Modern English Education, Vol. 17, No. 4, Winter 2016
http://dx.doi.org/10.18095/meeso.2016.17.4.01

Task-Induced Involvement Load and Low-Level Learners' Vocabulary Learning

HyunMi Sung
Kunsan National University

Sung, HyunMi. (2016). Task-induced involvement load and low-level learners' vocabulary learning. *Modern English Education*, *17*(4), 1-24.

This study purposed to testify the Involvement Load Hypothesis (Hulstijn & Laufer, 2001) especially focusing on the learners of low level of English proficiency. A total of 180 students attending K university were divided into two groups: intermediate-low and low. The participants were randomly assigned to one of the three vocabulary tasks which are different in terms of involvement load index established by Hulstijn and Laufer (2001): reading comprehension (1), reading comprehension plus gap-filling (2), and unscrambling sentences (3). Prior to the task performance, a pretest intended to check the learners' prior knowledge of the target words was executed. To measure the learners' short-term learning and long-term retention, two posttests were administered: one time immediately after and the other time two weeks after task performance. The results showed that task-induced involvement load affected the learners' short-term learning but did not influence on long-term retention. Likewise, proficiency level was proved significant in short-term learning but was not in long-term retention. Besides, no interaction effect between the two factors, task-induced involvement load and proficiency level, was found in this study. The results will be analyzed more in-depth and implications will be discussed based on the findings.

**[task-induced involvement load/short-term learning/long-term retention/
과업으로 인한 관여도/단기학습/장기기억]**

I. INTRODUCTION

It has been widely acknowledged that vocabulary knowledge is a critical competence in learning a second or foreign language since it plays a basic and fundamental role in

communication (Akbarian, 2010; H. Hahn, 2014). Over the past decades, in the field of English education, ESL or EFL vocabulary learning has been researched from various points of view (Folse, 2004; Nation, 2001; Schmitt, 2000).

In general, EFL learners need to strive for vocabulary learning because they are hardly exposed to L2 learning environment except classroom. Furthermore, the learners' outcomes might be different by how deeply they process the words when they encounter new vocabulary (Craik & Lockhart, 1972). This can be said that most learners are not convinced about whether they learn vocabulary in an efficient way or not (H. M. Sung, 2013). In particular, unlike high level learners, low level learners would have more difficulties in learning L2 vocabulary. This has been proven by some research findings on the relationship between the use of vocabulary learning strategies and the general English proficiency; high proficiency learners are more skillful at using vocabulary learning strategies than low proficiency learners (Gu & Johhson, 1996; Kojic-Sabo & Lightbown, 1999; S. Lee & M. Min, 2006).

According to Hulstijn and Laufer (2001), language teachers can support their learners by providing vocabulary tasks inducing different level of involvement load. The efficiency of such vocabulary tasks can be measured by whether the students recall target words after task performance and also by how long they retain those new words. Craik and Lockhart (1972) focused on the level of processing by which the learners can remember new information over time: deeper level of processing leads to long-term memory. Based on this concept, Hulstijn and Laufer (2001) developed a motivational and cognitive construct (ILH); the learners can have longer vocabulary retention by help of tasks inducing high level of mental efforts, so-called involvement load.

Since the Involvement Load Hypothesis (ILH) was suggested, many researchers have testified the effects of the ILH in classroom settings (Folse, 2006; Y. Kim, 2008; H. Kim & Y. Na, 2010; Laufer, 2003; Webb, 2005). However, the original contention posited by Hulstijn and Laufer that a task of higher level of involvement load yields higher vocabulary retention has not shown consistent results especially from the studies conducted with the participants of different proficiency levels (H. Kim & Y. Na, 2010; J. Kim, 2015; H. M. Sung, 2013). More specifically, H. Kim and Y. Na (2010) found that the task-induced involvement load had a main effect on the learners' vocabulary outcomes across proficiency level whereas H. M. Sung (2013) and J. Kim (2015) showed that low-level learners were not affected by the different types of vocabulary tasks. Then it seems to be essential for researchers to more focus on low-level learners so that teachers can support their students with distinguished methodologies from those for high-level learners. J. Kim (2015) also mentioned that it is necessary to design tasks for the low level learners' better vocabulary retention.

