

## **A TRIAL OF THE INVOLVEMENT LOAD HYPOTHESIS IN A MIXED-ABILITY ESP CLASS**

This study is another trial of the Involvement Load Hypothesis in task-based learning, stating that vocabulary acquisition and retention depend on the involvement load of a task. The novelty of this research lies in the hypothesis testing in a mixed ability ESP class and exploring the effects of involvement load added to a task on receptive and productive target word recall. Three tasks with different involvement indexes were given to a group of fifty-two students whose receptive and productive target word recall was tested immediately and two weeks later. The predictions of the hypothesis were confirmed, as the scores were the highest for the writing task, lower for the reading comprehension with a gap-fill and the lowest for the reading comprehension with words glossed. Most of the results comply with the hypothesis, while partial discrepancies draw our attention to the issue of time-on-task. This study introduces teachers of courses with a focus on vocabulary to some effective vocabulary learning task types which can accompany reading. Extensions left to future research are investigating the effect of time-on-task on task effectiveness and comparison of the effectiveness of two exercises with the same involvement load, but different output or input orientation.

**Key words:** vocabulary acquisition, involvement load, receptive word recall, productive word recall, task effectiveness

### **1. Introduction**

This paper investigates the effectiveness of three tasks with varying involvement loads on vocabulary acquisition in a mixed-ability ESP class. Besides testing the Involvement Load Hypothesis, the paper gives an insight into the effectiveness of these tasks for most students in a mixed-ability class and the effect they have on their productive and receptive knowledge of the target vocabulary.

---

\* [katarina.lazic@sfb.bg.ac.rs](mailto:katarina.lazic@sfb.bg.ac.rs)

It is necessary to devote a considerable part of a good ESP course to vocabulary acquisition, since “subject-matter vocabulary is part of the definition of what constitutes English for specific purposes“ (Bejan, 1989, p. 93). Although majority of word learning happens incidentally via extensive reading (Krashen, 1989), most studies on incidental vocabulary acquisition from reading find relatively small gains in vocabulary acquisition of about 1-5 words per text from a short text (of about 7000 words) (Laufer, 2005). Word learning is a cumulative process and acquisition and long-term retention of a new word require numerous encounters with it and rehearsal. Frequent enough exposure to the new word through reading would require far more time than available to teachers and students during language courses. Reading can be introduced in the form of certain task types or accompanied by certain types of vocabulary exercises with the aim of providing an opportunity for rehearsal and more frequent exposure to the newly acquired vocabulary during classes.

Therefore, this study aims at comparing certain task types which can either be added to reading or done instead of it in order to boost vocabulary acquisition. Moreover, it investigates and substantiates the postulation of the Involvement Load Hypothesis that vocabulary acquisition and retention are commensurate with the task-induced involvement, i.e. the involvement load composed of three components: need, search and evaluation (Hulstijn & Laufer, 2001a). In addition, it provides an insight into two more issues to which prior research in this tradition has failed to pay attention.

First, the study explores if tasks with a higher involvement induce better L2 lexical acquisition and retention than the ones with a lower involvement load in a mixed-ability ESP class, i.e. whether they are beneficial to most students regardless of their general language proficiency, for this group ranging from A2 to C1 CEF levels<sup>1</sup>.

Secondly, the study is focused on determining if tasks with different involvement load lead to distinctive gains in both receptive and productive word knowledge for the above-mentioned group. The study conceptually replicates the research Hulstijn & Laufer (2001b). Some time before Hulstijn & Laufer (2001a) formulated the motivational and cognitive construct of Task Induced Involvement occurring during incidental vocabulary acquisition, which led to the creation of the Involvement Load Hypothesis.

---

1 The Common European Framework of Reference for Languages

Participants in the study are freshmen students of the Faculty of Forestry in Belgrade in an academic mixed-ability ESP class. Three groups of students were given three different tasks with varying involvement load. Task 1 was a reading comprehension with words glossed and explained in the margin in L1 followed by comprehension questions. Task 2 was reading comprehension plus fill-in, with the target words replaced by blank spaces and given on a separate piece of paper with four other unrelated words. Task 3 was writing an original paragraph containing the target words. The students were given a list of words with explanations of the target words and four more previously known words in L1, and were asked to write an original paragraph containing the words.

The students' receptive and productive target word recall was tested twice, both immediately and two weeks later. The receptive word recall test was an English- Serbian word translation task, while the productive word recall test was a Serbian (L1) – English (L2) sentence translation task. The analysis of the data obtained through tests analyzed the performance of students in two groups. The group of lower (CEF levels A2 and B1) general language competence and of higher competence (CEF levels B2 and C1). After the testing the teacher divided the post-tests into two groups of higher and lower general language proficiency based on the teacher's records which were kept after the placement test held at the beginning of the term and their results were separately analyzed. The aim was to discover the pattern followed by each group in vocabulary acquisition under the same conditions. The scores of the receptive and productive tests for these two groups were submitted to ANOVA, which compared the effectiveness of the three tasks in each group.

Besides testing the Involvement Load Hypothesis, the paper speculates on task effectiveness in the light of the notions of incidental and intentional learning, receptive and productive knowledge of words, time-on-task and other fields tangential to this topic. Unfortunately due to the limitations, both physical research limitations and the format limitations of this study, some interesting issues which arose in the course of speculating on the above listed notions were omitted as an integral part of this paper but are highlighted as challenging suggestions for future research.

## 2. Literature review and related research

### 2.1. *Vocabulary acquisition from reading*

In order to successfully acquire new words from reading a student must know 98% of the words in the text (Hu & Nation, 2000; Nation, 2001). Extensive research has shown that only a small number of words can be acquired from exposure to texts without later vocabulary practice. Studies on incidental vocabulary learning from reading report very small gains of about (15% of unknown words) or one to five words per short text (up to 7000 words) (Hulstijn, 2003; Krashen, 1989). Krashen (1987, p. 23) noted that “new vocabulary should be supplied in enough frequency, and to help the acquirer determine the meaning.” Nevertheless, vocabulary learning depends on the repeated exposure to large quantities of input, and for the cumulative effect to take place a new word has to be seen ten times or more (Nation, 2001). Learning 2,000 words from reading only would take learners about 29 years (Zahar et al. in Laufer, 2006). On the other hand, teachers can offer only two, three or five hours of instruction per week and cannot involve such amounts of reading in the classroom context. For that reason, they need ways to stimulate vocabulary acquisition in a limited time during classes. The above mentioned learning can be increased by paying some deliberate attention to vocabulary (Hulstijn, 2003). Nation (2001, p. 238) noted that “second language learners should not rely solely on incidental vocabulary learning from context; there needs to be judicious attention to decontextualised learning to supplement and be supplemented by learning from context. Direct vocabulary learning and incidental learning are complementary activities”.

Numerous studies indicate that vocabulary acquisition is better when learners are involved in word-focused tasks than reading alone. Words guessed with some difficulty will be remembered better, because difficulty leads to processing effort and creates a more distinctive memory trace (Cho & Krashen, 1994; Luppescu & Day, 1993; Hulstijn, 1993; Laufer, 2003; Laufer, 2005). In these studies, increased involvement load generally implies a greater focus on form<sup>2</sup>.

---

2 This study does not pay attention to the focus on form approach in L2 vocabulary acquisition. For more information on this topic see the following papers, Laufer (2003), Laufer (2006), and Webb (2007).

All these studies suggest that new vocabulary should be introduced in tasks which involve the student in more processing effort. These tasks should leave a distinctive memory trace and serve as a more efficient and less time-consuming way to acquire new vocabulary compared to reading alone.

