

Higher Task-induced Involvement Load Enhances Students' EFL Vocabulary Learning

Mohammad Reza Ghorbani
University of Bojnord, Bojnord, Iran
Email: mrg872@yahoo.com

Maryam Rahmandoost
Islamic Azad University, Garmsar Branch, Iran
Email: maryam.rahmandoost@yahoo.com

Abstract—Foreign or second language researchers and teachers have been trying for hundreds of years to develop effective instructional tasks so that they can facilitate and enhance students' vocabulary learning. Accordingly, many claims have been made in this regard. For instance, the Involvement Load Hypothesis proposed by Hulstijn and Laufer suggests that higher task-induced involvement load leads to better vocabulary learning. This study was an attempt to test and substantiate this Hypothesis. Experiments were conducted with two classes of non-English major students who were assigned tasks with different involvement loads of reading comprehension passages. Scores of vocabulary tests were collected before and after the treatment. The reliability of the test was estimated 0.70 through KR-21 formula. Independent samples t-test was used to compare the possible differences between the means of the experimental and control groups based on the gain scores from the pretest and posttest. The final result of the independent samples t-test for the comparison of the mean increment between the two groups after treatment indicated that the experimental group who received the task with higher involvement load had a better performance than the control group who didn't. Thus, the findings suggest that vocabulary learning in the fill in the blanks and sentence writing condition is significantly higher than the true or false and multiple-choice comprehension questions condition. Since the task with higher involvement load seems to have facilitated the process of vocabulary learning, the Hypothesis is supported and its application is suggested to reinforce the word retention in English as a Foreign Language (EFL) or English as a Second Language (ESL) contexts.

Index Terms—EFL vocabulary learning, involvement load hypothesis, task, reading comprehension

I. INTRODUCTION

Vocabulary as a core component of language proficiency is the basis of reading, writing, listening, and speaking ability. Learning words and how to use them is quite challenging and EFL learners cannot achieve their potential without an extensive vocabulary. One of their main problems with vocabulary learning is that it is here today, gone tomorrow. There are many theories regarding vocabulary learning. Many researchers believe that modeling vocabulary development is a difficult task. Despite more than a century of research on this multi-faceted phenomenon, researchers and theorists are still coming up with new models to help teachers figure out the best ways of instruction that might foster EFL vocabulary learning. Teachers, in turn, have been applying these models to help students increase their vocabulary and ensure better retention. Many authors suggest that effective tasks require a deeper level of processing of new words. One of the most commonly accepted views of vocabulary learning is that tasks with higher involvement load lead to better vocabulary retention.

Hulstijn and Laufer (2001) proposed the Involvement Load Hypothesis which involves three factors (need, search, evaluation), where absence of a factor is marked as 0 (-), a moderate presence of a factor as 1 (+), and strong presence as 2 (+ +). The notion of involvement load includes both motivational and cognitive components. Need is the motivational component which is present if the unknown word is needed to complete the task. Search is the cognitive component which is present if the learner needs to search for or retrieve the meaning or form of a particular word. Evaluation is the cognitive component which is present if the task involves having to compare the form or meaning with other possible words or meanings to choose the most suitable one for the context. The involvement load of the task equals the sum of the strengths of the three factors. The greater the involvement load, the better the learning.

Some researchers state that task-induced involvement entails the type of processing crucial for learning. They suggest that effective tasks require a deeper level of processing of new words. However, more empirical evidence is needed to support this hypothesis. One of the major ways of researching vocabulary teaching techniques is doing experimental comparisons of vocabulary learning activities (Hulstijn and Laufer, 2001). Learning new words as a by-product of a non-vocabulary activity can happen when students do reading, listening, speaking and even writing tasks. This study is an attempt to explore if tasks with higher involvement load lead to better EFL vocabulary learning. It investigates the effect of two tasks with different degrees of involvement load on EFL vocabulary learning through reading.

