The Effect of Evaluation Factor on the Incidental Vocabulary Acquisition through Reading

Chao Wang[1], Kun Xu[1] & Ying Zuo[1]

[1] Foreign Language Department, Beijing Institute of Petrolchemical Technology, Beijing, China

Correspondence: Chao Wang, Foreign Language Department, Beijing Institute of Petrolchemical Technology, Beijing, No. 19, Qingyuan North Road, Daxing District, Beijing, China. Tel: 1-343-650-2487. E-mail: wangchao@bipt.edu.cn

Abstract

This empirical study investigates the respective effectiveness of three factors (need, search and evaluation) included in task-induced involvement load on the EFL vocabulary learning and retention. Three tasks with the same amount of involvement load but containing different factors are assigned to 108 non-English majors at Beijing Institute of Petrol-chemical Technology in China. After these reading tasks, the participants are given an unannounced immediate posttest. One week later, the participants are given the delayed posttest. A 3×2 analysis of variance (ANOVA) is employed to process the scores and identify the relationship between the EFL incidental vocabulary learning and the three factors contained in the involvement loads. The results are assumed to show that the Evaluation factor is more decisive and crucial than the other two factors (need and search). Learners benefit more by using the target words in their original contexts. That means vocabulary instruction should focus on tasks that require high degrees of evaluation.

Keywords: EFL learning, incidental vocabulary acquisition, task-induced involvement load

1. Introduction

1.1 Vocabulary Learning: An Enormous Task of EFL Learning

Vocabulary has always been regarded as a core component of language learning. EFL (English as a foreign language) learners and their teachers are all aware of the painful fact that EFL learning, especially to intermediate and advanced levels, involves an enormous task of learning a large number of vocabularies, which becomes the barrier to improve the EFL proficiency. Correspondingly, many researchers are motivated to find out how vocabularies are best learned in a foreign language learning context. One of the most influential studies is Task-induced Involvement Load Hypothesis (Hulstijn & Laufer, 2001), which suggests that the higher the degree of involvement is, the more possible the acquisition occurs. It is the first time for researchers to devise an observable method (involvement load index) to measure the degree to which readers process the knowledge of vocabulary through reading.

Based upon the theoretical framework in the Task-induced Involvement Load Hypothesis, an empirical study is conducted at Beijing Institute of Petrol-chemical Technology in China. The purpose is to further test whether the three factors in the involvement load hypothesis are equally important to enhance EFL learners' incidental vocabulary acquisition, and to provide practical feedback to EFL vocabulary teaching and learning.

1.2 Intentional Vocabulary Acquisition versus Incidental Vocabulary Acquisition

Intentional vocabulary learning means that learners deliberately commit to memory the knowledge and usage of the target words, such as pronunciation, meaning, form as well as their grammatical and social functions. But vocabulary learning can also occur incidentally (Hulstijn, 2003). Learners may acquire vocabulary by picking up the meaning, form or usage of unknown words while doing other inputting activities, like reading and listening. Actually, many vocabularies are learned incidentally through extensive reading (Hunt & Beglar, 1998). Such a process is completely subconscious, when learner's attention is not focused on vocabulary at all. Learners do not know that they acquire vocabulary while they read, but the by-product is quite rewarding.

1.3 Reading and Incidental Vocabulary Acquisition

Many L1 researches indicate that direct vocabulary instruction does not provide a significant proportion of the words learners acquire. The major way for L1 vocabulary increasing is by learning through context (Nagy, Herman, & Anderson, 1985; Nation & Coady, 1988).

Studies concerning EFL or L2 reading practice and vocabulary acquisition also indicate that if learners are motivated and focused on meaning, increasing the amount of reading leads to measurable vocabulary acquisition (Elley & Mangubhai, 1983; Ferris, 1988; Kiyochi, 1988, cited in Krashen, 1989; Pitts, White, & Krashen, 1989). That means vocabulary can be incidentally acquired through independent reading (Pitts, White, & Krashen, 1989; Day, Omura, & Hiramatsu, 1991; Day & Swan, 1998; Tekmen & Daloglu, 2006; Xu, 2010).