## 2.2. *Depth of processing and elaboration*

In the nineteen seventies cognitive psychologists Craik and Lockhart (1972) introduced the idea of the *depth of processing*<sup>3</sup>. They argued that the memory trace is a product of the processing of information. According to them, the persistence of a memory trace is a result of the depth of processing at which the information is analyzed. Craik and Tulving (1975) introduced the notion of *elaboration*. They argued that richness of encoding of the information is responsible for differences in the later recall. These two may be seen as preceding the idea of *task-induced involvement* and *involvement load*. However, that concept was challenged by many psychologists, because it was impossible to decide which information is processed at a deeper level than the other.

Numerous studies compare task types created to support incidental vocabulary acquisition. In a study by Hulstijn (1993), words whose meaning was guessed during reading led to better acquisition and retention than words explained by a synonym. Similarly, in the experiment which investigated reading and vocabulary acquisition, words which were looked up were better remembered than words which were not looked up (Cho & Krashen, 1994). In addition, the students who used dictionaries during reading achieved better results than the ones who did not use dictionaries (Luppesku & Day, 1993). Furthermore, a text with marginal glosses was more effective for vocabulary acquisition than the text without marginal

---

3 The level to which information is processed after being introduced into the memory. The concept is based on the rejection of the idea that information is stored in discrete memory stores. It adopts the view that an item enters the information-processing system and undergoes an increasing amount of processing, thereby moving deeper into the system. It is postulated that the retention characteristics of information are determined by the depth of processing, with deeper levels being associated with more abstract coding and greater retention. The notion of the depth of processing could not be discussed in detail in the current study. For more information see Craik & Lockhart (1972) and Craik & Tulving (1975).

glosses because the form-meaning connection in the readers' minds was strong, unlike in the text without glosses (Hulstijn, Hollander & Greidanus, 1996). Similarly, words negotiated in a discourse were better remembered than the nonnegotiated words, and in a study which investigated vocabulary acquisition after reading and retelling tasks, the exercise with a higher level of generation was more effective (Newton, 1995). Moreover, productive tasks, especially used in an original context, such as retelling in their own words, resulted in a better memorizing of words than in the case of non-productive tasks (Joe, 1995). Likewise, words learned in vocabulary exercises were more effective than words which had additional exposure in texts (Paribakht & Wesche, 2000).

All these studies suggest that a more involving task which engages the student in a deeper processing of lexical information is also a more effective one. However, there were no criteria which could grade the tasks in terms of depth of processing. The researchers needed an instrument to observe, measure and manipulate this component. For that reason, Hulstijn and Laufer developed their Involvement Load Hypothesis for L2 vocabulary learning in 2001.

### 2.3. *The Involvement Load Hypothesis*

Hulstijn and Laufer (2001a) tried to transfer the idea of the depth of processing to the depth of processing of lexical information and vocabulary acquisition. They studied vocabulary learning tasks and proposed the motivational-cognitive construct of *involvement*. The construct consisted of three components: need, search, and evaluation. *Need* is a motivational, non-cognitive dimension which depends on whether the students need to be familiar with the new word in order to do the task successfully. Need is treated as moderate when imposed by someone else and strong when it is self-imposed by the learners themselves and their eagerness to communicate. *Search* is the involvement of students in the search for an adequate translation of a word in a foreign language. This can be being engaged in looking up a word in a reference book e.g. a dictionary, or directly asking a competent person for help. *Evaluation* involves comparing the word with other words or distinguishing the specific meaning of a word from its other meanings, choosing the one which best fits the context, etc. When a student has to identify the word among a number of other words or select from several meanings of a word, the evaluation is moderate. However, if

the student has to use the word with other words in a new text or sentences, the evaluation will be treated as strong.

Table1 is an illustration of some tasks and their accompanying involvement load.

**Table1. Some tasks their involvement load**

Task	Need	Search	Evaluation
Reading with questions words glossed are relevant to the task	+	-	-
Read and fill in given words	+	-	+
Writing a composition	+	-	++
Notes: - = no involvement load + = moderate involvement load ++ = strong involvement load			

The factors of need search and evaluation can be absent or present. The combination of these factors and levels of their prominence comprise involvement load pertinent to a task. These features occur in combination and each of them has some influence on the involvement. Laufer and Hulstijn’s research and the Involvement Load Hypothesis imply that the greater the involvement load the more effective the learning.

A more precise estimate of a task is made by determining its *involvement index*. In that case absence of a factor is marked as (0), moderate presence as (1), and strong presence as (2).

#### 2.4. Experiments Supporting the Involvement Load Hypothesis

In many studies reading followed by word-focused tasks led to better results than reading alone. When given enough time for completion, writing with the use of new words was more effective than reading, while in another experiment, using new words in original sentences was more effective than reading three sentences with new words glossed (Webb, 2005). Likewise, using new words in original sentences was more effective than reading both when new words were glossed and looked up (Laufer, 2003). Similarly, unknown words acquired through a negotiation

of the target vocabulary are better learned than those learned through reading (de la Fuente, 2002).

There are several direct tests of the Involvement Load Hypothesis. Empirical evidence for the Involvement Load Hypothesis was given in a study by Hulstijn and Laufer (2001b). Two parallel experiments were carried out in two countries, Holland and Israel. In both experiments, Task 3 (composition writing) was the most effective in vocabulary acquisition. In accordance with the Involvement Load Hypothesis, the Hebrew – English experiment found that Task 2 (reading comprehension plus fill-in with marginal glosses) was more effective than Task 1 (reading comprehension with marginal glosses). Conversely, in the Dutch-English experiment, Task 2 did not produce significantly higher results than Task 1, which was not consonant with the hypothesis.

An empirical study of the Involvement Load Hypothesis in vocabulary acquisition in EFL listening (Jing & Jianbin, 2009) applies the hypothesis to incidental vocabulary acquisition in EFL listening comprehension. The experiments were carried out in three different classes whose learners did three different listening exercises with graded involvement load. Vocabulary acquisition and retention showed significant difference and superiority of more involving tasks. So, once again, the Involvement Load Hypothesis was confirmed.

A study by Keating (2008) tests and substantiates the Involvement Load Hypothesis investigating the focus on form component in vocabulary exercises. The study was not in full agreement with the hypothesis as Task 1 (reading with marginal glosses) and Task 2 (comprehension text plus fill-in) appeared to be more effective than Task 3 (sentence writing) in the delayed post-tests.

A number of recent studies, including Huang et al. (2012), Soleimani et al. (2014) and Douglass (2016) also support the predictions of the involvement load hypothesis, but suggest some additional factors that can have an impact on vocabulary acquisition in addition to involvement load.

#### *2.4. ESP Teaching and Vocabulary Acquisition*

The methodology of teaching vocabulary of English for specific purposes (ESP) shares much with the methodology of general vocabulary teaching (Bejan, 1989). Subject-specific vocabulary items of ESP vary



in their frequency and the most frequent ones which appear in various original texts help in establishing priorities in vocabulary teaching and production of adequate teaching materials. What vocabulary to focus on should be determined by the needs of the learners and the usefulness of vocabulary items. Meara (2002, p. 37) noted that “one way of measuring the usefulness of items is to discover their frequency and range in a relevant corpus”. When they establish the relevant corpus, teachers need effective ways and means to teach that vocabulary.

This study investigates the effectiveness of certain task types with different involvement load. The most effective task types from this research can contribute to the efforts of ESP teachers in the creation of adequate teaching materials by serving as a model regarding their amount of involvement or task type.

### **3. The current study**

The previously mentioned studies which preceded and inspired this research have certain limitations. The students who participated in previous studies were advanced learners of English and the studies did not test the predictions of the Involvement Load Hypothesis in a mixed-ability class. Although the students of higher general language competence are expected to show results which are consonant with the Involvement Load Hypothesis, the predictions of the hypothesis have not yet been tested in mixed ability-classes, i.e. students of lower general language competence and the groups consisting of both students of higher and lower general language competence which should function as a whole. The direct empirical test of the Involvement Load Hypothesis by Hulstijn and Laufer tests only the students’ receptive knowledge. The students’ productive knowledge is not tested. For that reason this study is a conceptual replication of Hulstijn and Laufer’s 2001 study which tries to find the answer to the following research questions:

1. Does adding involvement load to certain task types designed to support incidental vocabulary learning lead to gains in vocabulary acquisition and retention for most students in a mixed-ability ESP class?