II. LITERATURE REVIEW

According to Kim (2011), teachers are well aware of the fact that second language learners need to acquire so many words to become fluent, but they probably do not know how best to help their students in this regard. Therefore, identifying learning tasks which provide and lead to optimal vocabulary learning opportunities is pedagogically justified.

Many studies have been conducted in recent years regarding the enhancement of vocabulary learning through text-based tasks. According to Hulstijn (1992), if the meanings of unfamiliar words are correctly inferred during a reading task, they are remembered better than words which have been explained by their synonyms. According to Hulstijn, Hollander, and Greidanus (1996), if words are looked up in a dictionary during a reading task, they are remembered better than words that are glossed in text margin. Newton (1995) in a case study of task-based interaction found that negotiated words are retained better than non negotiated words during communicative activities. Joe (1995) discovered that tasks demanding higher levels of attention, retrieval, and generation led to a significantly higher level of vocabulary learning. In addition, Paribakht and Wesche (1997) underlined that task demands, learner attention, and learning outcome are closely related. In their study, they found that words are better retained if they are practiced after a reading task. That is, vocabulary retention in such exercises will be better than words that receive additional exposure in texts. Almost all of these authors focus on tasks that require a deeper level of processing.

In the cognitive psychology field, Craik and Lockhart (1972) stated that remembering information depends on not only attention and rehearsal but also on the levels of processing. They propose that if the information is processed at a deeper level, it will stay in memory for a longer period of time. That is, long retention depends on the depth of the processed information. Their focus was simply on orthographical, acoustic, and semantic information. This theory was expanded by Craik and Tulving (1975) by adding the notion of elaboration. They asserted that if new information is connected to the old information, it will lead to more robust memory traces. Baddeley (1999) referred to the term "richness" and argued that if a word is enriched by a variety of properties (denotation, connotation, association, use, usage, and general meaning), it will be remembered better.

This concept was applied to the EFL/ESL field by Laufer and Hulstijn (2001) who proposed the Involvement Load Hypothesis as an operationalization for the construct of depth of processing. According to Craik (2002), depth of processing is recently viewed to be related to elaboration and high degree of consciousness. This view is similar to the notion of levels of awareness proposed by Schmidt (1990; 1995; 2001) at two levels – noticing and understanding. Noticing needs attention and is necessary for intake, while understanding results in deeper learning.

According to Laufer and Hulstijn (2001), the notion of involvement includes three task-specific components – need, search, and evaluation. Need is a motivational component which is defined as “the drive to comply with task requirements” (p. 14). If the task requirements are externally imposed, need is moderate (+N), and if they are self-imposed, it is strong (++N). Search and evaluation are two cognitive components related to form-meaning relationships. Martinez-Fernandez, (2008) defines the search component as "the attempt to find the meaning of an unknown word when the meaning is not provided" which "may include a variety of strategies, such as contextual guessing, consulting a dictionary, or asking the teacher" (p. 211). According to Laufer and Hulstijn, (2001, p. 14), the evaluation component involves “a comparison of a given word with other words, a comparison of a specific meaning of a word with its other meanings, or combining the word with others in order to assess whether a word (i.e. a form-meaning pair) does or does not fit its context”. They use moderate evaluation (+E) for evaluating words that must fit in a given context and strong evaluation (++E) for words that must be combined with additional words in an original context. They suggest that vocabulary retention is influenced by the degree of the learner's involvement which, in turn, depends on the presence or absence of the three components. The following table clearly indicates the degrees of the components in the involvement load hypothesis.

THE DEGREES OF THE COMPONENTS IN THE INVOLVEMENT LOAD HYPOTHESIS

Components	Degrees of the Involvement Load	Explanations
Need	Index 0 (None)	The learner does not feel the need to learn the word.
	Index 1 (Moderate)	The learner is required to learn the word.
	Index 2 (Strong)	The learner decides to learn the word.
Search	Index 0 (None)	They do not need to learn the meanings or forms of the word.
	Index 1 (Moderate)	The meaning of the word is found.
	Index 2 (Strong)	The form of the word is found.
Evaluation	Index 0 (None)	The word is not compared with other words.
	Index 1 (Moderate)	The word is compared with other words in the provided context.
	Index 2 (Strong)	The word is compared with other words in self-provided context.