Such kind of incidental vocabulary acquisition has long been focused by EFL and L2 researchers. However, other group of scholars doubts the role of incidental vocabulary acquisition, arguing that it does not have effect a high pick-up rate (Schmitt, 2008). Thus, numerous vocabulary enhancement tasks have been researched to identify the effectiveness of incidental vocabulary acquisition while reading (Knight, 1994; Rott & Cameron, 2002; Hill & Laufer, 2003; Kim, 2008).

1.4 Task-Induced Involvement Load Hypothesis

The more deeply readers process vocabulary through reading, the greater possibility there is for vocabulary to be acquired. Whether a word can be acquired is dependent on the amount of attention allocated to it (Rott & Cameron, 2002). In order to provide an observable tool to measure the "depth of processing", Hulstijn and Laufer (2001) proposed the involvement load hypothesis. A motivational-cognitive construct of involvement was formulated to describe the degree of processing induced by the given task, which contains three factors (need, search, and evaluation). Need is the motivational factor. If the task needs readers to be aware of the meaning of the unknown word, need involvement is present. Search is the cognitive factor which is present if the task needs readers to search for the meaning or form of the target words. Evaluation is the cognitive factor which is present if the task needs readers to compare the form or meaning with other possible words or to assess whether its form or meaning is suitable in the context. An absence of a factor is marked as 0 (-), a moderate presence of a factor as 1 (+), and strong presence as 2 (+ +). Specific standards to identify the degree of the involvement load are shown in Table 1. The hypothesis suggests that the higher the degree of involvement is, the more possibly the acquisition occurs.

Table 1. Specific standards to identify the degree of the involvement load

Factors	Degree of the involvement load	Standard
Need	Index 0 (absent)	Readers do not feel the need to make clear the meaning of unknown words.
	Index 1 (moderate)	Readers are externally required to make clear the meaning of unknown words.
	Index 2 (strong)	Readers internally decide to make clear the meaning of unknown words.
Search	Index 0 (absent)	Readers do not have to search for the meaning or form of unknown words.
	Index 1 (moderate)	Readers have to search for the meaning of unknown words.
	Index 2 (strong)	Readers have to search for the form of unknown words.
Evaluation	Index 0 (absent)	Readers do not compare unknown words with others (meaning & form).
	Index 1 (moderate)	Readers compare unknown words with others in the provided context.
	Index 2 (strong)	Readers compare unknown words with others in the original context.

Thereafter, language scholars and educators' passion is inspired, whose researches focus on the difference between narrative and argumentative passages, between passages with a familiar theme and unfamiliar theme. Some other researches focus on the different effects of involvement loads on verbs, nouns, and adjectives. Although reading materials and involvement load have been researched widely (Hill & Laufer, 2003; Hulstijn et al., 1996; Hulstijn & Laufer, 2001; Knight, 1994; Kim, 2008; Rott, 2004), few studies have investigated the effectiveness of tasks with the same degree of involvement load but different factors on EFL vocabulary learning and retention through reading. The current research attempts to explore this question in an empirical way.

2. Methodology

2.1 Research Question

The present research is to investigate the respective effect of tasks with the same amount of involvement load but containing different factors on the learning and retention of target vocabularies among college non-English majors in China in an incidental learning setting (reading), and providing feedback to EFL learners and teachers on how to acquire vocabulary effectively through reading.

The research question is:

With the same amount of involvement load, are the three factors in the involvement load hypothesis equally important to learners' vocabulary learning and retention?

2.2 Participant (Subject) Characteristics

One hundred and eight Chinese intermediate EFL learners at Beijing Institute of Petrol-chemical Technology (BIPT) in China participate in this study. All participants are non-English majors with an English proficiency of level B (the intermediate level based on the Entrance Level Test hosted by BIPT). The age range of the students is between 18 and 20.

Two intact classes of second-year students (56 + 52) are selected for the experiment, respectively majoring in mechanical engineering and chemical engineering. Students in each class are randomly divided into three experimental groups. Each group is supposed to read the same selected passage and finish its specific task.

