

Efficacy of Task-Induced Involvement in Incidental Lexical Development of Iranian Senior EFL Students

Ali A. Pourakbari¹ & Reza Biria¹

¹ Faculty of English Language Department, Islamic Azad University, Khorasgan Branch, Isfahan, Iran

Correspondence: Ali A. Pourakbari, Faculty of English Language Department, Islamic Azad University, Khorasgan Branch, Isfahan, Iran. Tel: 98-939-289-7654. E-mail: alipourakbari2000@yahoo.com

Received: February 10, 2015 Accepted: March 12, 2015 Online Published: April 23, 2015

doi:10.5539/elt.v8n5p122 URL: <http://dx.doi.org/10.5539/elt.v8n5p122>

Abstract

One of the most significant current discussions in L2 classroom research has been related to the necessity of vocabulary in language learning. Undoubtedly, EFL learners with high levels of lexical knowledge perform better in writing and oral tasks. Accordingly, the present study sought to investigate the efficacy of Task-induced Involvement in Incidental Lexical Development of Iranian Senior EFL Students. For this purpose, based on the scores obtained from an Oxford Placement *Test (OPT)* administered to the population of senior EFL students at Khorasgan University, six samples, twenty five each, were selected and assigned to work with a list of English words by utilizing six different tasks. Each task was gauged by applying a different involvement load. Subsequently, a receptive and a productive vocabulary tests were administered as post-tests to specify the degree of learners acquisition of target words, through the role of incidental Task-induced involvement load. The results revealed that the group doing the task with the highest degree of involvement load obtained the best results on the vocabulary tests. Afterwards, the retention of unfamiliar words was claimed to be conditional upon the amount of involvement while processing these words.

Keywords: Iranian senior EFL students, intentional, lexical, task-induced involvement

1. Introduction

The notion of 'involvement load' includes both motivational and cognitive components. According to the Involvement Load Hypothesis proposed by Laufer and Hulstijn (2001), incidental tasks that trigger need, search and evaluation of the meaning of unfamiliar words will lead to higher vocabulary learning than those which do not trigger such processes. Studies suggest that incidental tasks with a higher degree of involvement load are more conducive to the type of processing that is crucial for learning. After saying that, studies investigating the effectiveness of different lexical intervention tasks during reading have led to conflicting results (e.g., Hulstijn, 1992; Watanabe, 1997). This hypothesis has important pedagogical implications, since it allows us to manipulate task features and predict what tasks will be more effective.

2. Background

2.1 The Involvement Load Hypothesis

Laufer and Hulstijn (2001) stated that the Involvement Load Hypothesis is an incidental vocabulary learning theory that formulated the criteria which explains why some specific tasks lead to better vocabulary retention than others. This construct comprised three principal components: 'need', 'search', and 'evaluation'.

The need component refers to whether, for task completing, the learner is supposed to know the meaning of the new words. Two levels of importance for need were offered: moderate and strong. Need is moderate when it is externally enforced by the teacher or the task, and strong when it is intrinsically imposed by the learner.

The search component signifies the endeavor of discovering the meaning of a new L2 word or discovering the L2 form of a word in L1. Unlike need, search may be present or absent. While learners attempt to discover the meaning of unfamiliar words to complete a task, the search is present; however, it is absent while such an attempt does not exist.

Evaluation entails reaching a conclusion about the meaning of a word during tasks, which can be moderate or strong. Evaluation is moderate while the learners are required to compare several lexical items with each other

(as in matching tasks), or compare different meanings of a lexical item in a provided text (as in a homonym). However, strong evaluation makes learners to combine new lexical items and create novel sentences. Combining all the three factors with their levels of importance in a task makes the task-induced involvement load.

