Task Effectiveness and Vocabulary Learning and Retention in a Foreign Language

Feng Teng
tengfeng@uni.canberra.edu.au
Nanning University, China

Abstract

The present study provides further evidence that vocabulary learning and retention in a foreign language are contingent on the involvement load that a task induces, (i.e., the amount of *need*, *search*, and *evaluation*), as proposed in Laufer and Hulstijn's (2001) Involvement Load Hypothesis (ILH). A total of 60 participants were selected based on their word level, and were divided equally into three groups. Each small group was required to complete one of the three tasks. Task 1 participants extensively read a prescribed amount of business materials provided with L1 marginal glosses. The participants in Task 2 completed the same requirements as in Task 1, along with the additional requirement of target word suppliance. Task 3 participants were required to create an original composition by using the target words after fulfilling the same requirements as in Task 1. The participants were not told that they would be administered post-tests after the tasks. In line with the predictions of the ILH, the findings of this study revealed that Task 3 was the most beneficial in terms of immediate learning and retention of target words, followed by Task 2, and then Task 1. The value of form-focused instruction and rehearsal of newly learned words was accentuated based on the results of this study.

Key words: Vocabulary learning, Involvement Load Hypothesis, incidental learning, task

Introduction

It is widely acknowledged that knowing a large amount of vocabulary facilitates reading comprehension. A great deal of research on vocabulary has had a considerable emphasis on developing vocabulary fluency for students learning English as a Foreign Language (EFL) in Asian contexts (Bell, 2001; Laufer & Nation, 2001; Zhang & Lu, 2014). To achieve this fluency, Laufer (2003) recommended word-focused tasks, which are superior to reading tasks because they help learners engage in more elaborate processing of previously known words. Put simply, intentional word learning is superior to incidental learning via reading. Some researchers have pointed out that a large amount of word learning does occur incidentally via extensive reading (Horst, 2005; Waring and Nation, 2004). Hence, learners are encouraged to extensively read both authentic (Nunan, 2002) and comprehensive materials (Krashen, 1985) in order to gain knowledge of word meanings (Hulstijn, 1992; Waring and Takaki, 2003).

The issue of incidental vs. intentional word learning aside, the primary concern of teachers is to organize tasks that provide more opportunities for learners to encounter the unknown words. However, the time for students to intentionally learn

every unknown word in class is limited. Thus, incidental vocabulary learning through reading is proposed (Teng, 2014; Warring & Takaki, 2003). Concerning this, Laufer and Hulstijn (2001) proposed the Involvement Load Hypothesis (ILH). The ILH states that word learning is subject to three task factors: *need, search,* and *evaluation*. Tasks with a higher degree of *need*, *search*, and *evaluation* will yield a better retention of learned words. Although tasks have been conducted with varying measures of the ILH (Huang, Eslami, & Willson, 2012; Keating, 2008; Kim, 2008), it is not known how active knowledge of newly learned words was affected by the ILH, which is the main purpose of this study.

Literature Review

Incidental vocabulary learning

The issue of incidental vocabulary learning has existed for decades and many studies have been conducted to confirm the position of incidental learning.

Incidental learning, which is achieved in an implicit way, is a "by-product" of any reading activity (Hulstijn, 2001, p.266). In other words, when learners try to understand the embedded meaning of the context, the acquisition of words occurs subconsciously (Paribakht and Wesche, 1999). Many studies have shown the effects of extensive reading on incidental vocabulary learning (Brown, Waring, & Donkaewbua, 2008; Coady, 1997; Teng, 2015a). For example, in order to enhance incidental vocabulary learning, some studies focused on incidental vocabulary gains through analyzing the frequency of word occurrence. Their results showed that words with a higher frequency level were retained better by EFL learners (Day, Omura and Hiramatsu, 1992; Horst, 2005; Pellicer-Sánchez &Schmitt, 2010; Pigada & Schmitt, 2006; Waring & Takaki, 2003). In another study, Chen and Truscott (2010) analyzed the effects of repetition on incidental vocabulary learning and found that repetition improved the outcome of incidentally learned productive knowledge. Webb and Chang (2012) also showed that repeated reading of material with assisted learning led to significantly greater incidental vocabulary learning than unassisted learning. However, the issue of context is one limitation mentioned in the above studies. It is argued that a word meaning that is difficult to infer in one context might be easy to infer in another context. Accordingly, Webb (2008) focused on the context and showed that more informative context had a greater effect on incidental vocabulary learning than less informative context.