2.3 Instruments

2.3.1 Reading Passage

Hu and Nation (2000) suggest that only when learners have an adequate comprehension of a text (containing 98% known vocabulary) is the incidental vocabulary acquisition likely to happen. The passage used in the present study includes 96% of words, which is assumed to be familiar to learners (confirmed by a trial test). This density was regarded as being at an appropriate level to make incidental vocabulary learning possible. The theme of this passage is the youth market in North America, which appeals to the college students in China. Therefore, in terms of the difficulty (containing 96% known vocabulary, a favorable theme) and length (512 words), participants are able to have an adequate comprehension of the selected reading material.

2.3.2 Target Words

Twenty students in other classes, who are not going to participate in the experiment, are required to underline any words in the passage that they do not know. Ten target words are selected from the passage, including four nouns, two adjectives and four verbs. A pilot test is conducted among another twenty students of a similar English proficiency to make sure that students are not familiar with the selected target words.

These words are: *budget, gigantic, allowances, estimate, purchase, market, manufacturer, latest-style, institution,* and *debt*.

2.3.3 Reading Tasks

The present research investigated three different tasks with the same amount of involvement loads, but containing different factors. They are designed based on Hulstijn and Laufer's experimental framework (2001). Each task is described in the following details.

Task 1: Reading comprehension with True or False questions

Participants are supposed to read the selected passage (the Chinese meaning of the target words are not offered) and to finish ten true/false questions according to the information in the passage. For example,

In North America, young people have ***gigantic*** buying power in today's economy. (True/False)

Participants are allowed to use dictionary. In terms of the involvement load, due to the fact that the knowledge of target words is essential to understand the passage and the true/false statement, Task 1 induces strong need, moderate search but no evaluation. Its involvement index is (2 + 1 + 0).

Task 2: Reading comprehension with multiple choices of the meaning of target words

Participants assigned to Task 2 are supposed to read the same passage (the Chinese meaning of the target words are not offered) and are required to choose among four choices the right Chinese meaning of the target word. For example,

In North America, young people have ***gigantic*** buying power in today's economy.　(　　　)

a. 巨大的   b. 有限的   c. 奇特的   d. 吸引人的

In terms of involvement load, participants are required to make clear the Chinese meaning of the target word by either consulting the dictionary or inferring through context, and decide which choice is the correct one. Therefore task 2 induces moderate need, moderate search, and moderate evaluation. Its involvement index is (1 + 1 + 1).

Task 3: Reading comprehension with forming new sentences with the target words

Participants assigned to Task 3 are also arranged to read the same passage (the Chinese meaning of the target words are offered) and are required to write original sentences using the ten target words. For example,

In North America, young people have **_gigantic_** buying power in today's economy.

Your own sentence: _____.

Similar to Task 2, need in Task 3 is moderate because it is imposed by the task and search is absent because the Chinese meaning is provided. Evaluation is strong because participants have to use the new words with previously known words in order to create original sentences. Thus, Task 3 induces moderate need, absent search, and strong evaluation. Its involvement index is 3 (1+0+2).

2.3.4 Testing and Scoring

Two unannounced tests are conducted to evaluate the initial learning and retention of the target words among all the participants in different groups, under the condition that participants do not know when they are doing their respective tasks.

The instant test is conducted immediately after participants complete the required tasks. The delayed test is conducted one week after their completion of the tasks.

In both tests, participants are required to give the Chinese translation or English explanation to evaluate their learning and retention of the target words.

The scoring standard is adopted from Hulstijn and Laufer (2001), in which not translated or wrongly translated words got zero, a correct response received 1 point and approximately correct response obtained half a point.

2.4 Data Collection

Participants are going to read the selected passage and finish their respective tasks on the same day in their normal reading classes to make sure that the reading passage is fresh and the immediate vocabulary test is credible. The delayed vocabulary test is to also be conducted on the same day one week later.