Laufer and Hulstijn (2001) declared that tasks with higher involvement loads promoted better vocabulary retention than tasks with lower involvement loads. But how may we determine one task's involvement load in a numerical fashion? In order to compare different tasks with each other in a numerical fashion the involvement index was offered. In that, thenumerical weightfor the absence of a factor is marked as 0, a moderate presence of a factor as 1, and strong presence as 2 are appointed. Therefore, each task can have an involvement index of 0 (lowest index) to 5(highest index). In their hypothesis, Laufer and Hulstijn (2001) declared that any special task type (e.g., output) does not consider more effective than other type (e.g., input). They asserted that this is just the level of involvement load of a task which determine task's efficacy. In other words, they stated that two input and output tasks with the same load conditions will act equally on vocabulary acquisition. So, the equality of the involvement loads for different tasks types (e.g., input vs. output) calls for further research.

2.2 Empirical Studies on Involvement Load Hypothesis

Keating (2008) investigated whether the low-proficiency learners may also benefit from the more involving tasks, and whether the learners may gain the same word knowledge on passive and active tests. In order to have certainty about these questions, the low-proficiency learners of Spanish randomly completed one of the three tasks. After task completion and two weeks later, the learners' knowledge of target words was assessed through two passive and active tests. Partially confirming the Involvement Load Hypothesis, the results of both immediate and delayed passive tests reported that Task 2 and 3 resulted in higher retention scores compared to Task 1. However, Task 3 was not more effective than Task 2. On the other hand, the results of immediate active test which firmly supported the hypothesis revealed that learners in Task 2 and 3 promoted better word retention than those in Task 1, and learners in Task 3 also performed better than those in Task 2. In the delayed active test, however, learners in Task 3 did not perform much better than those in Task 1 or Task 2. In short, Keating's (2008) study claimed that the Involvement Load Hypothesis may be generalized to low proficiency learners and may also affect similarly the learner's passive and active word knowledge.

Most studies in the field of the Involvement Load Hypothesis indicate that time on task has not been well considered. Folse (2006) claimed that the efficacy of one task over another might be due to the length of time needed for task completion. Also, Keating (2008) argued that when time on task was taken into account, the benefits connected to more involving tasks faded. Therefore, the interpretation of results about the role of time on task draws a conclusion that may be we still need further research to test this hypothesis with a controlled time on task from the outset of the study.

3. Statement of the Problem

With respect to the number of L2 words to be learned, some researchers propose that 5,000 words is the lowest lexical necessity for non-specialized L2 learners of English to understand in general (Laufer, 1997). However, for the understanding of specialized and academic texts, 7,000 (Groot, 2000) or 10,000 word stock is required (Schmitt, 2000). In other words, 5,000 words is the prerequisite for communicative skills in a second or foreign language (Nation, 1992 cited in Prince, 1996). Accordingly, the first step for many foreign or L2 learners are to grab and memorize a large stock of vocabulary. However, the issue is how?

The accepted view among most researchers (e.g., Nagy & Herman, 1985) is that it is not possible for L2 learners to learn such a large stock of vocabulary merely through the explicit instruction of vocabulary. As Schmitt (2000) states, it would be very time-consuming and too laborious. The majority of word learning by L2 learners occurs incidentally (Krashen, 1989). Studies in this area reveal an extensive diversityof factors which were effective in promoting incidental word learning. Prince (1996) put emphasis mainly on learner factors with regard to incidental vocabulary learning. Hulstijn (1992) examined the impact of contextual cues such as marginal glosses; Knight (1994) considered the use of the dictionary as an issue affecting incidental vocabulary learning. Joe (1998) investigated the effects of text-based. Laufer (2001), Wesche and Paribakht (1997) examined word-focused. Ellis (1995) on interactional tasks and Loschky (1994) studied incidental vocabulary learning.

In each of these studies, one task was superior to another in terms of incidental vocabulary learning. To illustrate this superiority Craik and Lockhart (1972) indicated that the more effective task requires a deeper level of processing than the other task. It is notable that Craik and Lockhart's (1972) depth of processing has been criticized by Baddeley (1999), Nelson (1977), Tulving (1975), Laufer and Hulstijn (2001) for not having a clear-cut and simple definition about different levels of processing. Accordingly, the Involvement Load Hypothesis was formulated by Laufer and Hulstijn (2001) to provide a more clear-cut definition of processing

depth.