2. Does adding involvement load to certain task types designed to support incidental vocabulary learning contribute to students’ receptive and productive knowledge of words?

## **4. Method**

### *4.1. Participants*

The study of task effectiveness in vocabulary acquisition was performed at the Faculty of Forestry in Belgrade. The participants involved in the study were freshmen students in a mixed-ability ESP class. The observed class consisted of male and female students aged 18-19, who belonged to the Department of Landscape Architecture and the Department of Wood Processing, forming a single group of about 100 students. At the very beginning of the term all freshmen students' general language proficiency was tested using a quick pen and paper placement test, and it was found that the students' knowledge varied from A2 to C1 CEF levels. Through this test the teacher became familiar with every individual's level of general language proficiency and a record of it was kept in the teacher's files.

Freshmen students in the investigated group at the Faculty of Forestry in Belgrade come from all over the country. Even though a foreign language, which is in most cases English, is a compulsory subject in all grammar schools and high schools in Serbia, students' knowledge when they reach university tends to vary greatly depending on the area where they come from, particular schools they attended, and most importantly, their extracurricular engagement with English. A large number of students of the Faculty of Forestry in Belgrade come from off-beat rural areas, where they had no opportunity to attend foreign language courses apart from their lessons at school. On the other hand, some other students in this group come from urban areas where extracurricular engagement with English is available and common for children and teenagers of all ages starting from the pre-school age. This usually offers additional exposure and consequently better knowledge. When the students' general language proficiency was tested, it was clear that the group was a large mixed-ability class. Consequently, the activities included in the course needed to be carefully chosen to enable improvement and successful learning for most students in such a heterogeneous group of learners. Finally, the conclusions reached in this research aim at finding effective task types which prove functional for vocabulary acquisition in such classes.

The experiment was held during one of the regular lessons of the course. Before the experiment, the students were informed that they would

be taking part in a research and told that the results of the immediate and delayed tests would not have an effect on their course grade.

#### 4.2. Target Words

The target words were chosen from an unfamiliar subject-matter text on soil maps. The lesson on soil maps and mapping and the text used in the study are a part of the regular curriculum of the ESP course at the Faculty of Forestry in Belgrade. During research design, students' familiarity with the target words was prevented by avoiding some units from the textbook (e.g. the unit on glass-houses) whose subject-specific vocabulary (items such as *humidity*, *pest*, *artificial fertilizer* etc.) may be familiar to students due to the media coverage of the topic or exposure to the vocabulary at school. The chosen unit on soil maps contains subject-specific vocabulary which is less likely to be familiar to students before the research than the subject specific vocabulary of most units from the textbook as most of the students were likely to learn about the topic for the first time. The target words are two verbs, two adjectives and four nouns and they are the following: *soil map*, *mapping*, *survey*, *criss-cross*, *exacting*, *overlay*, *scale*, and *customary*.

Before the experiment a test of familiarity with the words was given to a second year class which had never had the lesson on soil maps. The control group used for this test of familiarity with the target words was supposed to be very similar to the group used in the research, as the students belonged to the same departments and had never read the text on soil maps before. The group consisted of students of a similar profile, as they were also a mixed-ability class, at the same faculty and departments, with a different English language learning history. The mean score of the familiarity test was less than 1.5 words recognized out of the 8 target words. The prediction was that students in the research groups were a similar group, as unfamiliar with the target words as the control second-year group.

#### 4.3. Tasks

In this research *Task 1* was reading comprehension with marginal glosses of the target words followed by comprehension questions. The target words were printed in bold in the text and an equivalent to each word in L1 (Serbian) was given in the margin. The students were given 15 min-

utes to read the text and glosses. The leaflets with tasks were then taken away from them. After a short break, they were given seven True or False comprehension questions on the topic of soil maps and mapping. The students were supposed to answer the questions without access to the text of the tasks. The involvement index calculated from three components, need, search and evaluation for this task was 1. The need component was moderate, because it was imposed by the task and its value was 1. The components search and evaluation were absent and the involvement index for this task was derived using the following equation: need 1 + search 0 + evaluation 0 = 1.

*Task 2* was a reading comprehension plus fill-in, where bold target words were replaced by blanks and given on a separate piece of paper with four other unrelated words. The words were followed by a definition in L2 (English) plus an example sentence and a translation of the word in L1 (Serbian) in brackets. The students were given 20 minutes to study the words and fill in the blanks. After a short break, the task was followed by the same seven comprehension questions as Task1. The need in this exercise was moderate or imposed by the task, the evaluation was moderate since the students needed to compare the words given in order to fill in the blanks, and its value was 1 and the search component was absent or 0. The involvement load index was 2 (need 1 + search 0 + evaluation 1= 2).

*Task 3* was composition writing with the words given and the instruction said that the students had to use the words given. Target words were the same as in Task 2 with four more unrelated words and the same definitions in L2 (English) plus an example sentence and a translation of the word in L1 (Serbian) in brackets. Students had 25 minutes to complete the task. Task 3 had the highest involvement index. Need was moderate, imposed by the task, search was absent, and evaluation was strong, since the students had to use the words in sentences of their own. The involvement index was 3 (need 1 + search 0 + strong evaluation 2= 3).

#### *4.4. Procedure*

Although the tests were announced, which is one of the features of intentional learning (Keating, 2008), the students knew that the testing would not affect their course grade and we cannot say that learning which took place was intentional and not incidental only on the basis of that. Furthermore, it is doubtful that they were deliberately trying to engage in

learning strategies since there was no difference if their test results would be good or bad. They were also aware that the text was a part of the regular curriculum and that the vocabulary learned would be as important to them as any other lesson learned during the term. The students were also informed that the post-tests were going to be delivered twice, immediately and two weeks later.

The number of students who completed Tasks 1, 2 and 3 was much larger than the final number of participants whose scores were analyzed in the study. The reason was that as both lessons were regular lessons, the number of participants varied. Some of them were absent in the delayed post-test lesson two weeks later, whereas some other students who were not present in the first research lesson appeared in the delayed post-test lesson.

Although the tasks were done by a much larger number of students, only the ones that belonged to the students who were present in all research lessons were analyzed. The numbers of those students were the following: for Task 1  $n=17$ , Task 2  $n = 17$  and for Task 3  $n= 18$ . Means for each of the groups were calculated and they were compared using one-way ANOVA, and the posthoc Tukey test was applied for further data analysis.

Although the students performed Tasks 1, 2 and 3 in random mixed groups, the teacher would find a student's name in the records and put them in the lower or the higher general language proficiency group. Students at CEF levels A2 and B1 were classified as the lower general proficiency group and Students at B2 and C1 levels as the higher general proficiency group. After the classification, it was clear that Task 1 was performed by nine students from the more advanced group and eight students from the lower proficiency group. Task 2 was performed by nine students from the more advanced group and eight students from the lower proficiency group and Task 3 by nine students from the lower proficiency group and nine students from the more advanced group. Means for each of the groups were calculated, and they were submitted to one-way ANOVA for comparison.

#### *4.5. Testing and Scoring*

The participants were given a list of eight target words and were asked to provide their translation into L1 (Serbian). The immediate post-test was delivered during the same lesson as the tasks, after a short 10-minute break. The delayed post-test was delivered in a regular lesson two weeks later.

The productive word recall post-test was delivered after a short break following the receptive word recall test in the form of a Serbian-English sentence translation task. Each of the eight sentences contained one of the target words. In the delayed post-test the sentences were changed but they still contained the target words. The tests were scored on a 0-8 point scale. A semantically equivalent word and not the exact target word would receive half a point.