Participants are divided into three groups in each class. Each group is given the same worksheet with one reading passage and one specific reading task. The time on task was not controlled, which means participants are allowed as much time as they like to complete the task, because the focus of the experiment is to investigate the involvement load instead of the quality and efficiency of reading. Once tasks are completed, the worksheets are gathered. Afterwards, participants are arranged to complete an immediate unannounced vocabulary test. A week later, the same unannounced test is conducted again. Both tests were scored by one experienced teacher according to the scoring standard mentioned above.

3. Results and Discussion

Date collected from the two posttests is analyzed by using SPSS12.0.

3.1 Descriptive Statistics of the Immediate and Delayed Tests

The descriptive statistics of the scores on the immediate and delayed tests will be displayed in the following table.

Table 2. Descriptive statistics of the immediate and delayed tests

Group	N	Immediate		Delayed	
		Means	SD	Means	SD
Group 1 (Task 1: T/F questions)	34	3.958333333	2.349392019	3.180555556	2.046435151
Group 2 (Task 2: multiple-choice of Chinese meaning)	36	5.180555556	2.47314946	3.244444444	2.512192491
Group 3 (Task 3: original sentence making)	36	6.411764706	1.912758551	4.941176471	1.857681456

As Table 2 has shown, the difference between scores achieved by different groups is remarkable and obvious. Participants doing Task 3 achieve the highest score in both the immediate (M = 6.411764706) and delayed tests (M = 4.941176471). Participants doing Task 2 (M = 5.180555556) perform better than those doing Task 1 (M = 3.958333333). That indicates the three factors (need, search and evaluation) in the involvement load are not equally important to affect the vocabulary learning and retention through reading. The higher the degree of evaluation is, the greater possibility there is for vocabulary to be acquired. Task 1 involves the highest degree of the factor of need, but Group 1's scores are low in the two tests (M = 3.958333333, M = 3.180555556). It can be inferred that the effect of need factor on incidental vocabulary acquisition is not decisive.

What's more, although Group 3 and Group 2 did a fairly good job in the immediate test, every group's scores in the delayed test decline dramatically. This reminds us that vocabulary retention is by no means easy and needs to be reinforced from time to time. The influence of exposure is worth being further investigated.

There is another point worth to be mentioned. The standard deviation of the scores of Group 3 (SD = 1.912758551, SD = 1.857681456) in both tests is not as high as that of Group 1 and Group 2. This can be interfreted as a proof that by doing Task 3, which contains evaluation factor, readers tend to be more engaged into processing the vocabulary knowledge to the greatest extent. Even the relatively less diligent participants would acquire something in the process of doing the task.

3.2 The Relationship between the Word Learning and Factors in Task Involvement Load

A 3 × 2 analysis of variance (ANOVA) is employed to process the two test scores, with task as the between-subject factor (reading with T/F questions, reading with meaning choosing, and reading with sentence writing) and time (immediate and delayed tests) as the within-subjects factor (see Table 3). The measure of effect size isη2, expressing explained variance.

Table 3. Task effect, time effect and the effect of task × time interaction

	df	F	Sig.	η2
Time	2	72.183	.000	.340
Task	2	44.942	.000	.479
Time × task	4	17.964	.000	.235

As is shown in Table 3, the task effect [F (2, 105) = 72.18; sig = .000 < .01; η2 = .340], time effect [F(2, 105) = 44.94; sig = .000 < .01; η2 = .479], and the task × time interaction [F(4, 105) = 17.96; sig = .000 < .01; η2 = .235] are all significant. In other words, there exist obvious differences in the mean scores among the three groups. The statistics indicate that the degree of involvement load does have different effects on the learning and retention of new words through reading.

3.3 The Differnece Location

As is analyzed above, the F value is significant. Therefore Scheffé post hoc tests (see Table 4) are selected as an option in the ANOVA to locate the differences.