Although research has been successful in improving incidental vocabulary learning, Schmitt (2008) pointed out that incidental vocabulary learning alone does not increase total vocabulary learning at a significant rate. Vocabulary-focused tasks are superior to vocabulary learning through reading alone (Laufer, 2003; Paribakht & Wesche, 1997; Zimmerman, 1997) because learners can notice form, meaning, and usage. Hence, some scholars have proposed that reading plus intentional vocabulary learning might enhance overall vocabulary learning (Laufer & Hulstijn, 2001; Stoller & Grabe, 1993; Webb, 2005). Nakata's (2008) study analyzed the spacing effects in intentional learning by comparing vocabulary learning through word lists, word cards, and computers. The computer-based sequencing algorithm was found to be superior to the word lists. However, having students intentionally learn every word in class is unfeasible. Thus researchers agreed on the necessity of conducting more research on incidental vocabulary learning.

According to the research mentioned above, it appears that incidental vocabulary learning has shown to be effective. More research is needed to evaluate its results and reinforce those results based on a new variable, Involvement Load Hypothesis.

Involvement Load Hypothesis

Researchers widely agree with Craik and Lockhart's (1972) theory of 'depth of processing,' which states that a deeper processing of information yields a longer retention of words. Put succinctly, the deeper the level of the information processed, the longer the memory is maintained. Craik and Tulving (1975) added the idea of elaboration to the processing model, asserting that elaborating on new input information related to pre-existing information makes it easier for learners to improve learning and retention.

Based on the above theoretical ideas, Laufer and Hulstijn (2001) proposed the Involvement Load Hypothesis (ILH). The notion of the ILH was not originally posited in the context related to form-focused instruction. Form-focused instruction refers to the position that overt attention to linguistic elements is beneficial for learners to acquire native-like competence in English (Ellis, 2001). Literature up to date has focused the issue of form-focused instruction more on grammar acquisition. The assumption was that vocabulary is mainly acquired through extensive reading. Thus, Laufer (2005) introduced the position of form-focused instruction to lexical acquisition. Her study showed that vocabulary gains due to form-focused instruction were significantly better than reading alone. As mentioned above, the ILH was not originally related to form-focused instruction. However, when vocabulary tasks are contrasted in terms of the effort that a task induces, it is acknowledged that an increased involvement load generally entails a greater form-focused position (where *form* refers to lexical items). Hence, to some extent, the ILH was closely connected to form-focused instruction.

According to the ILH, the amount of involvement in a task determines the retention ability of unfamiliar words. Involvement is considered a motivational-cognitive construct, which includes three elements: *need, search,* and *evaluation*.

The element of *need* is the incentive to learn the vocabulary, and is a motivational, but non-cognitive, dimension of involvement that can be distinguished as a strong *need* (++) or a moderate *need* (+) based on intrinsic or extrinsic factors. In other words, a *need* is strong when the learner is self-motivated to learn the vocabulary. On the other hand, a *need* is a moderate factor when it is forced by extrinsic factors, such as, a task or a teacher.

Search and *evaluation* are the two cognitive dimensions of involvement. A similarity of the two constructs is that both require focusing on word forms and meanings. *Search* refers to finding the meaning of an unknown word by locating the definition using a dictionary or asking for help from a teacher. *Search* is absent when this attempt is not required.

Evaluation involves comparing the meaning or usage of the target word with other meanings, or comparing the word with other words to find out if the word can be used in a certain context or not. For instance, when learners are required to use a dictionary to look up the meaning of a word, they need to compare all the meanings of the word to confirm which one is more semantically appropriate for the specific context. *Evaluation* is moderate when learners are required to locate differences

between words in a certain context, and strong when learners are required to identify the meanings of unknown words and produce them in a new context. An example of moderate *evaluation* would be choosing the correct word in context; whereas strong *evaluation* would be writing a sentence or composition with the word.

According to the ILH, the three components of *need, search,* and *evaluation* do not always appear simultaneously during a reading task, as some of the constructs may be absent. Hustijn and Laufer (2001) proposed the "involvement index" to define the depth of processing for a task. According to this index, the absence of a component is marked as 0, the moderate presence of a component is marked as 1, and the strong presence of a component is marked as 2. According to the ILH, language learners who are engaged in a task with a higher degree of involvement according to the involvement index are more likely to yield a better performance in learning vocabulary.