## 5. Results

### 5.1. Receptive word recall

Scores of the immediate and delayed receptive word recall tests for groups 1 and 2 were submitted to one-way ANOVA, an analysis between groups with Task (reading with glosses, reading and fill-in or composition writing) as the between-subjects factor. Table 2 shows N- the number of participants, SD- standard deviation, and the immediate and delayed MEAN- (mean scores) for the receptive word recall tests of Group 1.

**Table 2. Receptive word recall tests for Group 1**

Tasks	Immediate		Delayed		
	N	Mean	SD	Mean	SD
Task 1	9	3.22	1.1331154	1.44	0.5270463
Task 2	9	2.89	0.993808	2.33	0.7071068
Task 3	9	5.33	1.1547005	3.33	0.942809

Table 3 shows N- the number of participants, SD- standard deviation, and the immediate and delayed MEAN-(mean scores) for the receptive word recall tests of Group 2.

**Table 3. Receptive word recall tests for Group 2**

Tasks	Immediate		Delayed		
	N	mean	SD	mean	SD
Task 1	9	1.67	0.6666667	0.89	0.781736
Task 2	9	1.56	0.6849349	1.56	0.5270463
Task 3	9	2.89	0.7370277	2.11	0.8748898

In both groups mean score in the immediate receptive word recall test was higher for the participants who performed Task 1 than for the ones who performed Task 2, unlike in the delayed test in which participants who performed Task 2 had a higher mean score than the ones who performed Task 1. In both tests, the participants who completed Task 3 outperformed the ones in the other two groups.

*5.2. Productive Word Recall*

Table 4 shows N- the number of participants, SD- standard deviation, and the immediate and delayed MEAN- (mean scores) for the productive word recall tests of Group 1.

**Table 4. Productive word recall tests - Group 1**

Tasks	Immediate		Delayed		
	N	mean	SD	mean	SD
Task 1	9	1.67	0.6666667	0.89	0.781736
Task 2	9	1.56	0.6849349	1.56	0.5270463
Task 3	9	2.89	0.7370277	2.11	0.8748898

Table 5 shows N- the number of participants, SD- standard deviation, and the immediate and delayed MEAN- (mean scores) for the productive word recall tests for Group 2.

**Table 5. Productive word recall tests- Group 2**

Tasks	Immediate		Delayed		
	N	mean	SD	mean	SD
Task 1	9	1.11	0.6372093	0.44	0.503077
Task 2	9	0.89	0.4829039	0.67	0.6666667
Task 3	9	2.33	0.942809	1.44	0.6849349

Once again the immediate productive word recall test showed that participants who completed Task 1 outperformed the participants who performed Task 2. Conversely, the delayed productive word recall test showed that students who completed Task 2 were superior to those who completed

Task 1. Students who completed Task 3 had the highest mean score in both the immediate and delayed productive recall post-tests.

The mean scores of groups 1 and 2 were submitted to one-way ANOVA, and the analysis of the scores of the immediate receptive word recall test for group 1 revealed the following results: the main effect of Task,  $F$ - statistic = 11.698630 and  $p$ -value = 0.00078 < 0.5, which means very strong evidence against the null hypothesis. The scores of the immediate receptive post-test for group 2 reached the following values:  $F$ - statistic = 8.9478261 and  $p$  = 0.00204 < 0.5. This suggests very strong evidence against the null hypothesis.

The scores of the delayed receptive post-test for group 1 also indicated very strong evidence against the null hypothesis when submitted to ANOVA. Interestingly, the scores of the same delayed test for Group 2 revealed the value of  $F$ - statistic= 0.5975236, but  $p$ = 0.55935 > 0.5, which suggests moderate evidence against the null hypothesis.

When mean scores of the immediate and delayed active word recall tests were submitted to ANOVA, all cases revealed evidence against the null hypothesis, which were very strong except for the delayed productive word recall test for Group 2, which found moderate evidence against the null hypothesis ( $F$ - statistic= 4.6589714 and  $p$ -value = 0.02105). As most tests revealed strong evidence against the null hypothesis, the next step was to determine the significance.

### 5.3. *The Posthoc Tukey Test*

The Posthoc Tukey Test was applied after the ANOVA to determine which groups differ significantly from each other.

The scores of the immediate receptive word recall post-tests for Group 1 are HSD=1.1809171 < M1-M2 =1.935648, M1-M3 = 2.136785 and M2-M3= 2.432897. The scores of Group 2 revealed similar results, and in both cases, the scores of Task 1 were significantly higher than the scores of Task 2 and the scores of Task 3 were significantly higher than the scores of Task 1.

In the delayed receptive Group 1 post-test the results are: HSD= 0.785453 < M1-M2 =0.8837489; M1-M3 = 1.8358889 and M2-M3= 2.3365783. Table 6 shows the application of the Posthoc Tukey Test on the delayed receptive word recall post-test of Group 1.



**Table 6. Posthoc Tukey test of the delayed receptive word recall post-test for Group 1**

	1		2		3	
	del.	$x^2$		$x^2$		$x^2$
	1	1	2	4	3	9
	2	4	3	9	2	4
	2	4	3	9	4	16
	1	1	3	9	3	9
	1	1	2	4	5	25
	2	4	3	9	2	4
	1	1	2	4	4	16
	1	1	2	4	3	9
	2	4	1	1	4	16
SUM	13	21	21	53	30	108
SUM <sup>2</sup>	169		441		900	
MEAN	1.4376821		2.3347923		3.3228675	
M1 vs. M2	3.4641016	q	M1 - M2	0.8837489	q	HSD
					3.553	
M1 vs. M3	7.3612159	q	M1 - M3	1.8358889	3.061	0.785453
M2 vs. M3	3.8971143	q	M2 - M3	2.3365783		

The scores of Task 2 were significantly higher than the scores of Task 1, and the scores of Task 3 were significantly higher than those of Task 2.

The productive word recall tests revealed the same model with differences between the mean scores of groups higher than the HSD value, which indicated significance. The scores of the immediate productive word recall post-test for Group 1 are  $HSD = 0.750681 < M1 - M2 = 0.98675$ ,  $M1 - M3 = 1.239725$  and  $M2 - M3 = 1.342837$ . In the delayed productive post-test of Group 1  $HSD = 0.758453 < M1 - M2 = 0.83935$ ,  $M1 - M3 = 1.43733$  and  $M2 - M3 = 2.33496$ , which meant that scores of Task 2 were significantly higher than those of Task 1 and scores of Task 3 were significantly higher than the ones of Task 2. Once again, Group 2 followed the same pattern.

## 6. Discussion

The groups of higher and lower proficiency levels followed the same pattern in vocabulary acquisition in both the receptive and produc-

tive word recall. The results show that in all cases Task 3 with the highest involvement load brought about the best results in lexical acquisition and retention. The delayed post-tests revealed that Task 2 was more effective than Task 1, which is totally in line with the predictions of the hypothesis, whereas the immediate post-tests indicated that Task 1 was superior to Task 2, which is not in accordance with it. Some explanations for such findings are discussed in the sections below.

### *6.1. Task effectiveness and time-on-task*

The results of the delayed post-test concur with the results of the Hulstijn and Laufer's (2001b) Hebrew-English experiment in the sense that Task 3 was superior to Task 2 and that Task 2 was superior to Task 1. Surprisingly, the results of the immediate post-tests are closer to the Dutch-English experiment in that study, which does not fully support the Involvement Load Hypothesis, in that Task 2 did not produce significantly higher retention than Task 1. What is more, the results of the immediate post-test in this study show that Task 1 (reading comprehension with marginal glosses) was superior to Task 2 (reading comprehension plus fill in with marginal glosses) and the posthoc Tukey test revealed significance in this regard (e.g. active word recall post-test for Group 1  $HSD = 0.750681 < M1 - M2 = 0.98675$ ), even though the involvement load was higher in Task 2.