Table 4. Scheffé post hoc multiple range test for immediate and delayed tests

Test	(I) Tasks	(J)Tasks	Mean Difference (I-J)	Std. Error	Sig.
Immediate Test	Task 1	Task 2	-2.13055*	.46355	.000
		Task 3	-3.46142*	.46355	.000
	Task 2	Task 1	2.13055*	.46355	.000
		Task 3	-1.28308*	.43549	.022
	Task 3	Task 1	3.46142*	.46355	.000
		Task 2	1.28308*	.43549	.022
Delayed Test	Task 1	Task 2	-.06727	.67212	.695
		Task 3	-1.93352*	.67212	.024
	Task 2	Task 1	.06727	.67212	.695
		Task 3	-1.84175*	.66526	.028
	Task 3	Task 2	1.93352*	.67212	.024
		Task 3	1.84175*	.66526	.028

As is demonstrated in Table 4, there is a significant difference among Task 1, Task 2 and Task 3 in the immediate posttest (p = .000 < .05, p = .000 < .05 and p= .022 < .05 respectively). However, statistics in the Scheffé post hoc tests for delayed posttest only shows difference between Task 1 and Task 3 (p = .024 < .05) and between Task 2 and Task 3 (p = .028 < .05). Although the mean score of Task 2 is higher than that of Task 1, no difference was found between Task 1 and Task 2 (p = .695 > .05).

The reason may lie in the fact that the evaluation factor in Task 2 (readning comprehension with meaning choosing) does stimulate better performance compared with Task 1 (reading comprehension with T/F questions), but it is not strong enough to guarantee learners' retention of new vocabularies.

4. Conclusion

Based on the above data and analysis, the results of the empirical study can be concluded as the follows:

First, with the same amount of involvement load, the three factors in the task-induced involvement load hypothesis are not equally important to learners' incidental vocabulary acquisition through reading. The Evaluation factor is more decisive and crucial than the other two factors (need and search). Task 2 (reading with meaning choosing) (moderate evaluation) and Task 3 (reading with sentence writing) (strong evaluation) may encourage greater gains of word knowledge compared to task 1(reading with T/F questions) (no evaluation). In other words, vocabulary that is processed with involvement load including higher degree of evaluation will be retained better than that with lower degree. The need factor is not as influencing as the search factor or the evaluation factor. That means the motivation factor is important, but cannot be competing with the cognitive factors.

Second, by comparing tasks with varying degrees of evaluation, the present research may indicate that learners benefit more by using the target words productively in original contexts. In the actual teaching and learning practice, vocabulary instruction should include meaning-focused tasks that require high degrees of evaluation.

Finally, learners need opportunities to be exposed to newly acquired words, especially in a productive way, such as creating their original context. Task 3 leads to the greatest gains, but still suffers the dramatic decrease one week later. Instructors can stimulate larger initial acquisition to prevent or delay declines in word retention by doing reading tasks with involvement load containing high degree of evaluation factor.

Acknowledgments

References

Day, R., & Swan, J. (1998). Incidental vocabulary of foreign language spelling through targeted reading. *TESL Reporter*, *31*, 1-9.

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

Elley, W., & Mangubhai, F. (1983). The impact of reading on second language learning. *Reading Research Quarterly*, *19*, 53-67. http://dx.doi.org/10.2307/747337

Ferris, D. (1988). *Reading and second language vocabulary acquisition* (Unpublished paper). Department of Linguistics, University of Southern California.

Hill, M., & Laufer, B. (2003). Type of task, time-on-task and electronic dictionaries in incidental vocabulary acquisition. *IRAL*, *41*, 87-106. http://dx.doi.org/10.1515/iral.2003.007

Hu Hsueh-chao, M., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, *13*(1), 403-430.

Hulstijn, J. H. (2003). Insidental and intentional learning. In C. Doughty, & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 349-381). Oxford: Blackwell. http://dx.doi.org/10.1002/9780470756492.ch12

Hulstijn, J., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *Modern Language Journal*, *80*, 327-339. http://dx.doi.org/10.1111/j.1540-4781.1996.tb01614.x

Hunt, A., & Beglar, D. (1998). Current research and practice in teaching vocabulary. *The Language Teacher*, *22*(1).