4. Research Questions

Laufer and Hulstijn (2001) claimed that any particular task type—be it input or output—is not considered superior or more effective, and that the only influential factor in task efficacy is the task's level of involvement load. Consequently, more research is needed to examine whether tasks with similar levels of involvement load but from different types—input vs. output—will have similar effects on vocabulary acquisition. To meet these two purposes, the researchers designed three receptive and three productive vocabulary tasks with varying involvement loads. In the light of the purposes of the study, the following research questions were posed:

- 1) On the basis of English receptive vocabulary tasks, will Iranian EFL learners obtain better retention of lexical in higher task load conditions compared to lower ones? If so, will the benefits of tasks hold up over time?
- 2) On the basis of English productive vocabulary tasks, will Iranian EFL learners obtain better retention of words in higher task load conditions compared to lower ones? If so, will the benefits of tasks hold up over time?
- 3) On the basis of English receptive and productive vocabulary tasks with the same levels of involvement index, will Iranian EFL learners obtain the same retention of new lexical on both types of tasks?

5. Method

5.1 Participants

Six groups, twenty five each, of Iranian senior Translation EFL students from Khorasgan University in Isfahan, Iran, homogenized by an Oxford Placement *Test (OPT)* were selected for this study. All of them English major and their first language was Persian. Each group was randomly assigned to one of the six experimental groups, in which three groups performed the receptive tasks, and the other three completed the productive tasks. Moreover, four participants' data were excluded from the study because they had the knowledge of more than two target words.

5.2 Materials

5.2.1 The Selected Lexical

The 10 unfamiliar selected lexical to the learners were chosen for examination from the GRE reading text by Kaplan. The selected lexical unfamiliarity was checked through a pilot study with a group of participants who did not participate in the experiment. These participants who had the same proficiency level of ours were given a list of 10 target words and asked to translate them. Out of 10 target words, the overall mean score was 0.2, which indicates the target words were unfamiliar within this proficiency level. Furthermore, the pre-knowledge of the participants in the main study was also checked in the immediate post-test. The chosen target words from the text were: (Cogent, Austere, Lament, Pedant, Loquacious, Vacillate, Repudiate, Capricious, Diffident, and Esoteric).

5.2.2 The Graphic Organizers

The graphic organizers designed by Kim (2011) were also used and modified according to the revised text. The participants in True-false, and matching task conditions were asked to answer the graphic organizers because they took less time than the other groups, as it was discovered in the pilot study.

5.2.3 Vocabulary Task Conditions

To address the first research questions, the researcher designed three receptive vocabulary tasks with varying involvement loads: Involvement 1=True-false; Involvement 2= Matching; and Involvement 3= Multiple-choice.

1) True-false task condition

In true-false task condition the Participants were asked to read the marginally glossed text and then complete the graphic organizers. Then, the participants were given the 10 True-false vocabulary tasks focused on the target words. In terms of the Involvement Load Hypothesis, this task induced a moderate need (the knowledge of target words was relevant to answering the tasks), but neither search nor evaluation. Thus, its involvement index was 1 (1 + 0 + 0).

2) Matching task condition

In Matching task condition the participants were also asked to read the text and completes the graphic organizers. After that, they were given 10 Matching vocabulary tasks focused on the target words. This task induced moderate need, moderate evaluation, and no search.

3) Multiple-choice task condition

Participants in the multiple-choice task condition were provided with the same text given to the last group; however, the text was not marginally glossed. The participants' task was to read the text by looking up the target words in a dictionary; then, they were given 10 multiple choice vocabulary tasks focused on the target words. This task induced moderate need and moderate evaluation (because four options in each of the multiple-choice vocabulary tasks must be assessed against each other). The search factor was also present here.