From The Involvement Load Hypothesis: An Inquiry into Vocabulary Learning, by Mayumi Tsubaki, as it appeared in http://cicero.u-bunkyo.ac.jp/lib/kiyo/fsell2006/EIBUN175_184.pdf

Hulstijn and Laufer (2001) conducted two parallel studies in the Netherlands and Israel to test their hypothesis. They randomly assigned three intact classes of advanced university foreign language learners to one of three conditions (gloss, fill-in, and writing) which varied in terms of task-induced involvement load. In the first condition [+N, -S, -E], subjects read a text with marginal glosses in their first language for ten targeted words. Then, they answered ten multiple-choice comprehension questions. In the second condition [+N, -S, +E], subjects read the same text and answered the same questions, they had to fill in ten blanks choosing the words from a list of 15 words with their first language translations and target language explanations. In the third condition [+N, -S, ++E], subjects wrote a composition using the targeted words. They were provided with grammatical categories, target language explanations, examples, and first language translation. Immediate and delayed posttests, in which the subjects provided a translation in their first language or an explanation in the target language, were administered to measure the targeted words. The targeted words had not been instructed. The results of their study showed that vocabulary retention in the writing condition was significantly higher than the fill-in and gloss conditions in both studies. They concluded that the hypothesis was fully supported in Israel and partially supported in the Netherlands. This study was designed to test the effectiveness of the involvement load hypothesis in an EEL setting.

III. METHOD

A. Research Design

Since random assignment was not possible, the nonequivalent group, pretest-posttest design was employed in this study. That is, subjects were tested in existing groups. The following diagram summarizes this quasi-experimental design in which the dotted line represents non-equivalent groups. Both groups are measured before and after treatment. Only one group receives the treatment. In this diagram GA and GB stand for experimental and control groups respectively. O1 and O3 stand for the tests before applying the treatment. O2 and O4 stand for the tests after the treatment and X stands for treatment.

GA	O1	X	O2
GB	O3		O4

B. Subjects

Sixty Iranian pre-intermediate EFL learners participated in this study. All subjects were female non-English major students at girls' professional-technical institute in Neishaboor. Two intact classes of third-semester students were selected for the experiment. There were 30 students in each class, majoring in computer field and food science. Originally, the two classes consisted of 69 students but 9 were excluded because of their prior knowledge and outside exposure to the targeted words as well as their failure to attend all sessions. A coin was tossed to choose the experimental and control groups. The computer class was considered as experimental group and the food science class as control group. A pre-test was administered to capture the initial differences between the two groups before starting the treatment. The age range of students was between 19 to 21 years old.

C. Instrument

The input passages were selected from pre-intermediate level "Select Readings" by Lee and Gunderson (2001). The subjects were asked to choose unknown words from a list of 100 words from the reading material. A total of 45 target words that were unfamiliar to both groups were chosen for the study. The target words consisted of 18 verbs, 14 adjectives, and 8 nouns.

A 40-item multiple-choice vocabulary pre-test was developed by the researchers. The items were designed based on the reading texts and the supplementary quizzes of the above mention book. The pre-test was administered to capture the initial differences between the two groups before commencing the main study. Since the time interval between the pre-test and post-test was long enough (about 7 weeks), the same pre-test with some changes in its arrangement was administered to the subjects as the post-test. For the pilot test, a class similar to those of this study responded to the items and helped the researcher establish the reliability of the test which was estimated 0.70 through KR-21 formula.

$$k_r = \frac{N}{N-1} \left(1 - \frac{M(N-M)}{NV}\right) = \frac{40}{39} \left(1 - \frac{9.033(40-9.033)}{40*22.440}\right) = 0.70601$$

D. Procedure

After administering the pre-test to capture the initial differences between the two groups, the experiment was carried out in seven sessions. Each group attended the class for a period of one and a half hours each week. One week after the last session, the participants completed the 40-item post-test. The degree of task involvement load in each group is described below.