Based on the rationales addressed above, the current study aims to verify whether the

Korean college students of low proficiency levels (intermediate low and low) are affected by the task-induced involvement in terms of short-term learning and long-term vocabulary retention when they are engaged in three different vocabulary tasks. In addition, this study examines whether the task-induced involvement load and proficiency level interact each other in the learners' short-term learning and long-term vocabulary retention. The research questions are as follows:

1) Is the task-induced involvement load influential in low-level learners' vocabulary learning and retention?
2) Is there any interaction effect between task-induced involvement load and proficiency level in low-level learners' vocabulary learning and retention?

II. LITERATURE REVIEW

1. Depth of Processing Hypothesis

The Involvement Load Hypothesis is the construct that Hulstijn and Laufer (2001) developed for vocabulary learning based on the Depth of Processing Hypothesis (Craik & Lockhart, 1972). The basic notion of Depth of Processing Hypothesis (DOPH) is that new information can be a long-term memory only if it is through deeper level of processing. Later, the depth of processing has been designated the level of task-induced involvement load by Hulstijn and Laufer (2001). Craik and Lockhart (1972) insisted that learners can remember words for a longer period of time under the condition the meaning of those words were processed through a greater degree of semantic analysis. In general, semantic or cognitive analysis requires deeper processing, while sensory or physical stimulus can be processed at a shallow level (Selfridge & Neisser, 1960; Sutherland, 1968).

Since the original framework had been introduced, the DOPH was criticized because of its abstractness in defining how the level of processing can be differentiated or measured. Thus, Craik and Tulving (1975) modified the DOPH by implementing the notion of elaboration. 'Elaboration' can be explained as richness by which the information like new words is encoded by learners. Thus, the idea of DOPH that new words processed elaborately can be successful retention to the learners has formulated the basis of the task-induced involvement load (Hulstijn & Laufer, 2001).

2. Task-Induced Involvement Load

The DOHPH has developed into a measurable construct, the Involvement Load

Hypothesis (Hulstijn & Laufer, 2001). The key concept of the Involvement Load Hypothesis (ILH) is 'task-induced involvement load' which can be referred as 'mental efforts' devoted to task performance by the learners. Thus, the ILH claims that learners produce higher retention of new words when they perform a task of a higher involvement load than they engaged in a less involving task. Table 1[1] presents 'task-induced involvement load' according to task types established by Hulstijn and Laufer (2001).

TABLE 1

Task-Induced Involvement Load (Hulstijn & Laufer, 2001)

Task	Status of Target Words	Need	Search	Evaluation
1. reading & comprehension questions	glossed in text but irrelevant to task	-	-	-
2. reading & comprehension questions	glossed in text and relevant to task	+	-	-
3. reading & comprehension questions	not glossed but relevant to task	+	+	-/+
4. reading & filling gaps	relevant to reading at the end of text listed with glosses	+	-	+
5. writing original sentences	listed with glosses	+	-	++
6. writing original sentences	concepts selected by the teacher (and provided in L1) L2 learners-writer must look up the L2 form	+	-	++
7. writing a composition	concepts selected and looked up by L2 learner-writer	++	+	++

Proposing this construct, Hulstijn and Laufer (2001) specified three measurable components: need, search, and evaluation. *Need* is a motivational and non-cognitive component and it gives rise to in two states: moderate (1) and strong (2). More concretely,

[1] The two symbols (+), (-) indicate 'presence' or 'absence' of each component respectively. The sum of the number of each symbol displays the involvement index for each task. For example, task 1 does not require any involvement load because it has only (-) symbols. In the case of task 7, it requires a total of 5 as the involvement index since it has 5 (+) symbols. A single (+) shows a 'moderate' degree and a double (+) signifies a 'strong' state of the concerned component.