To address the second research questions, the researcher designed three productive vocabulary tasks with different involvement loads: Involvement 1= Short response; Involvement 2= Fill-in the blank; and involvement 3 = Sentence writing.

1) Short-response task condition

Participants in the Short-response task condition received the same marginally glossed text to read, and then to complete the 10 Short-response vocabulary tasks focused on the target words. Need was moderate, but search and evaluation were absent.

2) Fill-in-the-blanks task condition

In Fill-in-the-blanks task condition, participants were asked to read the same text and then complete the graphic organizers. This task induced moderate need, no search, and moderate evaluation.

3) Sentence writing task condition

Participants in the Sentence writing task condition received the same marginally glossed text, and were asked to read the text. Then, they were required to write L2 (English) sentences by using the 10 target words. The evaluation was strong because the participants were required to assess the target words within appropriate collocations in order to generate a new context.

5.2.4 Vocabulary Tests

The present study administered two immediate and delayed post-tests to assess the participants' learning. After two weeks, upon the completion of tasks, through a modified version of the Vocabulary Knowledge Scale (Wesche and Paribakht, 1997), the participants' knowledge of target words in all six task conditions unexpectedly was tested (Figure 1).

<p>Self-report</p> <p>1) I can't recall having seen this word before.</p> <p>2) I have seen this word before, but I can't remember what it means.</p> <p>3) I have seen this word before, and I think it means:</p> <p>4) I can use this word in a sentence:</p> <p><i>(If you do this section, please also do Section 3.)</i></p>
--

Figure 1. Modified VKS self-report categories

The scoring procedure of the modified VKS may be presented in this way:

(Category I): The word is not familiar at all. (1 point)

(Category II): The word is familiar but its meaning is unknown. (2 points)

(Category III): A correct synonym or translation is given. (3 points)

(Category IV): The word is used with semantic appropriateness in a sentence. (4 points)

(Category IV): The word is used with semantic appropriateness and grammatical accuracy in a sentence. (5 points)

It should be noted that wrong responses in self-report categories III or IV would lead to a score of 2. The overall possible test score for both post-tests was 10-50. The learners were given the 10 target words in the form of VKS on both post-tests and asked them to complete it. The learners were also asked to point out if any of the words were familiar to them before doing the task.

5.2.5 Procedure

This study was initiated by administering the treatment and the immediate post-tests. After two weeks the

delayed post-test were performed. On the treatment day, each of the six groups was asked to complete one of the following task conditions: True-false, Matching, Multiple-choice, Short response, Fill-in-the-blanks, or Sentence writing. In each group, the participants were asked to read the text and complete the 10 vocabulary tasks. To control the time on task, we also added a set of graphic organizers to the True-false, Matching, and Fill-in-the-blanks groups. Each of the six task conditions took 50 minutes to complete. Due to the nature of the study- incidental learning, the participants were not informed of the upcoming immediate or delayed post-tests because according to Laufer and Hulstijn (2001), test announcement is an indication of intentional word learning. Accordingly, after task completion, and two weeks later, the participants were unexpectedly given the immediate and delayed post-test in a modified form of VKS in order to measure the initial learning and retention of target words, respectively.

6. Data Analysis

The dependent variable for the two research questions was the scores of the immediate and delayed post-tests, and the independent variable was the level of involvement load. In order to examine the impact of the independent variable on the dependent variable, the VKS scores of both post tests were submitted to four, one-way ANOVAs. The Scheffe post hoc contrasts were then computed to locate significant differences among pairs. Unlike the first two, the third research question examined whether the type of vocabulary task affected the learning of new words when two different types of task (receptive or productive) with the same involvement loads were administered. The dependent variable in this question was the scores of both post-tests, and the independent variable was the type of vocabulary task at two levels: receptive and productive. Six independent samples *t*-tests were performed to compare the receptive tasks with the productive ones of the same load condition with alpha level set at 0.05.

