2007) showed that if learners were exposed to unknown words ten times in context, sizeable learning gains occurred. More than ten repetitions may be needed to achieve full knowledge of a new word. On the other hand, Chen and Truscott (2010) demonstrated that a large distinction occurred between words that were encountered 3 times and 7 times, a small disparity was found between the words that were encountered 1 time and 3 times. The present study presented a picture of gradual, cumulative gains with increasing exposure, unambiguously supporting the role of repetition in incidental vocabulary acquisition. These differences might be attributable to the differences in target words and context. For example, Webb used nonsense words and short sentences in his study, Chen and Truscott used actual words and reading materials created by advanced bilingual Chinese EFL students, while pseudo-words and a graded reader were used in the present study. The participants may have found it difficult to learn new words in a condition where the pseudo-words were used to represent new vocabulary.

Results from the present study disagreed with Pigada and Schmitt's (2006) findings, where higher-frequency words did not necessarily display greater chances of being learned. However, in their case study of a learner of French, the participant had already acquired partial knowledge of some high-frequency words, which could have resulted in a ceiling effect. Put simply, there was less room for improvement. Similarly, in other previous studies (e.g., Horst et al., 1998; Saragi et al., 1978), some of the low-frequency words were easier for the learners to learn than some of the high-frequency words, supporting the assertion that frequency of word exposure was only one of the factors that affected word acquisition. This might have been due to the difference in the morphology or salience of the word or the context in which the word was encountered.

Therefore, although it has been accepted that the form–meaning link of a new lexical item could be strengthened when repetitions were increased, one thing to bear in mind was that since the number of target words was different and context for discerning meanings was different, the findings need to be verified by further research before any generalizations can be made.

RETENTION ABILITY

Vocabulary learning was easily forgotten when the learners had no exposure to target words for three months, even if they participated in a repeated reading of the materials. Concerning the receptive vocabulary knowledge, only 2.9 words (occurred 22 times) were remembered by the learners who did not receive any related reading for three months. Likewise, only 0.6 words (occurred 22 times) were successfully produced regarding the depth of vocabulary knowledge. This part of the results was worse than those that were found in Waring and Takaki's (2003) study, in which approximately half of the word knowledge gained in the

reading was lost after three months (*ibid.*) However, they measured only word form and meaning in their study. The present study attempted to measure the depth of vocabulary knowledge, which is more difficult to be acquired. Therefore, a high degree of word frequency does not mean a guarantee of maintaining learners' memory of the target words. This highlighted the value of recycling the newly-learned words. According to the results of the present study, participants suffered a significant loss in word knowledge gains over a three-month period, thus rehearsal of newly-learned words was regarded as an important aspect for learners to maintain their initial form-meaning link. It is suggested that teachers should provide opportunities for learners to recycle the new lexical items frequently in an attempt to curb precipitous decays in word retention.

CONCLUSIONS

Overall, based on data analysis and discussion, this study has shown that words can be learned incidentally from reading. However, many new words that were learned from this type of reading were soon lost. In addition, relatively few new words were learned from reading a graded reader as measured by a test on productive word knowledge. Nevertheless, more vocabulary knowledge was acquired receptively, demonstrating participants' relative weakness in productive vocabulary in spite of the repeated readings of a text with at least 95% coverage of known words.

The author had no intention of generalizing the results concluded from a group of Chinese EFL students in the present study for other learners in various Asian countries. However, this study could help teachers engaged in EFL vocabulary teaching find similarities between their students and the participants described in the study, and based on this, acknowledge the following suggestions for teaching and learning EFL vocabulary.

Learners' productive vocabulary knowledge tended to be slower than their receptive vocabulary size. The findings of this study appear promising since frequency of occurrence clearly has the potential to boost form recall of collocated words and synonyms for the target words. The findings should, however, be interpreted with caution given the following reasons:

First, a limited number of participants, a limited number of target items per frequency band, and target words might have been highlighted due to repeated reading. It can still be confidently stated that learners could establish initial form-meaning connections through reading a controlled text with different target word exposure frequency, and an increased frequency level helped learners to process target words elaborately and to consolidate their form-meaning link.