The cases of partial and not full correspondence of the results of this research with the Involvement Load Hypothesis may have a possible explanation in the effect of the time allotted for task completion. The allotted time for the first part of Task 1 (prior to delivering the comprehension questions) was 15 minutes, which could have been considerably longer than the time needed to complete the task. The students had enough time to revert to the text and the glosses at least twice if not several times. In that way, these learners could have received additional exposure to the target words. Better acquisition and short-term retention recorded in the immediate post-test may have been the result of repetition, which was absent in other exercises.

Nation (2001) compared spaced repetition to massed repetition and concluded that the results of spaced repetition are remembered for a longer time. Massed repetition involves spending a continuous period of time paying repeated attention to a word. According to Nation (2001) it only leads to short-term retention of vocabulary. Baddely, in Nation (2001, p.

77) speculated that “massed learning does not allow enough time for the regeneration of neuro-chemical substances which make physical changes on the brain and thus cannot continue to make physical changes needed for learning”. This may be one of the reasons why students who completed Task 1 in this research outperformed the learners who were engaged in Task 2 in the immediate post-test, but not in the delayed post-test.

The results inconsistent with the hypothesis, in this case, can be explained for by the inadequate time allowed for task completion. Although the tests were announced, the students were also informed that the tests would not influence their course grade. Their motivation to devote extra time to massed repetition of words during extra time in Task 1 may have been moderate. This study does not focus on the time-on-task component, but suggests time-on task and detailed study of words learned per minute as variables which should be taken into account when estimating task effectiveness in future research.

Conversely, the delayed post-tests showed that Task 2 scores were significantly higher than Task 1 scores. This finding is totally in line with the predictions of the Involvement Load Hypothesis. One more reason for differences in the delayed post-test results may be the expiry of the effect of massed repetition. Short-term retention, which appeared immediately, is likely to fade very soon in the absence of subsequent rehearsal.

On the other hand, Task 3 was the most effective of the three tasks in both post-tests. Some evidence suggest that benefits related to more effective tasks disappear when completion times are kept constant for all tasks in the research design (Webb, 2005). Completion times for this research were held constant, but time-on-task was well-designed for Task 3. The students had enough time to complete this exercise.

In the second experiment of a study by Webb (2007), where the allotted time-on-task depends on the time-on-task needed for completion of the writing task, that task was more effective than the reading task. If the second experiment is an example of authentic learning, a stronger argument can be made for the use of productive over receptive tasks. Similarly, in this research the time on Tasks 1, 2 and 3 varied and the tasks were designed to provide the learners with as much time as needed for the completion of a respective task based on the teacher's estimate. This turned out to be enough time for Task 3 and more than enough for Task 1. Task 3 being superior to Tasks 1 and 2 in all tests supports the findings of Webb (2005)

that when given enough time-on-task productive tasks yield better results and tend to be more effective than the receptive ones.

### *6.2. Task types and input or output orientation*

Prior to their experiments Hulstijn and Laufer (2001a) analyzed a considerable number of studies which investigated task effectiveness effect on vocabulary acquisition and retention. The hypothesis predicts that a task which involves the reader in deeper elaboration, i.e. greater involvement load, yields better results in lexical acquisition. It does not speculate on whether the tasks are input or output oriented. Hulstijn and Laufer's (2001 b) experiment and the current study gave results which lead to the presumption that output oriented tasks instigate better lexical acquisition. Conversely, Keating's (2008) study showed that learners involved in an input oriented task outperformed the ones involved in an output oriented task in the delayed post-test.

Interestingly, Laufer (2005) concluded that active knowledge is difficult to acquire since there is so much to learn about a word in order to use it properly (pronunciation, spelling, collocations, register etc.). According to Laufer (2005) if the word has been acquired productively, it will be less likely to be forgotten. This study supports that premise as the productive task (Task 3) appeared to be the most effective in all post-tests.

On the other hand, a study on task effectiveness by Keating (2008) found that Task 3 (sentence completion) was not more effective than Task 2 (reading plus fill-in). Keating (2008) suggested an explanation for such results in the type of writing exercise. While it was composition writing in Hulstijn and Laufer's (2001b) study, Task 3 was a sentence writing task in his study.

The current study chose composition writing for Task 3 like the Hulstijn and Laufer's (2001b) study, and its results indicate that the productive task was superior. The reason for such different results may be found in the assumption that producing connected discourse involves the student in more elaborate processing of words than producing disconnected sentences. Although that could be the explanation for such different results finding evidence to support this premise is one of the tasks left to future research.

Furthermore, the findings of this study are in favour of productive tasks, but proving definite superiority of productive over receptive tasks requires more research. The current research confined itself only to the

comparison of tasks of different involvement load, regardless of their productive or receptive nature. One of the suggestions left to future research is to compare the effectiveness of tasks with the same involvement index but different input/output orientation.

The findings of this study show that all students' word knowledge deteriorated in the period between the immediate and delayed post-tests. The task which recorded the highest retention of words in the immediate post-test also suffered the greatest decline in vocabulary growth after two weeks (e.g. productive word recall test for Group 1 immediate  $M3=2.33 >$  delayed  $M3= 0.68$ ; for Group 2 immediate  $M3=2.89 >$  delayed  $M3= 0.87$ ), which is similar to the Keating's (2008) experiment. This may be the result of the absence of repetition and rehearsal, or the opportunity to use these words in context. The thing which is more important for long term retention of the newly acquired vocabulary than the conditions under which new words have been come across for the first time are recurrent exposures to the words, i.e. practice (Hulstijn, 2003). Repetition is essential for vocabulary learning, because there is so much information to learn about each word that seeing it only once is not sufficient for retention (Nation, 2001). Repetition and particularly spaced repetition ensures acquisition which can be remembered for a long time.

### *6.3. Incidental and Intentional Learning*

Although the Involvement Load Hypothesis, which was tested in the current study, was constructed for incidental vocabulary acquisition, some conditions in the course of this resewrch, such as test announcement, were more typical of intentional learning.

However, Nation (2001) argues that intentional and incidental learning are complementary activities and that each can improve the learning that comes from the other. The distinction between message-focused, incidental learning and language-focused intentional learning is not distinct, and as Nation (2001, p. 232) noted "all learning involves conscious attention".

Learning vocabulary from context like in the tasks of this research is supposed to be incidental, because the tasks of this research are message-focused items rather than direct study of language items. Learning from context does not involve deliberate learning of words and their definitions. However, Keating (2008, p.16) noted that "test announcement is the main distinctive feature of intentional learning". In this research, the immedi-

ate and delayed post-tests were announced, but the learners were also informed that they were participating in a research and that the test results would not have impact on their course grade. That is why they were not consciously making an effort to acquire the words, and it is highly improbable that they were engaged in any kind of dictionary look-up of target words in the period between the immediate and the delayed post-test, since they were not motivated enough to do so. In addition to that, the results report on the relatively low acquisition of vocabulary which is typical of incidental vocabulary acquisition.

#### *6.4. Receptive and Productive Target Word Recall*

The results of this study also show that higher involvement load leads to gains in both receptive and productive knowledge of words. This is true for most students of the mixed-ability class in this study. Nation (2001) argues that receptive vocabulary use can be noticing a word form and recalling its meaning while reading, whereas productive vocabulary use implies expressing the meaning through speaking or writing and providing the suitable word form.

Nation (2001) considers Corson's idea of the *lexical bar*<sup>4</sup> (barrier) important for making receptive/productive distinction. "The lexical bar is a gap between everyday meaning systems and the high status meaning systems created by the introduction of an academic culture of literacy. This is a barrier that has to be crossed by everyone at some stage in their lives if they want to be successful in conventional forms of education." (Corson, as cited in Nation, 2001, p.26). Briefly, lexical bar is the consequence of being unfamiliar with academic meaning systems, and that is the reason why for some learners certain vocabulary items remain receptive.