6.1 Data Analysis of Three Receptive Tasks

The descriptive statistics of the three receptive vocabulary tasks in Table 1 demonstrate that, on both post-tests, the Multiple-choice group performed better than the Matching group, which, in turn, performed better than the True-false group. To determine if these differences were statistically significant, the scores of each posttest were then submitted to a one-way ANOVA.

Table 1. Descriptive statistics of the immediate and delayed Posttests for the three receptive vocabulary tasks

	N	Mean	SD	SE
Immediate Post-test				
True-false (1)	24	15.04	1.206	.213
Matching (2)	24	17.06	1.428	.297
Multiple-choice (3)	25	22.02	2.925	.584
Total	73	18.04	4.673	.365
Delayed Post-test				
True-false (1)	24	13.04	1.282	.241
Matching (2)	24	15.23	1.711	.338
Multiple-choice (3)	25	18.15	1.121	.211
Total	73	15.48	2.983	.314

Note. The indexes are in parentheses. The possible VKS scores in all three vocabulary tasks ranged from 10 to 50.

The results of both ANOVAs revealed a main effect for both the immediate [$F = 142.24, p < 0.001$] and the delayed post-test [$F = 124.263, p < 0.001$]. In fact, there was a significant difference among the tasks with different levels of involvement load on both post-tests. The results of two Scheffepost hoc tests also indicated that the Multiple-choice group significantly outscored both the Matching and the True-false groups, and the Matching group also significantly outscored the True-false group.

Comparing the means of the immediate with those of the delayed post-test for each of the three receptive vocabulary tasks, the results of three paired samples *t*-tests revealed that there was a significant decrease in the

mean scores of the delayed posttest for all the three receptive vocabulary tasks, that is, for the True-false task [$t = 9.365$, $p < 0.001$], for the Matching task [$t = 8.113$, $p < 0.001$] and for the Multiple-choice task, [$t = 11.482$, $p < 0.001$].

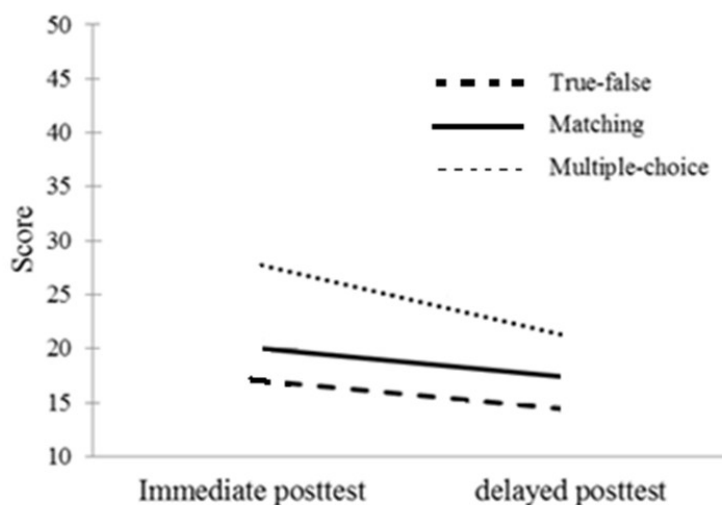


Figure 2. The scores of the immediate and delayed post-tests for the three receptive vocabulary tasks

6.2 Data Analysis of Three Productive Tasks

The descriptive statistics of the three productive vocabulary tasks in Table 2 suggested that the mean score of the Sentence writing group was higher than that of the Fill-in-the-blanks and the Short-response groups on both post-test; however, there was no great difference between the mean scores of the latter two groups on the delayed post-test.