The involvement load of evaluation can be 'absence (-)' or 'presence (+)' according to the type of word and the context the word is used in. If a new word has only one meaning, and learners are able to infer the meaning of the word directly from the context, no evaluation (-) is required. However, if the word has more than one meaning, and learners have to think and decide which meaning is suitable in the context, the involvement index for evaluation (+) is needed (i.e., a moderate degree of evaluation). *Source.* Adapted from Hulstijn and Laufer (2001).

when learners need words to finish the given task as per their teachers' request, moderate amount of *need* occurs. On the other hand, if learners need words to finish a certain task as per their own intrinsic motivation, then they have strong amount of *need.* The other two components, *search* and *evaluation* are cognitive dimensions. *Search* is generated when learners try to look up the dictionary for unknown words and it tasks places in two states: present (1) or absent (0). Finally, *evaluation* also comes about in two states: moderate (1) and strong (2). If learners are required to notice the difference between words considering the context, they are supposed to carry moderate amount of *evaluation*, while strong *evaluation* can occur when learners have to select the meaning of unknown words and integrate them into the known words.

3. Incidental Vocabulary Learning

The notion of task-induced involvement load presupposes 'incidental vocabulary learning.' In other words, learners are expected to acquire new words incidentally while they are engaged in a certain task. Over the past decades, incidental vocabulary learning has been researched constantly in the field of SLA (Ellis, 1994; Hustijn, 2003; Nagy & Herman, 1987). According to Hulstijn (2003), not only L1 but also L2 learners acquire a number words incidentally not explicitly. In particular, most of the L2 vocabulary is learned incidentally only except for the first few thousand words (Huckin & Coady, 1999). These two researchers also emphasized the benefits of incidental vocabulary learning in related to task performance: learners can perform two tasks, main task and vocabulary learning task, simultaneously. Based on this background, Hulstijn and Laufer (2001) developed the construct of task-induced involvement load so that learners can get benefits of incidental vocabulary learning while performing a certain task.

4. Previous Studies

1) Studies on the Task-Induced Involvement Load and Learners' Proficiency Level

Y. Kim (2008) also attempted to verify the ILH through two experiments. Her participants were divided into two levels based on their TOEFL scores. As for the tasks of the first experiment, she replicated what Hulstijn and Laufer (2001) did. The results revealed that task 3 group outperformed the other two groups both on the two posttests. In the second experiment, two different writing tasks of the same level of involvement load were compared, and they turned out to have the same amount of involvement load in the learners' short-term and long-term recall. The researcher also examined the role of proficiency level when the participants executed different type of learning tasks, but found

that it did not have main effect on vocabulary learning outcomes.

H. Kim and Y. Na (2010) examined the effects of task type and proficiency level in terms of task-induced involvement load. Their participants were two proficiency groups based on TOEIC scores, and the same tasks that Hultijn and Laufer (2001) and Y. Kim (2008) employed in their studies were used. The research findings generally supported the IHL in the learners' immediate recall but not in long-term retention. Unlike what Y. Kim (2008) found, these two researchers discovered the factor of proficiency level played an influential role both in the learners' vocabulary retention as well as immediate learning.

H. M. Sung (2013) also lends support to experimental investigations on the ILH. The participants were 217 high school students of two different proficiency levels: high and low. For the first two tasks inducing involvement load 1 and 2, the type of tasks (reading comprehension and reading comprehension plus gap-filling) used in the previous studies was replicated, but unscrambling sentences was employed instead of writing original sentences. The research results showed that involvement load only influenced on high-level learners' learning. In addition, involvement load and proficiency were proved as significant determiners in the learners' vocabulary learning.

Most recently, J. Kim (2015) conducted a similar experiment to investigate the roles of task-induced involvement load and proficiency level in vocabulary retention. As done in other previous studies, to measure the learners' vocabulary outcomes, a pair of posttest (receptive and productive) was administered twice: immediately and two weeks after the task performance. In her research, low proficiency learners demonstrated a significant difference only in the immediate receptive test among the four posttests, whereas high proficiency learners made significant differences in all posttests. In her study, task type and proficiency level turned out to have mostly have interaction effects each other.