In line with the Hulstijn and Laufer's (2001) research, this study randomly assigned the two intact classes of learners to two different conditions which varied in terms of task-induced involvement load. In the first condition [-N, -S, -E], the subjects in the control group read a text with marginal glosses in the target language irrelevant to the questions. Then, they answered the true or false and multiple-choice comprehension questions. In the second condition [+N, +S, ++E], the subjects in the experimental group read the same text with marginal glosses in the target language relevant to

the questions and answered the fill in the blanks items by choosing the words from a given list and wrote a sentence using the targeted words.

The need component, which is operationally defined as “the drive to comply with task requirements” (Laufer and Hulstijn, 2001, p. 14), is moderate (+N) in the experimental group and zero (-N) in the control group in this study. The search component is operationally defined as “the attempt to find the meaning of an unknown word when the meaning is not provided” (Martinez-Fernandez, 2008, p. 211) in this study. Based on this definition, it is moderate (+S) in the experimental group and zero (-S) in the control group. The evaluation component involves “a comparison of a given word with other words, a comparison of a specific meaning of a word with its other meanings, or combining the word with others in order to assess whether a word (i.e. a form-meaning pair) does or does not fit its context” (Laufer and Hulstijn, 2001, p. 14). According to this definition, the evaluation component in this study is zero (-E) for the control group but strong (++E) for the experimental group because the words are combined with additional words in an original context.

After introducing the topic and giving the students clear instructions on what they will have to do for their task, the teacher read the passage aloud once for both groups. In the control group class, the learners were asked to read and work on the passage in groups and were given enough time to do the true or false and multiple-choice comprehension questions. The involvement load was zero because all the questions could be answered without referring to the marginal glosses. That is, the task induced no need (the subjects did not need the target words to answer questions). It induced no search (the target words had already been glossed for them). And it induced no evaluation (the subjects did not have any chance to make comparisons).

As to the experimental group, the learners were asked to read and work on the passage in groups and were given enough time to do the fill in the blanks and sentence writing exercises which could only be correctly answered with the understanding of the unknown words. The involvement load was high because the questions could only be answered with referring to the marginal glosses. That is, the need component was moderate because it was imposed by the task. The search component was moderate because the target words had only been glossed in the target language for students. And the evaluation component was strong because learners had to use additional words in combination with the new word in a sentence.

IV. STATISTICAL ANALYSIS AND RESULTS

Before analyzing the data, a few preliminary steps were taken to ensure the validity and reliability of the results. To answer the research question, the raw scores taken from the pre-test and post-tests were submitted to the computer software Statistical Package of Social Sciences (SPSS version 15), using t-test. Paired-samples t-test was used to see whether there was a statistically significant difference in the mean scores for Time 1 (prior to the intervention) and Time 2 (after the intervention) of the same group. Independent samples t-test was conducted to compare the possible differences between the means of the experimental and control groups based on the gain scores from the pretest and posttest. The following tables indicate the summary of t-tests.

TABLE 1.
THE T-TEST FOR THE EXPERIMENTAL AND CONTROL GROUPS (PRE-TEST)

Group	N	Mean	Std Dev	df	t	Sig.
Control	30	5.30	2.61	58	2.47	0.016*
Experimental	30	3.81	1.98			

*Sig. p < .05

Table 1 shows a significant difference between the means of the two groups.

TABLE 2.
THE PAIRED SAMPLES T-TEST FOR THE CONTROL GROUP (PRE-TEST & POST-TEST)

Group	N	Mean	Std Dev	df	t	Sig.
Pre-test	30	5.30	2.61	29	8.17	0.000*
Post-test	30	8.80	3.63			

*Sig. p < .05

Table 2 shows that there is a significant difference between the means of the control group before and after the treatment.