Empirical studies for the ILH

Keating (2008) invited seventy-nine beginning learners of Spanish to complete one of three tasks with different involvement indexes: Reading comprehension (index of one), reading comprehension plus target word suppliance (index of two), and sentence writing (index of three). Passive and active word knowledge were measured in his study. Results showed that learning and retention of target words was highest in the sentence writing tasks. The next best was reading comprehension plus target word suppliance, followed by reading comprehension only. Similar results were found in Huang, Eslami, & Willson (2012).

Teng (2015b) went one step further to compare four vocabulary tasks with different involvement indexes. A total of 180 Chinese EFL students were placed into four word levels based on their test results: 40 learners were placed at the 2,000 word level; 60 learners at the 3,000 word level; 48 learners at the 5,000 word level; and 32 learners at the university word level. The participants in each level were equally and randomly divided into four groups, with each subgroup completing one of four vocabulary learning tasks that varied in the involvement load required: reading comprehension (index of one), reading comprehension plus supplied target words (index of two), reading comprehension plus composition writing (index of three), and reading comprehension plus dictionary look-up and composition writing (index of four). Results revealed that, in line with the predictions of the ILH, vocabulary learning was highest in the fourth task, and descended according to involvement load.

These previous studies shed light on how to identify a task that provides the best opportunity for learning new words. It is essential to design different tasks in Asian EFL contexts in order to pursue a range of related improvements in task-based language teaching.

Limitations of prior research

Research conducted to date that has measured incidental word learning based on the ILH was limited in two ways. First, most of the previous research focused on using a multiple-choice test, which is an instrument that measures receptive knowledge, not usage. It is acknowledged by some scholars that knowing a word is more than simply knowing its form and meaning (Nation, 2001; Schmitt, 2000). In addition, learners might make wild guesses on this type of test. Therefore, two tests of measuring active word knowledge were applied in the present study. Second, most of the previous

research has only studied advanced learners, who might have a better ability of inferring unknown words. For example, participants in Eckerth and Tavakoli's (2012) study were at an IELTS 7 level. Involving learners with a lower word level would make the research more inclusive. The current study attempts to cover these two issues. Therefore, two related research questions were proposed:

1. Do the predictions of the Involvement Load Hypothesis also apply to learners at the 3,000 word level?
2. Do tasks with different involvement indexes result in different learning scores in both dimensions of active word knowledge?

Methodology

Participants

60 business English majors, age 19–22, were invited to participate in this study. Before the study, 180 students were required to take the vocabulary level test replicated from Nation (1983, p. 19–24). Laufer (1992, 1996) used this test in her studies, and the test was widely applied in L2 vocabulary research. The reason for choosing this test was to assure that participants had a similar word level before completing the tasks. Learners achieved one point for each correct answer in matching a word and its definition, and needed at least 13 out of 18 possible points to show that they had achieved the related word level. Following is an example from the 3,000 word level:

1. administration
2. angel ____ managing business and affairs
3. frost ____ spirit who serves god
4. herd ____ group of animals
5. mate
6. bond

60 students at the 3,000 word level were selected as participants in the present study. The reason for selecting students in this level is that having a 3,000 word level is a basic requirement for adequate reading comprehension (Qian, 2002). Moreover, it is assumed that students with a 3,000 word level would not understand the target words at the 7,000, 8,000, and 9,000 word levels. Thus, the present study is workable.

Materials and target words

As mentioned above, participants were business English majors. To gain their interest, the materials were chosen from the business section of the BBC News. Twenty-five news articles were chosen and the total word count of all the articles was 9,465 words.

The current study used the *Range* program to analyze the word level for the chosen materials (Heatley, Nation, & Coxhead, 2002). This program has been widely used in vocabulary research (Horst, 2005; Pellicer-Sánchez & Schmitt, 2010; Pigada & Schmitt, 2006; Waring and Takaki, 2003). Out of 3,000 word level, 26 words were chosen as the test items (Table 1). Although it is possible that participants might have

known some of the selected words, it is considered to be unlikely based on the researcher's teaching experience.