2) Studies Only Focused on the Effect of Task-Induced Involvement Load

Hulstijn and Laufer (2001) conducted two experiments with groups of EFL college learners whose mother tongues were Dutch and Hebrew, respectively. In their studies, three vocabulary learning tasks of different level of involvement load were used: reading with glosses (1), gap-filling (2), and composition (3). However, the results of the two studies were not consistent. That is, the Hebrew group supported the ILH, whereas the Dutch group only partially did.

Laufer (2003) testified the effect of the task involvement load through a series of experiments. In the first and second experiments in which reading comprehension (1) and two different writing tasks (3) were compared, vocabulary outcomes of 90 EFL high school learners with Arabic background confirmed the Involvement Load Hypothesis. On the other hand, from the third experiment compared three different tasks inducing the same

level of involvement load, the results did not support the Involvement Load Hypothesis entirely.

Keating (2008) carried out a study with the beginner level of college students learning Spanish by employing three tasks of different level of involvement load: reading (1) and gap-filling (2) and writing original sentences (3). He found that as for the learners' passive recall on the target words (writing translation of Spanish words in L1), the results confirmed the ILH although the active recall (producing Spanish words for L1 correspondance) in the delayed test did not.

The studies presented in this section have attempted to verify the Involvement Load Hypothesis (Hulstijn & Laufer, 2001) through vocabulary learning tasks inducing different level of involvement load. The results of the studies, however, so far have not achieved an entire consensus on the Involvement Load Hypothesis, particularly on its effect on the learners' vocabulary retention. Moreover, the research investigated on the role of learners' different proficiency level has reported that low-level learners' vocabulary learning is not affected by the different type of tasks. In the classroom, learners who are in need of instructional support in vocabulary learning is not high level but low level. In this light, the current study would provide meaningful answers so that teachers can get practical benefits in guiding low level learners.

III. METHOD

1. Participants

The population for the current study was 180 students who were taking the basic English course (TOEIC) at K university located in a local province in Korea. All the participants were freshmen whose majors varied. Although the researcher taught 217 students (6 classes), 37 students were excluded from the participants due to their absence during the treatment period. The participants of this study had belonged to their own level based on the placement test administered by the school. The researcher, however, reclassified them into intermediate-low and low levels on the basis of the mock-TOEIC scores that the participants. More concretely, in this study the students who obtained over and under 400 from the TOEIC test were assigned into two different levels: 87 students into intermediate-low and 93 students into low level. The two groups' average TOEIC scores were 524 and 290, respectively.

2. Procedures and Tasks

The current study was conducted from the third week of May to the second week of June in 2016. Over the three weeks, the experiment on the task-induced involvement load was carried out. On the first day, after the pre-test (prior knowledge test) about the ten target words, three different tasks were introduced to the participants, and one of the tasks was distributed randomly to each student. Immediately after the task performance, a post-test was executed to check the learners' short-term learning. Three weeks later, a delayed post-test was conducted to measure the learners' retention on the target words. In order to prevent the students from memory effect, a different arrangement of the word order was used for the delayed post-test. Both the two post-tests were not notified in advance to the participants so that they could perform pure learning outcomes only from the tasks.

Unlike some studies in which non-sense words were used for target vocabulary (Keating, 2008; J. Kim, 2015; J. Park & J. Oh, 2015), in the current study ten target words were selected from the reading text actually used for tasks 1 and 2. For this reason, as described above, to exclude the number of words the participants already knew before the treatment, a pre-test was conducted. The ten target English words varied in parts of speech like verb, noun, adjective, and adverb. The highest possible score of the two post-tests was 10 points. In order to measure the participants' pure learning outcomes from the treatment, the number of words marked 'known' along with Korean meaning in the pretest was excluded from the two post-tests. For the correct answer, 1 point was given, while 0 point was given to the incorrect one. For the cases of writing only Korean translation without designating the part of speech or vice versa, they were considered as incorrect answers. Korean translation without designating the part of speech or vice versa, they were considered as incorrect answers.