TABLE 3.
THE T-TEST FOR THE EXPERIMENTAL AND CONTROL GROUPS (POST-TEST)

Group	N	Mean	Std Dev	df	t	Sig.
Control	30	8.80	3.63	58	0.46	0.64*
Experimental	30	9.21	3.31			

*Sig. p < .05

Table 3 shows that there is no significant difference between the means of the two groups in the post-test.

TABLE 4.
THE PAIRED SAMPLES T-TEST FOR THE EXPERIMENTAL GROUP (PRE-TEST & POST-TEST)

Group	N	Mean	Std Dev	df	t	Sig.
Pre-test	30	3.81	1.98	29	9.27	0.000*
Post-test	30	9.21	3.31			

*Sig. p < .05

Table 4 shows that there is a significant difference between the mean scores of the experimental group in pre-test and post-test.

TABLE 5.
THE T-TEST FOR THE COMPARISON OF THE MEAN INCREMENT BETWEEN THE TWO GROUPS AFTER TREATMENT (POST-TEST)

Group	N	Mean	Std Dev	df	t	Sig.
Control	30	3.50	2.34	2.58	2.63	0.011*
Experimental	30	5.40	3.18			

*Sig. p < .05

As indicated in Table 5, there is a significant difference between the gain scores for the experimental group ($M = 5.40$, $SD = 3.18$) and the gain scores for the control group ($M = 3.50$, $SD = 2.34$; $t(2.58) = 12.00$, $p < .05$). This final result shows that the mean increment of the experimental group after treatment is more than the control group. That is, the participants who received the task with higher involvement load learned significantly more vocabulary than others. Since there is a significant difference between the means of the two groups, the null hypothesis (tasks with higher involvement load do not lead to better EFL vocabulary learning) is rejected. Accordingly, the results of this study show that vocabulary learning in the fill in the blanks and sentence writing condition is significantly higher than the true or false and multiple-choice comprehension questions condition. Therefore, the involvement load hypothesis is supported.

Since the focus of this research like the first study by Hulstijn and Laufer (2001) and the similar study by Kim (2011) was on word meaning, its findings regarding the effects of task-induced involvement cannot be generalized to other aspects of vocabulary learning (e.g. word-form learning). The results of this study generally confirm Hulstijn and Laufer's (2001) findings. However, the current study was not as strict as the above mentioned two studies due to the limitations concerning subject selection and test characteristics. Since the involvement construct can be operationally defined and investigated in different ways, devising tasks with different involvement loads will shed more light on its application.

V. CONCLUSION

The final result of the independent samples t-test for the comparison of the mean increment between the two groups after treatment from the posttest administration indicated that the experimental group who received the task with higher involvement load had a better performance than the control group who didn't. Thus, the findings suggest that vocabulary learning in the fill in the blanks and sentence writing condition is significantly higher than the true or false and multiple-choice comprehension questions condition. This result is in line with Hustijn and Laufer's (2001) research done among adult EFL learners in Israel and the Netherlands. Since the task with higher involvement load seems to have facilitated the process of vocabulary learning, its application can be suggested to reinforce the word retention in EFL/ESL contexts.

While this study was being conducted, higher involvement load tasks turned to be time-consuming. Therefore, if there is not enough time for doing such tasks, they won't probably produce desirable retention effects. Due to the fact that the same amount of time was allocated to two groups with different tasks, the application of the Involvement Load Hypothesis was not as rigorous as the original study by Hustijn and Laufer's (2001). Therefore, the results should be interpreted cautiously. In fact, the application of such tasks in the classroom is worth further research.