This research was carried out in an academic ESP class, in which both the lower and the higher general language ability students were not inhibited by the lexical bar, since they are educated enough in conventional forms of education. This can serve as the explanation for no considerable difference in the scores of the receptive and productive posttests for most students in the mixed-ability ESP class.

---

4 Corson in Nation (2001) presents the theory of lexical bar—a hypothesized semantic barrier in the English (of England and Australia) lexicon which is the reason why some vocabulary items remain receptive for learners that are unfamiliar with academic meaning systems.

### *6.5. Pedagogical Implications*

As McCarthy (1990, p. 9) noted “advances in language teaching stem from the independent efforts of teachers in their own classrooms“. Teachers of ESP academic classes sometimes have to focus on a challenging task of teaching subject-specific vocabulary to mixed-ability classes, which is why they need to design teaching materials for that purpose. Bejan (1989) argues that subject specific vocabulary used in the register of a particular field is one of the main focuses of a good ESP course. Subject-matter reading, as part of ESP courses, contains subject-specific vocabulary, which can be followed by effective tasks aimed at vocabulary acquisition. This research, which investigates task effectiveness of three different task types, reached the conclusion that tasks with a higher involvement load lead to better vocabulary acquisition and longer retention of that vocabulary. This conclusion can contribute to the creation of tasks with high involvement load as effective materials to introduce reading or create effective tasks and vocabulary exercises to accompany subject-matter reading. In addition, the findings of this research refer to a mixed-ability class as a whole, since the two groups of lower and higher general language proficiency followed the same pattern in vocabulary acquisition.

In the first place, the results of this study indicate that an output oriented productive task with a strong evaluation component (Task 3) is superior to tasks which are input oriented receptive tasks (reading comprehension exercises, i.e. reading and reading plus fill-in). These findings are in line with the studies which suggest that productive, i.e. output oriented tasks tend to be more effective than the receptive ones, i.e. input oriented tasks, although more studies need to be conducted to prove that.

Secondly, the thing which is definitely true for Task 3 is that it had the highest involvement index. The findings of this study suggest tasks with high involvement load as the most effective ones, and this can help teachers prioritize in the choice of tasks used after or instead of reading for the necessary practice and rehearsal of the target vocabulary.

Thirdly, the delayed vocabulary post-test fully supported the Involvement Load Hypothesis. Students had enough time to complete Task 3 (writing), while for Task 1 the students had some extra time for massed repetition. Consequently, the immediate vocabulary post-test revealed only short-term better retention of the newly-acquired vocabulary for stu-

dents who performed a task with a lower involvement index. Both findings suggest that time-on-task should be adequately allotted.

Finally, vocabulary knowledge of all students deteriorated over time, which is why the current study puts emphasis on the necessity of repetition and rehearsal after lexical acquisition in order to prevent fading of the knowledge.

## **7. Conclusions**

The research results of this study have implications which are not limited to ESP courses. They are generalizable to different types of foreign language courses, since most of them deal with vocabulary acquisition to a certain extent and need effective task types for vocabulary acquisition. They are also not limited to mixed-ability classes, as both groups of lower and higher English language proficiency followed the same pattern in vocabulary acquisition when analyzed independently. Therefore, it can be concluded that findings of this research may help language researchers and teachers outside the specific context of this research in finding effective task types and vocabulary exercises which can accompany reading. These are productive tasks, such as writing of a paragraph with the words given, and non-productive tasks, such as reading comprehension with a gap-fill rather than reading comprehension with the words given and explained in the margins. The first two tasks, and primarily the productive task, are recommended due to their involvement load.

During the research, it was shown that exercises with higher involvement load lead to more successful vocabulary acquisition and longer retention for most students in a mixed-ability ESP class, when compared to exercises with a lower involvement index. The study also implies that these exercises contribute to both students' receptive and productive word recall.

The aim of this study was to be one more trial of the Involvement Load Hypothesis, in this case in a mixed-ability ESP class. The higher level of processing of information turned out to be more effective in vocabulary acquisition, which is in line with the predictions of the hypothesis, while partial discrepancies between the hypothesis and the results draw our attention to the issue of time-on-task.

It was concluded that the acquired vocabulary tends to fade rather quickly in the absence of repetition. That is why, it is suggested to provide



long term retention of the acquired vocabulary with appropriate rehearsal to prevent deterioration of the knowledge.

In the course of this research a lot of interesting issues arose that could be a useful follow up to this study. Due to the format and time limits, it was not possible to include these potential extensions in this study. These possible foundations for future research can be summed up as the following: investigating the effect of time-on-task (i.e. limited, sufficient or extra time-on-task) on task effectiveness and taking into account words learned per minute when estimating task effectiveness, as well as investigating the impact of task effectiveness in vocabulary acquisition on different types of word knowledge. The study also suggests comparison of two exercises with the same involvement load, but different output or input orientation, and comparison of different types of writing exercises, i.e. composition writing and writing of separate sentences on vocabulary acquisition.

## REFERENCES

- Bejan, N. (1989). Teaching Vocabulary in English for Special Purposes. In: Peterson, Pat Wilcox, *ESP in Practice* (pp. 93-99) Washington, D. C.: English Language Programs Division. Bureau of Educational and Cultural Affairs. United States Information Agency.
- Cho, K. & Krashen, S.D. (1994). Acquisition of Vocabulary from the Sweet Valley Kids Series: Adult ESL Acquisition. *Journal of Reading*, 37 (8), 662-667. Retrieved from [http://www.sdkrashen.com/content/articles/1994\\_sweet\\_valley\\_high\\_with\\_kscho.pdf](http://www.sdkrashen.com/content/articles/1994_sweet_valley_high_with_kscho.pdf)
- Corson, D. (1983). Social dialect, the semantic barrier, and access to curricular knowledge. *Language in Society*, 12, 213-222. <https://doi.org/10.1017/S0047404500009817>
- 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. [https://doi.org/10.1016/S0022-5371\(72\)80001-X](https://doi.org/10.1016/S0022-5371(72)80001-X)
- Craik, F.I.M. & Tulving, E. (1975). Depth of Processing and the Retention of Words in Episodic Memory. *Journal of Experimental Psychology: General*, 104 (3), 268-294. <https://doi.org/10.1037/0096-3445.104.3.268>