Second, these results also suggested that the nature of incidental vocabulary learning from reading is more complex than can be determined from this study. Indeed, it suggested that a considerable amount of vocabulary knowledge was gained, but was not retained. One thing that was worth noting was that in EFL learning context, students' performance is seldom tested immediately after a reading program. Thus, rehearsal of newly-learned words was still important for learners to maintain their initial form-meaning link. However, many previous studies have shown the attrition of vocabulary knowledge gain in EFL students (e.g., Teng, 2015b; Webb, 2007). The issue of how to reduce EFL learners' disappointing forgetting rate may be worthy of further research.

Finally, the low acquisition rate of vocabulary knowledge found through this study, as well as in other research studies focusing on incidental vocabulary learning, suggested a value in emphasising the importance of combining incidental learning with some sort of explicit focus (Eckerth & Tavakoli, 2012). If incidental learning is not the focus of teaching, it may be helpful for teachers to guide learners to focus on explicit elaborate processing of some difficult words (Mirzaii, 2012). Teachers can also pre-teach the meaning of difficult

words that are necessary for comprehension but unlikely to be known from the contexts in which they occur. In addition, there is a value in instructing learners to utilize both receptive and productive learning opportunities meaningfully beyond their English language course (Kaur, 2015).

LIMITATIONS AND FUTURE RESEARCH

One of the limitations experienced was that only two dimensions regarding the depth of vocabulary knowledge were covered in the present study, which was far from enough to demonstrate that the learners have acquired full aspects of vocabulary knowledge (Nation, 2001). Another limitation was that the substituted words were expressed in a natural way; any substantive evidence was not presented on the point of whether the words should be highlighted or expressed in a natural way. Future studies regarding this controversial issue would be beneficial. The third limitation was that the study did not analyse the difficulties of learning nouns, verbs and adjectives, respectively (Zimmerman, 2009). Future studies should explore this point as well as other factors that might determine word learning: for example, the context in which a target word is introduced (Teng & He, 2015; Webb, 2008), or the specific needs of particular learners.

REFERENCES

- Baddeley, A. (1990). *Human Memory*. London, England: Lawrence Erlbaum and Associates.
- Chen, C., & Truscott, J. (2010). The Effects of Repetition and L1 Lexicalization on Incidental Vocabulary Acquisition. *Applied Linguistics*, 31, 693-713.
- Cobb, T. (n.d.). The Compleat Lexical Tutor [Computer Software]. Retrieved October 14th, 2015, from <http://www.lex tutor.ca/>.
- College English Test (CET) Band 4 and 6. (2005). Retrieved October 14th, 2015, from <http://www.cet.edu.cn/>
- Currie, L. (1997). Why Use a Novel? *Reading*, 31, 11-16.
- Day, R., Omura, C., & Hiramatsu, M. (1991). Incidental EFL Vocabulary Learning and Reading. *Reading in a Foreign Language*, 7(2), 541-551.
- Eckerth, J., & Tavakoli, P. (2012). The Effects of Word Exposure Frequency and Elaboration of Word Processing on Incidental L2 Vocabulary Acquisition through Reading. *Language Teaching Research*, 16(2), 227-252.
- Heatley, A., Nation, I. S. P. and Coxhead, A. (2002). RANGE and FREQUENCY programs. Achieved October 14th, 2015 from http://www.vuw.ac.nz/lals/staff/Paul_Nation
- Henriksen, B. (1999). Three Dimensions of Vocabulary Development. *Studies in Second Language Acquisition*, 21, 303-317.
- Hill, D. R. (2013). Graded Readers. *ELT Journal*, 67, 85-125.
- Horst, M. (2005). Learning L2 Vocabulary through Extensive Reading: A Measurement Study. *The Canadian Modern Language Review*, 61(3), 355-382.
- Horst, M., Cobb, T., & Meara, P. (1998). Beyond a *Clockwork Orange*: Acquiring Second Language Vocabulary through Reading. *Reading in a Foreign Language*, 11(2), 207-223.
- Hu, M., & Nation, I. S. P. (2000). Unknown Vocabulary Density and Reading Comprehension. *Reading in a Foreign Language*, 13, 403-30.
- Hulstijn, J. H. (2013). Incidental learning in second language acquisition. In C. A. Chapelle (Ed.). *The Encyclopedia of Applied Linguistics* (pp. 2632-2640). Chichester: Wiley-Blackwell.
- Kaur, N. (2015). Making Meaning of Vocabulary Learning: Seizing Opportunities at Opportune Moments. *GEMA Online® Journal of Language Studies*, 15(2), 1-16.