Table 2. Descriptive statistics of the immediate and delayed Posttests for the three productive vocabulary tasks

	N	Mean	SD	SE
Immediate Post-test				
Short-response (1)	25	19.76	1.517	.295
Fill-in-the-blanks (2)	24	22.35	2.724	.414
Sentence writing (3)	25	25.84	2.113	.328
Total	74	22.65	3.427	.382
Delayed Post-test				
Short-response (1)	25	15.14	1.502	.289
Fill-in-the-blanks (2)	24	16.32	.703	.146
Sentence writing (3)	25	18.64	1.418	.254
Total	74	16.72	2.117	.186

Note. The indexes are in parentheses. The possible VKS scores in all three vocabulary tasks ranged from 10 to 50.

To determine the statistical differences among groups, two one-way ANOVAs were conducted. The ANOVA results indicated that significant differences were found among the three productive vocabulary tasks on both the immediate [$F = 39.862$, $p < 0.001$] and the delayed post-test [$F = 56.432$, $p < 0.001$]. The results of Scheffe tests also demonstrated that the Sentence writing group performed significantly better than the Fill-in-the-blanks and the Short-response groups on both posttests, but the Fill-in-the-blanks group performed significantly better than

the Short response group only on the immediate post-test.

Regarding the means of the immediate and delayed post-tests for each of the three productive vocabulary tasks, the *t*-tests results revealed a significant decrease in the mean score of the delayed post-test for the Short-response [$t = 10.204, p < 0.001$], for the Fill-in-the-blanks [$t = 10.284, p < 0.001$], and for the Sentence writing group [$t = 13.275, p < 0.001$].

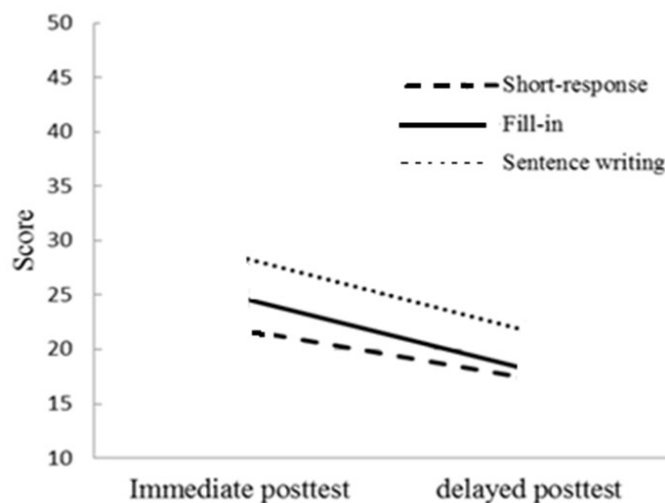


Figure 3. The scores of immediate and delayed post-tests for the three productive vocabulary tasks

The comparison between receptive and productive tasks:

Regarding the comparison between the true-false (load = 1) and the short-response group (load = 1), the results of two independent *t*-tests showed a significantly better performance for the Short response group on both the immediate [$t = -9.811, p < 0.001$] and the delayed post-test [$t = -6.026, p < 0.001$]. In the case of comparison between the matching (load = 2) and the Fill-in-the-blanks group (load = 2), the results of *t*-tests revealed that the Fill-in-the-blanks group performed significantly better than the Matching group on the immediate post-test [$t = -5.834, p < 0.001$]. However, this preference of the Fill-in-the-blanks group was not observed in the delayed post-test [$t = -1.428, p = 0.058 > 0.05$]. Unlike the last two pairs, the *t*-tests' results of the comparison between the multiple choice (load = 3) and the sentence writing group (load = 3) revealed that there was no significant difference between these two groups on both the immediate [$t = -.688, p = 0.382 > 0.05$] and the delayed post-test [$t = -1.236, p = 0.118 > 0.05$].