Table1. Test Items at Low Frequency Level

Word level	Selected target words
7,000	Adamant, anonymity, insolvent, infuriate, pessimistic, resurgence, sluggish, swathe, indebted, prelude, grapple, entrant, beverage, deflate
8,000	Enumerate, inversion, longevity, reticence, revamp, stagnant, zest
9,000	Brandish, eavesdrop, snub, gnarled, impunity

Procedure

The 60 business English major students invited to participate in this study were divided into three groups. Each group included 20 learners. Each group was then randomly assigned with one of three different tasks. One English teacher was responsible for one task, with three teachers total.

The participants were not informed of the purpose of the study. They were also not told that vocabulary tests would be administered after the reading program (immediate tests directly after completion of the tasks and delayed tests two weeks later), so as to not motivate participants to memorize the words in anticipation of the tests. Therefore, the results of the present study were in line with the requirements of incidental vocabulary learning. During the two weeks after the immediate tests, the participants were not exposed to the reading materials.

The time for finishing the tasks varied among the treatment groups. Tasks 1, 2, and 3 took an average of 90, 120, and 150 minutes to complete, respectively. After the participants completed their tasks, the teachers collected their materials and distributed the post-tests, as described below.

Vocabulary Proficiency Test

The vocabulary proficiency test was designed to measure the participants' active word knowledge of the 26 target items. The test included two parts (see Appendix I). The first part was an active recall test. It provided prompt sentences for the learners to guess and write down the test words. An example is shown below:

If a company has no money to pay their debts, it means they are a(an) ____ company. (insolvent)

Participants were given one point for writing a correct word and zero points for an incorrect word.

The second part was an active production test. Learners were required to produce a sentence (on any theme) with the test words. Similarly, the learners received one point for producing a sentence that correctly incorporated the target word in terms of usage and zero points for an incorrect answer. Grammar was not judged since it was

not the purpose of the present study.

For example:
He lose his job, because the company is insolvent.
They lost all the orders, and the company became insolvent within weeks.
He insolvent his company.

The above three sentences were from the participants' sentence writing. The first two sentences were rated correct (one point). Although student who wrote the first sentence should have used the past tense of the word "lose," they were still given one point because the test only measured the students' ability in using the target words. The third sentence, however, clearly showed that student was confused about the usage of the word "insolvent." Therefore, the third sentence was rated incorrect (zero points).

The maximum possible score for the test was 52 points (26 points each for parts one and two). The results of the test were used to establish the proficiency of the participants. This vocabulary test was administered twice: immediately after the reading program to assess the learners' vocabulary growth and two weeks after the program to assess learner's retention ability. To preempt ordering effects, the second tests were identical to the first tests with the exception of the item order.

Tasks

Three tasks were designed and each task was given a different load index for the present study, as shown in Table 2.

Table 2. Three Tasks with Different Involvement Load Indexes

Tasks	Conditions for target words	Procedures	Involvement Load			
			need	search	evaluation	index
Task1 RC	Marginally Glossed In L2	Read the news	+	--	--	1
Task2 RC plus fill-in	Marginally Glossed In L2	1, Read the news 2, Fill-in requirements	+	--	+	2
Task3 RC plus composition writing with the target words	Marginally Glossed In L2	1, Read the news 2, Write composition with the target words	+	--	++	3

RC=reading comprehension

The involvement construct of *need* was held constant (+), because the requirements of the three tasks were imposed by the tasks.

Task 1 was glossed reading. Participants were not required to search for the

meaning of the words or compare the meanings with other words. Thus, *search* and *evaluation* were absent (-). Therefore, the involvement index of Task 1 was one.

Participants assigned to Task 2 were also not required to search for the meaning of the word by using a dictionary or asking for help from teachers. Thus, *search* was absent (-). However, in Task 2, the target words were deleted from the text and replaced with uniformly-sized blanks. The target words not appearing in the original text were reprinted in alphabetical order on a separate page. Each target word was followed by a part of speech, as well as a brief English definition. Thus, *evaluation* was modest (+) in Task 2. Therefore, the involvement index was two in Task 2.

In Task 3, participants were required to produce the words in context (write a composition on any theme), so *evaluation* was strong (++). However, *search* was absent because participants were not required to locate the meaning of the new words. Therefore, the involvement index in Task 3 was three (Examples of the three tasks were provided in Appendix II).

Results

The results of the two tests are provided in Table 3. Figures 1 and 2 provide the same results graphically.