REFERENCES

- [1] Baddeley, A. D. (1999). *Essentials of human memory*. Hove, UK: Psychology Press.
- [2] Craik, F. I. M. (2002). Levels of processing: Past, present... and future? *Memory*, 10, (5/6), 305-318.
- [3] Craik, F. I. M., & R. S. Lockhart. (1972). Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671-684.
- [4] 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.
- [5] Hosenfeld, C. (1976) Learning about learning: Discovering our students' strategies. *Foreign Language Annals*, 9 (2), 117-129.
- [6] Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental learning. In. P. J. Arnaud & H. B'joint (Eds.), *Vocabulary and Applied Linguistics*. London: Macmillan, 113-125.
- [7] Hulstijn, J. H., M. Hollander, & T. Greidanus. (1996). Incidental vocabulary learning by advanced foreign language students: the influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal*, 80, 327-339.
- [8] Hulstijn, J., & Laufer, B., (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning*, 51 (3), 539-558.

- [9] Joe, A. (1995). The value of retelling activities for vocabulary learning. *Guidelines*, 17, 1-8.
- [10] Kim, Y. J. (2011). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language Learning: A Journal of Research in Language Studies*, 61, 100-140.
- [11] Laufer, B., & Hulstijn, J., (2001) Incidental vocabulary acquisition in a second language: the construct of task-induced involvement. *Applied Linguistics*, 22 (1), 1-26.
- [12] Lee, L & Gunderson, E. (2001). Select readings: Pre-intermediate level. New York: Oxford University Press.
- [13] Martinez-Fernandez, A. (2008). Revisiting the Involvement Load Hypothesis: Awareness, Type of Task and Type of Item. In *Selected Proceedings of the 2007 Second Language Research Forum*, ed. Melissa Bowles, Rebecca Foote, Silvia Perpinan, and Rakesh Bhatt, 210-228. Somerville, MA: Cascadilla Proceedings Project.
- [14] Nation, I.S.P. (2001). Learning vocabulary in another language. Cambridge: Cambridge University Press.
- [15] Newton, J. (1995). Task-based interaction and incidental vocabulary learning: A case study. *Second Language Research*, 11, 159-177.
- [16] Paribakht, T. S., & M. Wesche. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In Coody, J. & T. Huckin (Eds.), *Second Language Vocabulary Acquisition*. Cambridge: Cambridge University Press, 174-200.
- [17] Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129-158.
- [18] Schmidt, R. (1995). Consciousness and foreign language learning. In R. Schmidt (Ed.), *Attention and Awareness in Foreign Language Learning* (pp. 1-63). University of Hawai'i at Manoa: Second Language Teaching and Curriculum Center.
- [19] Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.



Mohammad Reza Ghorbani was born in Bojnord, Iran, in 1969. He finished his BA in English Literature in Tabriz University in 1995, his MA in Teaching English as a Foreign Language (TEFL) in Isfahan University in 1999, and obtained his PhD in Teaching English as a Second Language (TESL) from the Faculty of Educational Studies, Universiti Putra Malaysia (UPM), Malaysia, in 2008.

He is currently a member of academic staff and assistant professor of Teaching English as a Second Language (TESL) at the Department of English, University of Bojnord, Bojnord, Iran. He has worked as an English as a Foreign Language (EFL) teacher and researcher in Iran, Japan, and Malaysia since 1990. He has published three books on educational issues and eleven articles in specialized international journals. He has also presented six papers in international conferences. His interests are English Teaching, Learning, Testing, and Evaluation.

Dr. Ghorbani was awarded PhD Fellowship at Universiti Putra Malaysia (UPM) in 2007. He is currently Director of Research and Technology at University of Bojnord, Iran. He has received the following honors: 2005: Teacher of the Year in Bojnord, Capital of Northern Khorasan, Iran 2009: Provincial Researcher of the Year in Iran 2010: Provincial Teacher of the Year in Iran 2011: Provincial Researcher of the Year in Iran

Maryam Rahmadoost was born in 1979 in Neishabor, Iran. She finished her BA in Teaching English as a Foreign Language (TEFL) in Ferdowsi University of Mashhad, and obtained her MA in TEFL the Islamic Azad University, Garmsar Branch, Iran. She is currently teaching at senior high school and university level in Neishabor.