- Douglas, V. (2016). Enhancing English academic vocabulary acquisition and retention in intensive English programs with the Involvement Load Hypothesis. *Theory and practice in language studies*, 6 (12), 2237-2244. <http://dx.doi.org/10.17507/tpls.0612.01>
- de la Fuente, M. J.(2006). Classroom L2 vocabulary acquisition: Investigating the role of pedagogical tasks and form-focused instruction. *Language Teaching Research*, 3 (10), 263–295. <https://doi.org/10.1191/1362168806lr196oa>
- Haratmeh, M.S. (2012). Involvement load and task type in task effectiveness: Two aspects of vocabulary knowledge. *International Journal of Academic Research*, 4 (4), 86-116. <http://dx.doi.org/10.7813/2075-4124.2012/4-4/B.13>
- Hu, M. & Nation, P. (2000). Unknown Vocabulary Density and Reading Comprehension. *Reading in a Foreign Language*, 13 (1), 403- 430. Retrieved from <https://www.victoria.ac.nz/lals/about/staff/publications/paul-nation/2000-Hu-Density-and-comprehension.pdf>
- Huang, S., Eslami, Z. & Willson, V. (2012). The Effects of Task Involvement Load on L2 Incidental Vocabulary Learning: A Meta-Analytic Study. *The Modern Language Journal*, 96 (4), 544-557.
- Hulstijn, J.H. (2003). Incidental and intentional learning. In: C. Doughty & M.H. Long (Eds.), *The handbook of second language research* (pp.349-381). London: Blackwell. Retrieved from [https://books.google.rs/books?hl=en&lr=&id=xmLoVScagwYC&oi=fnd&pg=PA349&ots=jlCDUgJYn6&sig=01il1TG2aQlUO2vJ4kG7uEL26J8&redir\\_esc=y#v=onepage&q&f=false](https://books.google.rs/books?hl=en&lr=&id=xmLoVScagwYC&oi=fnd&pg=PA349&ots=jlCDUgJYn6&sig=01il1TG2aQlUO2vJ4kG7uEL26J8&redir_esc=y#v=onepage&q&f=false)
- Hulstijn, J. (2002). Towards a Unified Account of the Representation, Processing and Acquisition of Second Language Knowledge. *Second Language Research*, 18 (3), 193-223. <https://doi.org/10.1191/0267658302sr207oa>
- Hulstijn, J. (1993). When Do Foreign-Language Readers Look up the Meaning of Unfamiliar Words? The Influence of Task and Learner Variables. *The Modern Language Journal*, 77 (2), 139-147. <https://doi.org/10.1111/j.1540-4781.1993.tb01957.x>
- Hulstijn, J. & Laufer, B. (2001a). Incidental Vocabulary Acquisition in a Second Language: The Construct of Task-Induced Involvement. *Applied Linguistics*, 1 (22), 1-26. <https://doi.org/10.1093/applin/22.1.1>
- Hulstijn, J. & Laufer, B. (2001b). Some Empirical Evidence for the Involvement Load hypothesis in Vocabulary Acquisition. *Language Learning*, 51 (3), 539-558. <https://doi.org/10.1111/0023-8333.00164>
- 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. *The Modern Language Journal*, 3 (80), 327-339. <https://doi.org/10.1111/j.1540-4781.1996.tb01614.x>
- Jing, L. & Jianbin, H. (2009). An Empirical Study of the Involvement Load Hypothesis in Incidental Vocabulary Acquisition in EFL Listening. *Polyglossia*, 16, 1-11. Retrieved from [http://www.apu.ac.jp/rcaps/uploads/fckeditor/publications/polyglossia/Polyglossia\\_V16\\_Lu\\_Huang.pdf](http://www.apu.ac.jp/rcaps/uploads/fckeditor/publications/polyglossia/Polyglossia_V16_Lu_Huang.pdf).
- Joe, A. (1995). Text-based Tasks and Incidental Vocabulary Learning. *Second Language Research*, 11 (2), 149-158.
- 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. <https://doi.org/10.1177/026765839501100206>
- Krashen, S. (1987). *Principles and Practice in Second Language Acquisition*. Englewood Cliffs, N. J.: Prentice-Hall International.  
Retrieved from [http://www.sdkrashen.com/content/books/principles\\_and\\_practice.pdf](http://www.sdkrashen.com/content/books/principles_and_practice.pdf)
- Krashen, S. (1989). We Acquire Vocabulary and Spelling by Reading: Additional Evidence for the Input Hypothesis. *The Modern Language Journal*, 73 (4), 440-464. <https://doi.org/10.1111/j.1540-4781.1989.tb05325.x>
- Laufer, B. & Paribakht, S.T. (1998). The Relationship Between Passive and Active Vocabularies: Effects of Language Learning Context. *Language Learning*, 48 (3), 365-391. Retrieved from [http://www.academia.edu/9833737/The\\_Relationship\\_Between\\_Passive\\_and\\_Active\\_Vocabularies\\_Effects\\_of\\_LanguageLearning\\_Context](http://www.academia.edu/9833737/The_Relationship_Between_Passive_and_Active_Vocabularies_Effects_of_LanguageLearning_Context)
- Laufer, B. (2006). Comparing Focus on Form and Focus on FormS in Second language Vocabulary Learning. *The Canadian Modern Language Review*, 63 (1), 149- 166. <https://doi.org/10.3138/cmlr.63.1.149>
- Laufer, B. (2005). Focus on form in second language vocabulary learning. *EUROSLA Yearbook*, 5, 223–250. <https://doi.org/10.1075/eurosla.5.11lau>
- Laufer, B. (2003). Vocabulary Acquisition in a Second Language: Do Learners Really Acquire Most Vocabulary by Reading: Some Empirical Evidence. *The Canadian Modern Language Review*, 59 (4), 567-587. <https://doi.org/10.3138/cmlr.59.4.567>
- Luppescu, S. & Day, R. (1993). Reading, Dictionaries, and Vocabulary Learning. *Language Learning*, 43 (2), 263-278. <https://doi.org/10.1111/j.1467-1770.1992.tb00717.x>
- McCarthy, M. (1990). *Vocabulary*. Oxford: Oxford University Press.
- Meara, P. (2002). Vocabulary. In Schmitt, N., *An Introduction to Applied Linguistics* (pp. 35- 55). London: Arnold. A member of the Hodder Headline Group.

- Nation, P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press. Retrieved from <http://catdir.loc.gov/catdir/samples/cam031/2001269892.pdf>
- Newton, J. (1995). Task-based Interaction and Incidental Vocabulary learning: A Case Study. *Second Language Research*, 11 (2), 159-177. <https://doi.org/10.1177/026765839501100207>
- Paribakht, S. & Wesche, M. (2000). Reading-Based Exercises in Second Language Vocabulary Learning: An Introspective Study. *The Modern Language Journal*, 84 (2), 196-213. <https://doi.org/10.1111/0026-7902.00062>
- Soleimani, H., Rahmanian, M. & Sajedi, K. (2014). A Revisit to Vocabulary Acquisition in Involvement Load Hypothesis. In *Procedia – Social and behavioral sciences* (2015), 192 (pp. 388-397). Second Global Conference on Linguistics and Foreign Language Teaching LINELT-2014, United Arab Emirates, Dubai. Retrieved from [http://www.academia.edu/14169986/Soleimani\\_H.\\_Rahmanian\\_M.\\_and\\_Sajedi\\_Kh.\\_2015\\_.A\\_revisit\\_to\\_vocabulary\\_acquisition\\_in\\_involvement\\_load\\_hypothesis\\_Procedia\\_Social\\_and\\_Behavioral\\_Sciences\\_192\\_388-397](http://www.academia.edu/14169986/Soleimani_H._Rahmanian_M._and_Sajedi_Kh._2015_.A_revisit_to_vocabulary_acquisition_in_involvement_load_hypothesis_Procedia_Social_and_Behavioral_Sciences_192_388-397)
- Tonić, A. (1989). *Engleski jezik*. Beograd: Šumarski fakultet, Univerzitet u Beogradu, Kneza Višeslava br. 1, Beograd.
- Webb, S. (2007). Learning Word pairs and Glossed Sentences: The Effect of a Single Context on Vocabulary Knowledge. *Language Teaching Research*, 11 (1), 63-81. <https://doi.org/10.1177/1362168806072463>
- 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. <https://doi.org/10.1017/S0272263105050023>

**Katarina O. Lazić**

**TEST HIPOTEZE O OPTEREĆENJU UČEŠĆEM  
U GRUPI MEŠOVITOG NIVOVA ZNANJA NA KURSU ENGLESKOG JEZIKA  
ZA POSEBNE NAMENE**

**Sažetak**

Ovo istraživanje testira postulat Hipoteze o opterećenju učešćem koju su 2001. godine postavili Hulstijn i Laufer, koja pretpostavlja da usvajanje i zadržavanje novousvojenog vokabulara zavise od opterećenja učešćem u određenom zadatku. Hipoteza je testirana u razredu mešovitog predznanja, koji pohađa kurs engleskog jezika za specifične potrebe, sa ciljem da ispita njena predviđanja za većinu studenata u takvom razredu i njihovo receptivno i produktivno znanje usvojenih reči. Tri zadatka sa različitim ideksima učešća (čitanje i razumevanje teksta sa ciljnim rečima objašnjenim na margini, čitanje i razumevanje sa popunjavanjem praznina u tekstu i pisanje pasusa sa datim rečima) zadata su grupi od pedeset i dva studenta na Šumarskom fakultetu u Beogradu. Njihovo receptivno i produktivno znanje ciljnog vokabulara je testirano odmah i nakon dve nedelje. Rezultati testova potvrđuju predviđanja hipoteze. Najbolji su kod studenata koji su radili zadatak pisanja, lošiji kod onih koji su radili čitanje sa razumevanjem i popunjavanje praznina i najlošiji kod onih koji su radili čitanje sa razumevanjem i objašnjenim rečima. Mada je većina rezultata u saglasnosti sa predviđanjima hipoteze u radu se razmatraju i odstupanja za koja su data neka moguća objašnjenja.