This study used three vocabulary learning tasks, partially adapting what Hulstijn and Laufer (2001) established as presented in Table1 in the previous section. That is, according to the involvement load index, three tasks of different level of involvement load were employed: reading & comprehension questions and reading & gap gilling tasks inducing involvement load 1 and 2, respectively, and one writing task inducing involvement load 3. As for the writing task, the researcher designed 'unscrambling sentences' instead of 'writing original sentences' or 'writing a composition. As H. M. Sung (2013) mentioned, for most Korean high school learners are not competent in structuring sentences in English. This may lie in the lack of consciousness or negligence on writing instruction (Lu, 2013). Although the participants of the current study were university students, writing seemed to be an undoable task considering the overall proficiency level of English of the students. Table 2 shows the detailed description of the tasks used in the current study. With regard to task 1, the participants were required to read the given passage, referring to the glosses

which include meanings in Korean and parts of the speech of 10 target words. After that, they had to solve five comprehension questions. As for task 2, the participants had to carry out the same job as task 1 group did basically, but while reading the passage, they had to fill 10 gaps where the target words were printed for task 1 group.

TABLE 2

Vocabulary Tasks With Involvement Load

Task Description	Composition of Involvement Load
Task 1: Reading Comprehension	Involvement Load (1): moderate need (1) + no search (0) + no evaluation (0)
Task 2: Reading & Gap-filling	Involvement Load (2): moderate need (1) + no search (0) + moderate evaluation (1)
Task 3: Unscrambling Sentences	Involvement Load (3): moderate need (1) + no search (0) + strong evaluation (2)

For tasks 1 and 2, a news article about IKEA, a global furniture enterprise was chosen for the reading passage because at the time this research was being conducted, the participants were taking TOEIC course in which they encounter various business related articles or e-mails. Taking the participants' reading comprehension level into account, the article was modified in its length and difficulty level. As Keating (2008) pointed out that 'time on task' was a determinant factor in the effectiveness of involvement load, in the present study time (20 min.) was also thoroughly controlled.

Regarding task 3, the participants had to complete a total of 10 English sentences by unscrambling the given words in a correct order based on 10 Korean sentences. By doing this, they were supposed to use one target word to make each sentence.

3. Data Analysis

The results from the pre-test and two post-tests were analyzed by using SPSS 18. To examine the first research question on whether the vocabulary task types inducing different level of involvement load have main effect on low-level learners' short-term learning and long-term retention, two sets of two-way ANOVA were used: one time for the gains from the immediate posttest and the other time for the gains from the delayed posttest. More concretely, learners' proficiency and task type were set as two independent variables, and their individual effect on the two post-tests was examined. Consequently, post hoc comparisons were conducted to see whether there are any significant differences among the tasks.

With reference to the second research question, the results of the ANOVA were examined in terms of interaction effect between the two factors, task type and proficiency level. That is, the data was analyzed to see whether the factor of task-induced involvement

load influenced on the factor of proficiency level in learners' vocabulary gains from the treatment and vice versa.

IV. RESULTS AND DISCUSSION

1. Task-Induced Involvement Load and Vocabulary Learning

To investigate the research question one, the effects of task-induced involvement load on the learners' vocabulary outcomes were analyzed based on the results of the two posttests: the immediate posttest scores for short-term learning and the delayed posttest scores for long-term retention. In this study task type indicates different level of involvement load, and proficiency level means two groups of low-level learners who were subdivided into intermediate-low and low level.

1) Effects on Short-term Learning

The mean scores of the immediate posttest were analyzed and the descriptive statistics are as follows.

TABLE 3

Descriptive Statistics for Short-term Learning

Prof. Level	Task Type	*n*	*M*	*SD*
Intermediate Low (A)	Reading(1)	30	4.30	1.75
	Gap-filling(2)	30	5.20	1.93
	Unscrambling(3)	27	4.48	2.10
	Total	87	4.67	1.95
Low (B)	Reading(1)	34	4.21	2.32
	Gap-filling(2)	27	4.11	2.06
	Unscrambling(3)	32	2.81	1.86
	Total	93	3.70	2.13
Grand Total		180	4.17	2.12

As Table 3 shows, the mean scores of group A are a bit higher than those of group B across task types. More concretely, mean difference of task 3(1.67) between group A and B was the highest, and that of task 2(1.09) and of task 1(0.09) were followed in order. In the case of group A, they learned more words from task 2 (M = 5.20) than from task 3 (M = 4.48), which only partially supports the Involvement Load Hypothesis (Hulstijn & Laufer, 2001). On the other hand, group B learners' achievements from three tasks were inversely proportional to the level of the involvement load each task induces.