Table3. Descriptive Statistics for the Two Tests Administered at Different Time

Task		Immediate test		Delayed test	
	N	Mean	S.D.	Mean	S.D.
Part I (active recall)					
Task 1	20	4.35	1.75	2.10	0.97
Task 2	20	7.25	2.45	4.65	1.35
Task 3	20	11.80	2.30	9.50	1.15
Part 2 (active production)					
Task 1	20	2.12	0.89	1.05	0.88
Task 2	20	4.85	1.10	2.50	0.90
Task 3	20	9.95	2.50	7.21	1.33

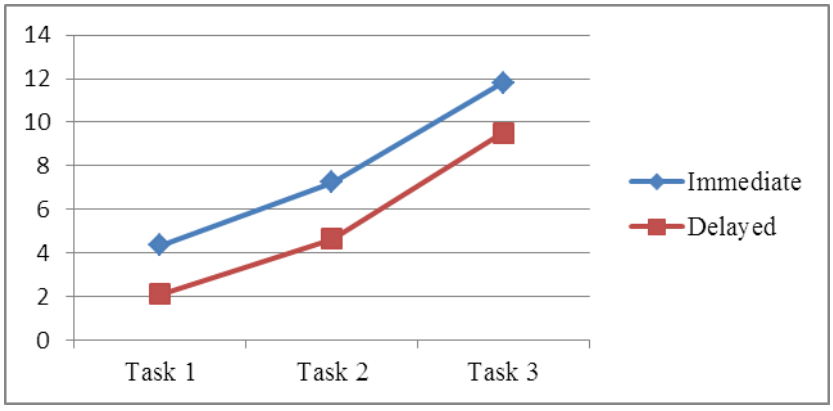

Figure 1. The Graphical Results for Part I (Active Recall) Administered at Two

Different Times

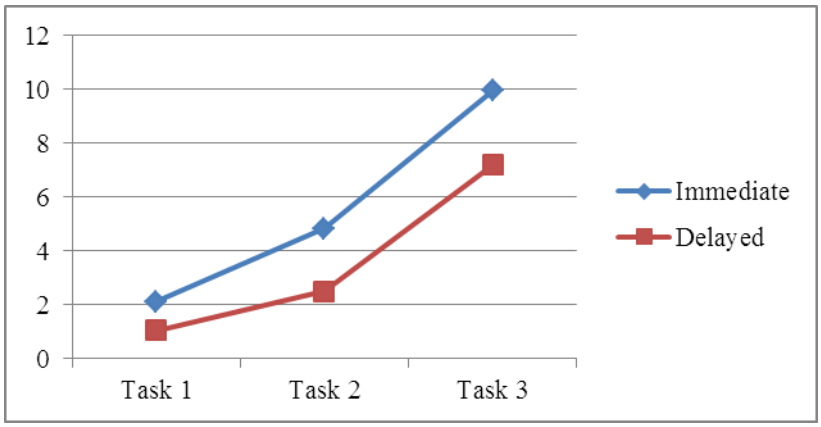

Figure 1. The Graphical Results for Part II (Active Production) Administered at Two Different Times

As shown in the above table and figures, the participants in each task showed a better learning score on the immediate tests than the delayed tests. The optimal task in producing the best learning outcomes was Task 3, followed by Task 2, and then Task 1. In addition, there was a larger standard deviation in the delayed tests than in the immediate tests, which means substantial attrition occurred for many participants when they were not exposed to the reading materials for the two weeks.

To verify this, the scores on the active recall test were submitted to a two-way mixed ANOVA analysis, with the three tasks as the between-subject factor and the two different administration times as the within-subjects factor. The results showed a main effect of Task, $F(2,76)=24.106$, $p<0.001$, and also a main effect of Time, $F(2,76)=18.535$, $P<0.001$. A post-hoc Turkey analysis also revealed that Task 3 was superior in active recall of word knowledge, followed by Task 2 and Task 1, respectively. The ANOVA analysis conducted on the delayed test also revealed a significant effect of Task, $F(1,76)=4.562$, $P=0.02$, and a significant effect of Time, $F(2,76)=6.463$, $P=0.01$. A post-hoc Turkey analysis also revealed that Task 3 yielded better results in retention of active word form and meaning than Task 2, and that the mean scores of Task 2 were significantly better than Task 1.