- Laufer, B. (1989). What percentage of lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.). *Special Language: From Human Thinking to Thinking Machines* (pp. 316–23). Clevedon, England: Multilingual Matters.
- Lee, S. H. (2003). ESL Learners' Vocabulary Use in Writing and the Effects of Explicit Vocabulary Instruction. *System*. 31, 537-561.
- Lessard-Clouston, M. (2013). *Teaching Vocabulary*. Alexandria, VA: TESOL International Association.
- Mirzaii, M. (2012). Implicit vs. Explicit Vocabulary Learning: Which Approach Serves Long-term Recall Better? *3L: Language Linguistics Literature®*, *Southeast Asian Journal of English Language Studies*. 18(2), 1-12.
- 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.
- Nassaji, H. (2004). The Relationship between Depth of Vocabulary Knowledge and L2 Learners' Lexical Inferencing Strategy Use and Success. *The Canadian Modern Language Review*. 61. 107-134.
- Nation, I. S. P. (1990). *Teaching and Learning Vocabulary*. Rowley, MA.: Newbury House.
- Nation, I. S. P. (2001). *Learning Vocabulary in another Language*. Cambridge, England: Cambridge University Press.
- Nation, I. S. P. (2007). The Four Strands. *Innovation in Language Learning and Teaching*. 1, 1-12.
- Nation, I. S. P., & Wang, K. (1999). Graded Readers and Vocabulary. *Reading in a Foreign Language*. 12(2), 355-380.
- Paribakht, T. S., & Wesche, J. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In J. Coady & T. Huckin (Eds.), *Second Language Vocabulary Acquisition: A Rationale for Pedagogy* (pp. 174-202). Cambridge, England: Cambridge University Press.
- Pellicer-Sa'nchez, A., & Schmitt, N. (2010). Incidental Vocabulary Acquisition from an Authentic Novel: Do Things Fall Apart? *Reading in a Foreign Language*. 22, 31-55.
- Pigada, M., & Schmitt, N. (2006). Vocabulary Acquisition from Extensive Reading: A Case Study. *Reading in a Foreign Language*. 18, 1-28.
- Pulido, D. (2007). The Effects of Topic Familiarity and Passage Sight Vocabulary on L2 Lexical Inferencing and Retention through Reading. *Applied Linguistics*. 28, 66-86.
- Qian, D. D. (1998). *Depth of vocabulary knowledge: Assessing its role in adults' reading comprehension in English as a second language*. Unpublished doctoral dissertation. University of Toronto, Toronto, Canada.
- Qian, D. D. (1999). Assessing the Roles of Depth and Breadth of Vocabulary Knowledge in Reading Comprehension. *Canadian Modern Language Review*. 56(2), 282-307.
- Qian, D. D. (2002). Investigating the Relationship between Vocabulary Knowledge and Academic Reading Performance: An Assessment Perspective. *Language Learning*. 52(3), 513-536.
- Read, J. (1993). The Development of a New Measure of L2 Vocabulary Knowledge. *Language Testing*. 10, 355-371.
- Read, J. (2004). Word Associates Test (Version 4.0). Achieved November, 14th, 2015 at <http://www.lex tutor.ca/>
- Richards, J. C. (1976). The Role of Vocabulary Teaching. *TESOL Quarterly*. 10, 77-89.
- Rott, S. (1999). The Effect of Exposure Frequency on Intermediate Language Learners' Incidental Vocabulary Acquisition through Reading. *SSLA*. 21, 589-619.
- Rott, S. (2007). The Effect of Frequency of Input-enhancements on Word Learning and Text