In order to investigate whether the task type affects the participants' vocabulary learning

immediately after treatment, two-way ANOVA was conducted. According to Table 4, not only task type (p = .027) but also proficiency level (p = .002) affected the participants' short-term learning. This means the different level of involvement load had a main effect on the learners' immediate vocabulary outcomes when they were engaged in vocabulary learning tasks. Besides, proficiency level (intermediate-low and low) played a role in differentiating the participants' learning outcomes just after they performed different type of tasks.

TABLE 4

Effects of Involvement Load and Proficiency Level on Short-term Learning

Source	Type III Sum of Squares	*df*	Mean Square	*F*	*Sig.*	Partial Eta Squared
Corrected Model	94.06	5	18.81	4.630	.001	.117
Intercept	3130.85	1	3130.85	770.599	.000	.816
Task Type	29.83	2	14.92	3.692	.027*	.040
Prof. Level	40.39	1	40.39	9.940	.002**	.054
Task * Level	19.49	2	9.75	2.399	.084	.027
Error	709.94	174	4.06			
Total	3926.00	180				
Corrected Total	801.00	179				

*p < .05, **p < .005

This finding suggests that vocabulary learning tasks established by Hulstijn and Laufer (2001) might even novice learners to concentrate on word forms. It is assumed that in the current study almost half (group B) of the participants are not able to understand a single English sentence properly. Thus, it is considered that the significant role of task-induced involvement load in the immediate learning can be attributed to the outcomes of intermediate-low level (group A) learners. Group B learners might have focused more on the glossary part than the reading passage. This means it is highly likely that the novice learners spent more time in getting information provided in Korean from the word list than struggling to translate incomprehensible English sentences of the reading passage. In short, it is interpreted that the results of the research question one may be caused by the feature of the tasks not by the involvement load induced by the tasks.

A post hoc test was carried out to examine whether any significant differences exist among the three different tasks. Table 5 shows the results of post hoc comparison (Tukey HSD) on the vocabulary outcomes gained immediately after task performance.

TABLE 5

Post Hoc Comparison Among Tasks on Short-term Learning

			M	*SE*	*Sig.*
Tukey HSD	Reading(1)	Gap-filling(2)	-.43	.367	.465
		Unscrambling(3)	.67	.364	.156
	Gap-filling(2)	Reading(1)	.43	.367	.465
		Unscrambling(3)	1.11	.374	.010*
	Unscrambling(3)	Reading(1)	-.67	.364	.156
		Gap-filling(2)	-1.11	.374	.010*

Note. The numbers in the parentheses indicate the involvement load index.
*$p < .05$

The multiple comparison results reveal that there exist significant differences between the mean scores of task 2 and task 3. However, no significant gaps were found between tasks 1and 2 as well as between tasks 1and 3.

2) Effects on Long-term Retention

Table 6 demonstrates the results of the descriptive statistics on the scores obtained from the delayed posttest conducted three weeks after task treatment.

TABLE 6

Descriptive Statistics for Long-term Retention

Prof. Level	Task Type	*n*	*M*	*SD*
Intermediate Low (A)	Reading(1)	30	1.77	1.74
	Gap-filling(2)	30	1.90	1.97
	Unscrambling(3)	27	2.11	1.48
	Total	87	1.92	1.73
Low (B)	Reading(1)	34	1.74	1.60
	Gap-filling(2)	27	1.37	1.57
	Unscrambling(3)	32	1.38	1.66
	Total	93	1.51	1.61
Grand Total		180	1.71	1.68

As were in the short-term learning, the mean scores of group A are higher than those of group B across task types. Among the tasks, mean difference of task 1 (0.03) between group A and B was the lowest, and that of task 2 (0.53) and of task 3 (0.73) were followed in order. Unlike short-term learning immediately after task performance, in long-term retention group A showed the highest mean score ($M = 2.11$) from task 3, which entirely confirms the Involvement Load Hypothesis (Hulstijn & Laufer, 2001). As for group B, as did in the immediate learning, the learners retained target words best from task 1 ($M = 1.74$) and showed similar mean scores from tasks 2 and 3 ($M = 1.37$, $M = 1.38$).