The scores on the active production test were also submitted to a two-way mixed ANOVA analysis, with the three tasks as the between-subject factor and the two different administration times as the within-subjects factor. The results showed a main effect of Task, $F(2,76)=14.481$, $p<0.001$, and a main effect of Time, $F(2,76)=12.281$, $P<0.001$. A post-hoc Turkey analysis also revealed that Task 3 yielded best results in free production, followed by Task 2, and then Task 1. The ANOVA analysis conducted on the delayed test also revealed a significant effect of Task, $F(1,76)=4.069$, $P=0.021$, and a significant effect of Time, $F(2,76)=6.063$, $P=0.012$. A post-hoc Turkey analysis revealed that Task 3 was the optimal task, followed by Task 2, and then Task 1.

Repeated Wilcoxon signed-rank analyses of comparing the two test scores were then applied. The comparisons between the two tests were significant at the $p<.05$ level.

Discussion

The first question of this research was: Do the predictions of the Involvement Load Hypothesis apply to EFL learners at the 3,000 word level?

It appears that the three tasks consisting of different degrees of involvement load yielded differing results. The results of the Chinese EFL students who completed Task 3 were remarkably better than those students who completed Task 2, followed by those who completed Task 1. The results were consistent in both the immediate and delayed tests. The effects of the most involving tasks were greater than those of the least involving tasks after two weeks. Thus, the results of the present study fully support the predictions of the Involvement Load Hypothesis. This is in line with previous studies on the matter (Hulstijn & Laufer, 2001). However, the results were different from Keating (2008), where Task 3 was not more effective than Task 2. One possible explanation may lie in the differences of designing Task 3 across the two studies. Although Task 1 and Task 2 were similar in both studies, Task 3 in the present study required participants to write an original composition on any theme, while Task 3 in Keating's (2008) study was concerned with creating an original sentence. This might mean that producing connected discourse (writing a composition) facilitated learners in more elaborate processing of words than creating sentences (unrelated). As proposed in Laufer's (2003) study, the mean scores of the sentence writing task were significantly lower than that of the composition writing task (2.14 vs. 3.73).

The second question was: Do tasks with different involvement indexes result in different learning scores on the immediate and delayed tests?

Armed with the above data analysis, the results indicated that free production using the target words was more difficult than active recall of the target words for learners at the 3,000 word level. This means that learners can only learn the use of an unknown word after they have understood the word form and meaning (Nation, 2001). The results highlighted the importance of active word knowledge. While many teachers focus on receptive word knowledge when teaching new words, it may be the active word knowledge which is the most difficult dimension to learn. Therefore, it is highly suggested that, when teachers organize different tasks and expect students to learn words incidentally from reading, the aspect of active word knowledge should be taken into account.

Pedagogical Implications

The results of the present study indicated that the *evaluation* construct of task-induced involvement played a critical role in incidental vocabulary learning. The reading plus composition writing task (strong *evaluation*) resulted in significantly better learning outcomes than the reading plus fill-in-the-blank (moderate *evaluation*), and the reading plus fill-in-the-blank task yielded better results than the reading comprehension only task (no *evaluation*). Therefore, EFL learners benefited more from the task that used target words productively in an original composition, and pushing learners to compare new words with words already known was more beneficial than general glossed reading.

The second implication drawn from the current study was related to the positive, reinforcing effect of form-focused instruction on incidental vocabulary learning. Previous studies have strongly supported this (Keating, 2008; Laufer, 2005;

Laufer &Rozovski-Roitblat, 2011). Two of the reading tasks in the present study included a form-focused component. Task 2 included target word suppliance and Task 3 included using the target words in an original composition. These two tasks were more beneficial than the reading-only task (without form-focused instruction). Accordingly, in future teaching of vocabulary, form-focused tasks that require high degrees of *evaluation* should be primarily used. Teachers might consider the value of form-focused instruction to consolidate learning done in the classroom.

A final implication to be drawn from the present study is the importance of repeated exposure to target words. There was substantial attrition in productive word knowledge over the course of two weeks. Thus, the improved learning gained from the task-induced involvement needs to be recycled in order to be maintained; otherwise the advantage of form-focused instruction may be lost. Previous studies have also fully supported the importance of rehearsal during the process of learning (Folse, 2006; Sonbul & Schmitt, 2010). Therefore, in order to curb precipitous declines in initial learning gains, teachers should provide more opportunities for learners to rehearse the newly learned words.

Conclusion

In conclusion, it should be emphasized that there is value for task-induced involvement in EFL vocabulary learning. The present study demonstrated that EFL learners at the 3,000 word level benefit from more involving tasks, and that a more involving task yields better results in active word knowledge than a less involving task. The present study also highlighted the value of form-focused instruction and rehearsal of newly acquired words.