## **Appendix A:** Vocabulary acquisition tasks used in this study

### **Task 1**

Read the text and the explanations below. You have 15 minutes.

After that answer True or False.

#### **SOIL MAPS AND MAPPING**

One of the major occupations of soil scientists is surveying and **mapping** soils and the production of **soil maps**. To **survey** soils is a most **exacting** scientific task because soils do not have sharp boundaries but gradually grade from one into another.

There are various methods for conducting soil surveys. Most are “free surveys” based on the experience of the soil surveyor who **criss-crosses** the area making inspections at points in the landscape where he considers it necessary to determine the nature of the soils and to establish boundaries.

The type of map on which the data are recorded can vary very much from place to place. In countries such as Britain, ordnance survey maps are used, but for many places such maps have not been produced. Then it is **customary** to use aerial photographs and to record data either directly on to the photograph or on to a transparent **overlay**. Aerial photographs have a number of advantages, particularly with regard to finding one’s location, but aerial photographs have a varying **scale** across the photograph, therefore they cannot be used as complete substitutes for maps.

The surveyor draws lines on a map or photograph to enclose areas of relative uniformity and so to produce a soil map. The field map is then submitted to the drawing office for the preparation of the final coloured map and its accompanying legend.

**Mapping-** kartiranje

**Soil map-** pedološka karta

**Survey-** snimati teren

**Exacting-** precizan

**Criss-cross-** ići cik-cak po terenu

**Customary-** uobičajen

**Overlay-** folija

**Scale-** razmera

## Task 2

Read the text and the explanations below and fill in the blanks. You have 20 minutes.

### SOIL MAPS AND MAPPING

One of the major occupations of soil scientists is surveying and 1. \_\_\_\_\_ soils and the production of 2 \_\_\_\_\_. To 3. \_\_\_\_\_ soils is a most 4. \_\_\_\_\_ scientific task because soils do not have sharp boundaries but gradually 5. \_\_\_\_\_ from one into another.

There are various methods for conducting soil surveys. Most are “free surveys” based on the experience of the soil surveyor who 6. \_\_\_\_\_ the area making inspections at points in the landscape where he considers it necessary to determine the nature of the soils and to establish boundaries.

The type of map on which the data are recorded can vary very much from place to place. In countries such as Britain, ordnance survey maps are used, but for many places such maps have not been produced. Then it is 7. \_\_\_\_\_ to use aerial photographs and to record data either directly on to the photograph or on to a transparent 8. \_\_\_\_\_. 9. \_\_\_\_\_ photographs have a number of advantages, particularly with regard to finding one’s location, but aerial photographs have a varying 10. \_\_\_\_\_ across the photograph, therefore they cannot be used as complete substitutes for maps.

The surveyor draws lines on a map or photograph to enclose areas of relative 11. \_\_\_\_\_ and so to produce a soil map. The field map is then submitted to the drawing office for the preparation of the final coloured map and its 12. \_\_\_\_\_ legend.

**Criss-cross-** To move back and forth through or over: *ići cik-cak po terenu*

**Aerial-** Of, in, or produced by the air. (Sj.) *vazdušni*, *iz vazduha*

**Soil map-** Soil map is a map showing distribution of soil types and/or soil properties. (Sj.) *pedološka karta*

**Survey-** To survey soils, is to engage in the process of classifying soil types and other soil properties in a given area (Sj.) *snimati teren*

**Exactig-** strictly accurate or correct: *an exact likeness; an exact description.* (Sj.) *precizan*

**Customary-** According to or depending on custom; usual; habitual.- (Sj.) *uobičajen*

**Uniformity-** The state or quality of being uniform; overall sameness, homogeneity, or regularity. (Sj.) *Jednobraznost*

**Scale-** the ratio of distances or sometimes of areas on a map to the corresponding values on the earth. (Sj.) *razmera*

**Mapping-** To explore or make a survey of (a region) for the purpose of making a map. (Sj.) *kartiranje*

**Accompanying-** To go along or in company with; join in action: *to accompany a friend on a walk.* (Sj.) *Prateći*

**Overlay-** A sheet of transparent paper placed over a photograph – (Sj.) *folija*

**Grade-** To cause to pass by degrees, as from one color or shade to another. (Sj.) *Postepeno prelaziti.*

### Task 3

Write a paragraph of about 100 words on the production of soil maps using these words. You must use the words given! You have 25 minutes.

**Criss-cross-** To move back and forth through or over: *ići cik-cak po terenu*

**Aerial-** Of, in, or produced by the air. (Sj.) *vazdušni, iz vazduha*

**Soil map-** Soil map is a map showing distribution of soil types and/or soil properties. (Sj.) *pedološka karta*

**Survey-** To survey soils, is to engage in the process of classifying soil types and other soil properties in a given area (Sj.) *snimati teren*

**Exacting-** strictly accurate or correct: *an exact likeness; an exact description.* (Sj.) *precizan*

**Customary-** According to or depending on custom; usual; habitual.- (Sj.) *uobičajen*

**Uniformity-** The state or quality of being uniform; overall sameness, homogeneity, or regularity. (Sj.) *Jednobraznost*

**Scale-** the ratio of distances or sometimes of areas on a map to the corresponding values on the earth. (Sj.) *razmera*

**Mapping-** To explore or make a survey of (a region) for the purpose of making a map. (Sj.) *kartiranje*

**Accompanying-** To go along or in company with; join in action: *to accompany a friend on a walk.* (Sj.) *Prateći*

**Overlay-** A sheet of transparent paper placed over a photograph – (Sj.) *folija*

**Grade-** To cause to pass by degrees, as from one color or shade to another. (Sj.) *Postepeno prelaziti.*



**Appendix B:** The subject matter text on soil maps and mapping as the target vocabulary source

### **Soil Maps and Mapping**

One of the major occupations of soil scientists is surveying and mapping soils and the production of soil maps. Surveying soils is a most exacting scientific task because soils do not have sharp boundaries but gradually grade from one into another. Further, this graduation is seldom clearly expressed on the surface of the ground and has to be determined by making numerous examinations which take the form of small inspection pits or auger holes.

There are various methods for conducting soil surveys. Most are “free surveys” based on the experience of the soil surveyor who criss-crosses the area making inspections at points in the landscape where he considers it necessary to determine the nature of the soils and to establish boundaries. At each point of inspection the properties of the soil are recorded on a map and when necessary in a notebook. Alternatively, he carries out inspections at predetermined locations which are based on some type of statistical programme. This technique is more objective, but is often time consuming.

The type of map on which the data are recorded can vary very much from place to place. In countries such as Britain, ordnance survey maps are used, but for many places such maps have not been produced. Then it is customary to use aerial photographs and to record the data either directly onto the photograph or on to a transparent overlay. Aerial photographs have a number of advantages, particularly with regard to finding one’s location, but aerial photographs have a varying scale across the photograph, therefore they cannot be used as complete substitutes for maps. The scale of the map or photograph varies depending upon the object of the investigation and time available. The scale seldom exceeds 10 cm to 1 km except for special purposes such as irrigation. From the data recorded on the map or photograph and in his notebook, the surveyor draws lines on a map or photograph to enclose areas of relative uniformity and so to produce a soil map. The field map is then submitted to the drawing office for the preparation of the final coloured map and its accompanying legend.

The unit on soil maps and mapping taken from “Engleski jezik” Tonić (1989).