To examine the effects of task type on the learners' long-term retention, another two-

way ANOVA was conducted. As Table 7 displays, unlike the results on short-term retention, neither task type nor proficiency level affected the learners' recall on target words 3 weeks after treatment. In other words, different level of involvement load induced by vocabulary tasks did not function in differentiating the learners' retention on target words. Likewise, different proficiency level did not play an effective role in facilitating the learners' 3-week retention of new words.

TABLE 7

Effects of Involvement Load and Proficiency Level on Long-term Retention

Source	Type III Sum of Squares	*df*	Mean Square	*F*	*Sig.*	Partial Eta Squared
Corrected Model	12.25	5	2.45	.868	.504	.024
Intercept	522.51	1	522.51	185.112	.000	.515
Task Type	0.49	2	0.24	.807	.917	.001
Prof. Level	8.35	1	8.35	2.960	.087	.017
Task * Level	4.04	2	2.02	.716	.490	.008
Error	491.15	174	2.82			
Total	1027.00	180				
Corrected Total	503.39	179				

* $p < .05$

These results are not in line with the findings by some studies recently conducted (H. Kim & Y. Na, 2010; H. M. Sung, 2013). In those studies, according to task type (involvement load), the participants' achievements varied significantly in their long-term retention as well as the short-term learning. Such discrepant findings can be addressed in related to the participants' proficiency level in each study. Unlike the previous studies that recruited all proficiency levels as participants, the current study targeted only low-level learners. Thus, the difference of involvement load might not have been enough for low-level learners to remember target words for a longer period of time although they benefitted from it for their immediate learning after task performance. Furthermore, as discussed in the previous section, it is assumed that even the influential role of task-induced involvement load in the immediate learning might be attributed to the achievement of intermediate-low level (group A) learners. By this token, it can be carefully assumed that Involvement Load Hypothesis is only applicable to those who are above a certain level of English proficiency: the learners with at least sentence writing ability like group A learners in the current study can get benefits from task-induced involvement load.

As done for the immediate test, post-hoc test was administered to see whether any significant differences took place among the tasks. Table 8 presents the multi-comparison results of post hoc comparison (Tukey HSD) on the learners' retention, but no significant differences among the tasks were found.

TABLE 8

Post Hoc Comparison Among Tasks on Long-term Retention

			M	*SE*	*Sig.*
Tukey HSD	Reading(1)	Gap-filling(2)	.10	.306	.942
		Unscrambling(3)	.04	.303	.991
	Gap-filling(2)	Reading(1)	-.10	.306	.942
		Unscrambling(3)	-.06	.312	.978
	Unscrambling(3)	Reading(1)	-.04	.303	.991
		Gap-filling(2)	.06	.312	.978

* $p < .05$

Apart from the effects of the task-induced involvement on the two posttests dealt with so far, the role of task-induced involvement load in this study can be discussed based on the figures of the descriptive statistics presented on Tables 3 and 6. In the present study the participants were all low-level learners and were subdivided into two groups: intermediate-low and low. This is because the researcher intended to examine how the task-induced involvement functions to low proficiency students in a fine-tuned manner. Although group A participants were labeled as 'intermediate-low,' their average TOEIC score 524 was almost similar to that of beginner level in the study of H. Kim and Y. Na (2010). Group B labeled as 'low' in this study was novice learners whose average TOEIC score was 290. Methodologically speaking, the distinction between group A and B was made based on the criteria whether the learners have sentence constructing ability or not. From the two posttests, on the whole group A produced better gains than group B even though the gap of the score range was not substantial. In particular, group A recalled the highest number of words from task 2 in the immediate posttest but from task 3 in the delayed posttest. This can be interpreted that the learners with at least the basic skill of sentence writing can get benefits from a task of a higher involvement load in remembering target words over time. On the other hand, group B demonstrated the opposite tendency in terms of the IHL: the participants yielded the highest mean score from task 1 and the lowest mean score from task 3 in the immediate learning. That is, reading task is the most facilitative one among the three tasks in recalling new words for novice proficiency learners who are not able to construct even an individual sentence. In other words, there is a high possibility that novice learners in the current study might just have seen and learned the given information of the target words regardless of understandability on the reading passage.