Although this study was conducted in provincial China, the results have resonance for the many global contexts where English is taught and learned as a foreign language.

Limitations and Future Research

First, because of the methodological limitations of the current study (e.g., completion times were not held constant across the four tasks), generalizations may not be able to be made concerning task-induced involvement. Some evidence has suggested that the benefits associated with more effective tasks are negated when completion times are held constant across tasks (e.g., Folse, 2006; Webb, 2005). Furthermore, the issue of word exposure frequency was not taken into consideration (Laufer & Rozovski-Roitblat, 2011). Adding this factor would make the current study more inclusive. Finally, if the present study involved participants with a lower proficiency level (e.g., 2,000 word level), the effects of task-induced involvement might be different (Teng, 2015b).

Acknowledgement

Thanks for the research funding provided by Nanning University (grant number: 2014JSGC10). Many thanks also go to the two anonymous reviewers for their very constructive suggestions on this article.

References

Bell, T. (2001). Extensive reading: Speed and comprehension. *Reading Matrix: An International Online Journal, 1*(1), 1-13.

Brown, R., Waring, R., & Dongkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language, 20*(2), 136-163.

Chen, C., & Truscott, J. (2010). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition. *Applied Linguistics, 31*(5), 693-715.

Coady, J. (1997). L2 vocabulary acquisition through extensive reading. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp.225-237). Cambridge, England: Cambridge University Press.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671-684.

Craik, F. I. M., & Tulving, E.(1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104,* 268-294.

Day, R. R., Omura, C., & Hiramatsu, M. (1992). 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.

Ellis, R. (2001). Introduction: Investigating form-focused instruction. *Language Learning, 51(supplement1),* 1-46.

Folse, K. S. (2006). The effect of type of written exercise on L2 vocabulary retention. *TESOL Quarterly, 40,* 273–293.

Heatley, A., Nation, I. S. P., & Coxhead, A. (2002). RANGE and FREQUENCY programs. Retrieved March 18th, 2015 from http://www.vuw.ac.nz/lals/staff/Paul_Nation

Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study. *The Canadian Modern Language Review, 61*(3), 355–382.

Huang, S., Eslami, Z., & Willson, V. (2012). The effects of task involvement load on L2 incidental vocabulary learning: A meta-analytic study. *The Modern Language Journal, 96*(4), 544-557.

Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In P. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp. 113–125). London, England: Macmillan.

Hulstijn, J. H. (2001). Intentional and incidental second-language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 258–286). Cambridge, England: Cambridge University Press.

Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning, 51*(3), 539-558.

Keating, G. D. (2008). Task effectiveness and word learning in a second language: The involvement load hypothesis on trial. *Language Teaching Research, 12*(3), 365-386.

Kim, Y. (2008). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language Learning, 58*(2), 285-325.

Krashen, S. D. (1985). *The input hypothesis: Issues and implications.* New York: Longman.

Laufer, B. (1992). How much lexis is necessary for reading comprehension? In H. Bejoint & P. Arnuad (Eds.), *Vocabulary and applied linguistics* (pp. 126-132). London, England: Macmillan.

Laufer, B. (1996). The lexical threshold of second language reading comprehension: What it is and how it relates to LI reading ability. In K. Sajavaara & C. Fairweather (Eds.), *Approaches to second language acquisition* (pp. 55-62). Jyviiskyla: University of Jyviiskyla.

Laufer, B. (2003). Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence. *The Canadian Modern Language Review, 59,* 567–587.

Laufer, B. (2005). Focus on form in second language vocabulary learning. *EUROSLA Yearbook, 5,* 223–250.

Laufer, B., & Hulstijn, J. H. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics, 22,* 1-26.

Laufer, B., & Nation, I. S. P. (2001). Passive vocabulary size and speed of meaning recognition: Are they related? *EUROSLA Yearbook, 1,* 7–28.

Laufer, B., & Rozovski-Roitblat, B. (2011). Incidental vocabulary acquisition: The effects of task type, word occurrence and their combination. *Language Teaching Research, 15*(4), 391-411.

Nakata, T. (2008). English vocabulary learning with word lists, word cards and computers:
Implications from cognitive psychology research for optimal spaced learning. *ReCALL*, *20*(1), 3-20.