7. Discussion

The aim of the first two research questions were to assess whether tasks with a higher involvement load achieved better vocabulary retention than tasks with a lower involvement load while time on task was controlled across different groups. On both post-tests, the results of the first research question fully supported the Involvement Load Hypothesis. In that, the Multiple-choice group with the highest involvement load (3) produced better initial retention of target words than the Matching group with the lower involvement load (2), which, in turn, performed better than the True-false group with the lowest involvement load (1). However, the results of the second research question partly supported the Involvement Load Hypothesis. In that, the Sentence writing group with involvement load (3) performed significantly better than the Short response with involvement load (1) and the Fill-in-the-blanks group with involvement load (2) on both post-tests, but the Fill-in-the-blanks group performed significantly better than the Short-response group only on the immediate, but not the delayed post-test.

The third research question was posed to investigate Laufer and Hulstijn's (2001) claim that no particular task type—be it input or output—was considered superior or more effective, and that the only determining factor in task efficacy was the degree of involvement load that a task induced.

To meet this end, the researchers compared the receptive tasks with the productive ones of the same load condition. Contrary to the predictions of the Involvement Load Hypothesis, the results of the first pair

comparison revealed the better performance of the Short-response (a productive task) over the True-false (a receptive task) on both post-tests. Similarly, contrary to the Hypothesis, the Fill-in-the blanks (a productive task) performed significantly better than the Matching (a receptive task) on the immediate post-test; however, this preference of the Fill-in-the-blanks group was not observed on the delayed post-test.

Unlike the last two pairs, the results of the third pair comparison completely fulfilled the predictions of the Hypothesis in that the Sentence writing (a productive task) performed as well as the multiple choice (a receptive task) on both post-tests. Overall, the results of the first research question on both post-tests, and the results of the second research question on the immediate post-test were in harmony with those obtained in Hulstijn and Laufer's (2001) Hebrew-English Experiment and Keating's (2008) active word recall on the immediate post-test in that they all supported the Hypothesis. Similarly, the results of the second research question on the delayed post-test were exactly the same as those obtained in Hulstijn and Laufer's (2001) Dutch-English Experiment and Kim's (2011) first Experiment on the immediate post-test.

Nevertheless, the results of the third research question were considerably in conflict with the predictions of the Involvement Load Hypothesis. This hypothesis did not predict that any output task would lead to better results than any input task when they both had the same involvement load. On the contrary, we found that despite the involvement load induced by the task, the type of task was also effective in learning new words. In other words, two different types of tasks (receptive and productive) with the same level of involvement load might not have the same results in L2 vocabulary retention. The results of the study also showed a significant decrease in the performance of all six groups on the delayed posttest.

This explanation provided support for Swain's (1985) Output Hypothesis which claimed that the act of production demanded deeper cognitive effort and could contribute more to word learning than the mere reading of a text which is an act of reception.

In general, the findings of this study clearly run counter to some of the previous studies (e.g., Hulstijn; Laufer, 2001; Keating, 2008; Ellis, 1995; Webb, 2005) which claimed that controlling for time on task would diminish the effect of more involving tasks on vocabulary learning. However, similar to Kim (2011), we found that even if the time on task was controlled across different groups, the more involving tasks would perform better than the less involving ones in vocabulary scores.

8. Conclusion

In short, it can be concluded from the evidence of this study that task-induced involvement in incidental lexical development is a major factor of task efficacy in vocabulary learning. However, in testing the hypothesis with different types of tasks, the involvement load is not the only factor of task efficacy; the task type also has some role in vocabulary retention.

References

- Baddeley, A. D. (1999). *Essentials of human memory*. Hove: Psychology Press.
- Coady, J., & Huckin, T. (1997). *Second language vocabulary acquisition: A rationale for pedagogy*. Cambridge, 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.
- Donley, K. M., & Reppen, R. (2001). Using Corpus Tools to Highlight Academic Vocabulary in SCLT. *TESOL journal*, 10(1), 7-12.
- Ellis, R. (1995). Modified oral input and the acquisition of word meanings. *Applied Linguistics*, 16, 409-441. <http://dx.doi.org/10.1093/applin/16.4.409>
- Established and strengthened. (2005). *Reading in a Foreign Language*, 17(2), 95-124.
- Folse, K. S. (2006). The effect of type of written exercise on L2 vocabulary retention. *TESOL Quarterly*, 40(2), 273-293.
- Groot, P. J. M. (2000). Computer assisted second language vocabulary acquisition. *Language Learning & Technology*, 4(1), 60-81.
- Hirsh, D., & Nation, P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language*, 8(2), 689-696.