- Comprehension. *Language Learning*. 57, 165–99.
- Saragi, T., Nation, I. S. P., & Meister, G. F. (1978). Vocabulary Learning and Reading. *System*. 6, 72–8.
- Schmitt, N. (1998). Tracking the Incremental Acquisition of Second Language Vocabulary: A Longitudinal Study. *Language Learning*. 48, 281-317.
- Schmitt, N. (2008). Review Article: Instructed Second Language Vocabulary Learning. *Language Teaching Research*. 12(3), 329-363
- Tekmen, E. A. F., & Daloglu, A. (2006). An Investigation of Incidental Vocabulary Acquisition in Relation to Learner Proficiency Level and Word Frequency. *Foreign Language Annals*. 39, 220–43.
- Teng, F. (2014). Incidental Vocabulary Learning by Assessing Frequency of Word Occurrence in a Graded Reader: Love or Money. *Language Education and Acquisition Research Network (LEARN) Journal*. 7(2), 36-50.
- Teng, F. (2015a). The Effectiveness of Extensive Reading on EFL Learners' Vocabulary Learning: Incidental versus Intentional Learning. *Brazilian English Language Teaching Journal*. 6, 82-96.
- Teng, F. (2015b). Extensive Reading Plus Explicit Vocabulary Exercises: Is it Better than Extensive Reading-only? *Malaysian Journal of ELT Research*. 11(2), 82-101.
- Teng, F. (2016). Incidental vocabulary acquisition from reading and listening: The effects of word exposure frequency. In E. Dorman & J. Bidal (Eds.). *Departing from Tradition: Innovations in English Language Teaching and Learning* (pp.182-207). Cambridge: Cambridge Scholars Publishing.
- Teng, F., & He, F. (2015). Towards Effective Reading Instruction for Chinese EFL Students: Perceptions and Practices of Lexical Inferencing. *The English Teacher*. 44(2), 56-73.
- van Zeeland, H., & Schmitt, N. (2013). Incidental Vocabulary Acquisition through L2 Listening: A Dimensions Approach. *System*. 41, 609-24.
- Waring, R., & Nation, I. S. P. (2004). Second Language Reading and Incidental Vocabulary Learning. *Angles on the English Speaking World*. 4, 97-110.
- Waring, R., & Takaki, M. (2003). At What Rate do Learners Learn and Retain New Vocabulary from Reading a Graded Reader? *Reading in a Foreign Language*. 15(2), 1–27.
- Webb, S. (2007). The Effects of Repetition on Vocabulary Knowledge. *Applied Linguistics*. 28, 46-65.
- Webb, S. (2008). The Effects of Context on Incidental Vocabulary Learning. *Reading in a Foreign Language*. 20(2), 232–245.
- Webb, S., & Chang, A. C-S. (2012). Vocabulary Learning through Assisted and Unassisted Repeated Reading. *The Canadian Modern Language Review*. 68(3), 267-290.
- Webb, S., & Chang, A. C-S. (2014). Second Language Vocabulary Learning through Extensive Reading: How does Frequency and Distribution of Occurrence Affect Learning? *Language Teaching Research*. 18, 1-20.
- Webb, S., Newton, J., & Chang, A. (2013). Incidental Learning of Collocation. *Language Learning*. 63, 91-120.
- Wesche, M. B., & Paribakht, T. S. (1996). Enhancing Vocabulary Acquisition through Reading: A Hierarchy of Text-related Exercise Types. *The Canadian Modern Language Review*. 52(2), 155-178.
- Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring Vocabulary through Reading: Effects of Frequency and Contextual Richness. *The Canadian Modern Language Review*. 57(3), 541-572.
- Zimmerman, C. B. (2009). *Word Knowledge: A Vocabulary Teacher's Handbook*. Oxford, England: Oxford University Press.

APPENDIX

Examples of the measurement tool used in this study (3 out of 36 items).

Test 1.

Directions: Please produce a collocated word or a synonym for each given word.

1. doonful _____ (collocation)
doonful _____ (synonym)
2. tuidy _____ (collocation)
tuidy _____ (synonym)
3. daccit _____ (collocation)
daccit _____ (synonym)

Test 2.

Directions: Please choose one collocation, and one synonym for each given word. The eight options were put into two groups, with six being distracters, distributed either in the left or right box.

1. Doomful

old youthful beautiful clever

shirt house people family

2. tuidy

glad sad good lucky

plant child painting classroom

3. daccit

tall hot false real

planet sky words weather

ABOUT THE AUTHOR

Feng Teng is a language teacher educator with extensive experience in China. He is interested in doing research on vocabulary studies, metacognition and autonomy. His recent publications appeared in *Innovation in Language Learning and Teaching* and *TESL-EJ*.