Nation, I. S. P. (1983). Testing and teaching vocabulary. *Guidelines*, *5*(1), 12-25.

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

Nunan, D. (2002). Listening in language learning. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 238–241). Cambridge: Cambridge University Press.

Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary development. In. J. Coady & T. Huckin (Eds.). *Second language vocabulary acquisition: A rational for pedagogy* (pp.174-200). Cambridge, England: Cambridge University Press.

Paribakht, T. S., & Wesche, M. (1999). Reading and "incidental" L2 vocabulary acquisition. *Studies in Second Language Acquisition, 21*(2), 195–224.

Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do Things Fall Apart? *Reading in a Foreign Language, 22*(1), 31-55.Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition through extensive reading: A case study. *Reading in a Foreign Language, 18*(1), 1-28.

Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading comprehension: An assessment perspective. *Language Learning, 52,* 513-536.

Schmitt, N. (2000). *Vocabulary in language teaching.* Cambridge: Cambridge University Press.

Schmitt, N. (2008). Instructed second language learning. *Language Teaching Research, 12,* 329-363.

Sonbul, S., & Schmitt, N. (2010). Direct teaching of vocabulary after reading: It is worth the effort? *ELT Journal, 64*(3), 253-260.

Stoller, F., & Grabe, W. (1993). Implications for L2 vocabulary acquisition and instruction from L1 vocabulary research. In T. Huckin, M. Haynes & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 29-45). Norwood, NJ: Ables.

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, January). The effectiveness of extensive reading on EFL learners' vocabulary learning. *Paper presented at the 3rd English Language Teaching International Conference, Macau S.A.R., P.R. China*.

Teng, F. (2015b). EFL vocabulary learning through reading BBC news: An analysis based on Involvement Load Hypothesis. *English as a Global Language Education (EaGLE) Journal, 2,* 20-40.

Waring R., & Nation, I. S. P. (2004). Second language reading and incidental vocabulary learning. *Angles on the English-Speaking World, 04,* 11-23.

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

Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language, 20*(2), 232–245.

Webb, S., & Chang, A. (2012). Vocabulary learning through assisted and unassisted repeated reading. *The Canadian Modern Language Review, 68* (3), 267-290.

Zhang, X., & Lu, X.F. (2014). A longitudinal study of receptive vocabulary breadth knowledge growth and vocabulary fluency development. *Applied Linguistics, 35*(3), 283-304.

Zimmerman, C. (1997). Do reading and interactive vocabulary instruction make a difference? An empirical study. *TESOL Quarterly, 31*(1), 121-14.

Appendix I
Vocabulary proficiency test

Sample test items (2 out of 26 items)

1. a. active recall

If a company has not enough money to pay their debts, it means they are a(an) ____ company.

b. active production

Please use the word that you have written down to create a new sentence.

_____

2. a. active recall

If someone is determined not to change their mind, it means s/he is_____ about something.

b. active production

Please use the word that you have written down to create a new sentence.

_____

Appendix II
Sample tasks

Task 1: Glossed reading (1 item out of 26)
A pre-pack administration is one in which the **insolvent** company has already lined up a buyer for its profitable assets before it enters administration, allowing a sale within days.

Insolvent: A person or an organization that has not enough money to pay their debts

Task 2: Reading and filling-in the gaps
A pre-pack administration is one in which the _____ company has already lined up a buyer for its profitable assets before it enters administration, allowing a sale within days.

Word list (5 items out of 26)

Words	Part of speech	Definition
Adamant	adj.	Unwilling to change the minds
Anonymity	n.	The state of remaining unknown to most other people
Infuriate	v.	Make someone extremely angry
Insolvent	adj.	A person or an organization that has not enough money to pay their debts
Pessimistic	adj.	Someone who believes bad things will happen

Task 3: Glossed reading and writing a composition
A pre-pack administration is one in which the **insolvent** company has already lined up a buyer for its profitable assets before it enters administration, allowing a sale within days.

Insolvent: A person or an organization that has not enough money to pay their debts

Follow-up writing
Please write a composition on any theme. Please note that you need to use all the words in the word list.

Word list

Words	Part of speech	Definition
Adamant	adj.	Unwilling to change the minds
Anonymity	n.	The state of remaining unknown to most other people
Infuriate	v.	Make someone extremely angry
Insolvent	adj.	A person or an organization that has not enough money to pay their debts
Pessimistic	adj.	Someone who believes bad things will happen