- Huckin, T. (1997). *Second language vocabulary acquisition: A rationale for pedagogy*. Cambridge: Cambridge University Press.
- 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: Macmillan.
- Hulstijn, J., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning, 51*(3), 539-558.
- Jacobs, G. M., Dufon, P., & Hong, F. C. (1994). L1 and L2 vocabulary glosses in L2 reading passages: Their effectiveness for increasing comprehension and vocabulary knowledge. *Journal of Research in Reading, 17*(1), 19-28.
- Joe, A. (1998). What effects do text-based tasks promoting generation have on incidental vocabulary acquisition? *Applied Linguistics, 19*(3), 357-377.
- Keating, G. (2008). Task effectiveness and word learning in a second language: The involvement load hypothesis on trial. *Language Teaching Research, 12*(3), 365-386.
- Keating, G. (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. (2011). "The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition". *Language Learning, 61*(1), 100-140.
- Knight, S. M. (1994). Dictionary use while reading: The effects on comprehension and vocabulary acquisition for students of different verbal abilities. *The Modern Language Journal, 78*(3), 285-299.
- Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *Modern Language Journal, 73*(4), 440-464.
- Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess. In J. Coady, & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 20-34). New York: Cambridge University Press.
- Laufer, B. (2001). Reading, word-focused activities and incidental vocabulary acquisition in a second language. *Prospect, 16*(3), 44-54.
- Laufer, B., & Hill, M. (2000). What lexical information do L2 learners select in a CALL dictionary and how does it affect word retention? *Language Learning & Technology, 3*(2), 58-76.
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics, 22*(1), 1-26.
- Loschky, L. (1994). Comprehensible input and second language acquisition: What's the relationship? *Studies in Second Language Acquisition, 16*(3), 303-323.
- Nagy, W. E., & Herman, P. (1985). Incidental vs. instructional approaches to increasing reading vocabulary. *Educational Perspectives, 23*(1), 16-21.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Boston, MA: Heinle & Heinle Publishers.
- Nelson, T. O. (1977). Repetition and depth of processing. *Journal of Verbal Learning and Verbal Behavior, 16*(2), 151-171.
- Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In J. Coady, & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 174-200). United Kingdom: Cambridge University Press.
- Prince, P. (1996). Second language vocabulary learning: The role of context versus translations as a function of proficiency. *The Modern Language Journal, 80*(4), 478-493.
- Read, J. (2004). *Assessing Vocabulary*. Cambridge: Cambridge University Press.
- Richards, J. C. (2002). *Methodology in language teaching*. Cambridge: Cambridge University Press.
- Rott, S. (2005). Processing glosses: A qualitative exploration of how form-meaning connections are established and strengthened. *Reading in a Foreign Language, 17*(2), 95-124.
- Rott, S., & Williams, J. (2003). Making form-meaning connections while reading: A qualitative analysis of word processing. *Reading in a Foreign Language, 15*(1), 45-75.

- 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.
- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass, & C. Madden (Eds.), *Input and second language acquisition*. Rowley, MA: Newbury House.
- Watanabe, Y. (1997). Input, intake, and retention: Effects of increasing processing on incidental learning of foreign language vocabulary. *Studies in Second Language Acquisition*, 19(3), 287-307.
- Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in Second Language Acquisition*, 27(1), 33-52.
- Wesche, M., & Paribakht, T. S. (Summer 2000). Reading-Based Vocabulary Exercises in Second Language Vocabulary Acquisition. *The Modern Language Journal*, 84(2), 196-213.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).