Kim, Y. (2008). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language Learning*, *58*, 285-325. http://dx.doi.org/10.1111/j.1467-9922.2008.00442.x

Knight, S. (1994). Dictionary use while reading: The effects on comprehension and vocabulary acquisition for students of different verbal ability. *Modern Language Journal*, *78*, 285-299. http://dx.doi.org/10.1111/j.1540-4781.1994.tb02043.x

Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *Modern Language Journal*, *73*(4), 440-463. http://dx.doi.org/10.1111/j.1540-4781.1989.tb05325.x

Nagy, W., Herman, P., & Anderson, R. (1985). Learning Words from context. *Reading Research Quarterly*, *20*, 233-253. http://dx.doi.org/10.2307/747758

Nation, I. S. P., & Coady, J. (1988). Vocabulary and reading. In R. Carter, & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 97-110). London: Longman.

Pitts, M., White, H., & Krashen, S. (1989). Acquiring second language vocabulary through reading: A replication of the Clockwork Orange study using second language acquirers. *Reading in a Foreign Language*, *5*(2), 271-275.

Rott, S. (2004). A comparison of output interventions and un-enhanced reading conditions on vocabulary acquisition and text comprehension. *The Canadian Modern Language Review*, *61*, 169-202. http://dx.doi.org/10.3138/cmlr.61.2.169

Rott, S., Williams, J., & Cameron, R. (2002). The effect of multiple-choice L1 glosses and input-output cycles on lexical acquisition and retention. *Language Teaching Research*, *6*, 183-222. http://dx.doi.org/10.1191/1362168802lr108oa

Schmitt, N. (2008). Instructed second language learning. *Language Teaching Research*, *12*, 329-363. http://dx.doi.org/10.1177/1362168808089921

Tekmen, E., & Daloglu, A. (2006). An investigation of incidental vocabulary acquisition in relation to learner proficiency level and word frequency. *Foreign Language Annals*, *39*, 220-243. http://dx.doi.org/10.1111/j.1944-9720.2006.tb02263.x

Xu, X. (2010). An empirical study on the effect of task on L2 incidental vocabulary acquisition through reading. *Asian Social Science*, *6*(7), 126-131.

Appendix A

The Selected Reading Passage

The Youth Market Just Keeps on Growing

It's well-known that money is very important to every one of us, but many people, especially youths, do not know how to spend money wisely. For example, over-consumption is one of the problems for every family. A good **budget** will help guide us to the right price range as well as prevent us from spending more than we could afford. It's wise of us to make budget as early as possible.

In North America, young people have **gigantic** buying power in today's economy. This has an effect on the media and on many companies that expect this market to increase their profits. First, the amount of money that teens have to spend is large and growing. The money they have comes from their **allowances**. The amount of money children receive for their allowance has increased by 15 percent in the past ten years. Parents give their children large allowances when the economy is strong. Today, more parents are working. Their work keeps them so busy that they don't have enough time to shop with their children. Many parents are just too tired when they come home at the end of the day to discuss spending with their children. In fact, today many children influence their parents' decisions about what to buy for their family. Nine out of ten kids have some power over family spending decisions. Some researchers **estimate** that up to $565 billion of family **purchases** on such items as cars, furniture, vacation, and home entertainment items are influenced by children.

Many companies are now **marketing** directly to young teens and pre-teens—especially girls. Major cosmetic and clothing **manufacturers** are producing brands that are attractive to a youth market. The clothes that young people want to buy are expensive—$150-250 for the **latest-style** blue jeans, $200 for a pair of shoes. The music and entertainment industries have traditionally marketed their products to the young, but now even the oldest **institutions**—banks—are looking to this new market. They are promoting themselves to children as young as seven years old with special accounts and programs designed to teach them about savings and checking. Clearly, the youth market is important.

Of course, all of this offers parents and others a great opportunity to teach children about responsible attitudes toward money and consumption. Some studies show that the sooner kids learn how to budget, invest, and save money, the more successful they will be at avoiding **debt** later in life. But there are, of course, the dangers. Other studies suggest that kids typically spend 90 percent of the money they have, and their credit card debt is a problem that many parents have.

Learning how to budget and save money should start early. Living within our means (量入为出) is a lesson that should be learned at a young age and reinforced (加强) throughout our lives. Staying away from debt is better than getting out of debt. But in our consumer society, it isn't easy.

Copyrights