To put the findings in another way, the involvement load index of individual component, in particular, evaluation, set for each task can be a moot point. The three tasks employed in this study varied in terms of the involvement load: they differed only because of the presence or degree of 'evaluation' component. According to the task-induced involvement index proposed by Hulstijn and Laufer (2001), gap-filling and unscrambling sentences are expected to cause 'moderate' evaluation (1) and 'strong' evaluation (2), respectively.

However, the degree of moderate (1) seemed to function as greater degree than it was supposed to do. In the current study, the participants' vocabulary gains from task 2 were noticeable. More concretely, group A performed the highest score from task 2 in their short-term learning, which accords with the study by H. M. Sung (2013). Learners can remember new words well if vocabulary instruction involves a focus-on form factor (Laufer, 2005). In this vein, gap-filling task requires the learners to pay attention to the lexical items as much as writing task does. Thus, the learners can more focus on the words in the process of selecting a proper word to fill each gap, which leads to greater vocabulary gains (Keating, 2008). In addition, group B showed almost similar mean scores from tasks 2 and 3 in their long-term retention. This phenomenon might be explained that the learners' mental efforts (evaluation) while performing gap-filling task could be stronger than that originally suggested by Hulstijn and Laufer (2001). This point can be plausible considering what the learners actually have to do when they are engaged in gap-filling task: they need to read a reading passage with 10 gaps, compare all the words given, and select the appropriate one according to the context of each gap, understanding the whole sentence at the same time. That being said, it is suggested that gap-filling task can be used for those who have sentence writing skill like group A in this study if teachers aim to help their students to get an immediate learning effect for a certain group of new words. On the other hand, to support novice learners like group B in this study, teachers can use reading tasks positively because at least the current study proved that novice learners recalled the target words best not only in the immediate test (short-term learning) but also in the delayed test (long-term retention).

2. Interaction Effect Between Involvement Load and Proficiency Level

To check whether the interaction effect took place between the two factors, involvement load and proficiency level, in the learners' vocabulary outcomes, the results of the two-way ANOVAs presented in the previous section were examined from a different point of view. The following two tables (Tables 9 and 10) are basically same as Tables 4 and 7 but with the underline for the wording, Task * Level.

TABLE 9

Interaction Effect Between Involvement Load and Proficiency Level on Short-term Learning

Source	Type III Sum of Squares	*df*	Mean Square	*F*	*Sig.*	Partial Eta Squared(η^2)
Corrected Model	94.06	5	18.81	4.630	.001	.117
Intercept	3130.85	1	3130.85	770.599	.000	.816
Task Type	29.83	2	14.92	3.692	.027*	.040
Prof. Level	40.39	1	40.39	9.940	.002**	.054
Task * Level	19.49	2	9.75	2.399	.084	.027
Error	709.94	174	4.06			
Total	3926.00	180				
Corrected Total	801.00	179				

*$p < .05$, **$p < .005$

According to Table 9, no interaction effect ($F = 2.399$, $p > .05$) between task type and proficiency level occurred in the learners' vocabulary learning immediately after task performance. This means that the ratios of mean differences between the same task groups of two different proficiency levels were constant. In other words, the effect of involvement load induced by three vocabulary tasks in the learners' short-term learning was not influenced by the factor of proficiency level.

The numbers on the right column labeled 'Partial Eta Squared (η^2) show the strength of association. In the short-term learning, the effect of proficiency level was much stronger (0.54) than that of involvement load (0.04), which caused a significant effect of proficiency level ($p = 0.002$) on short-term learning.

TABLE 10

Interaction Effect Between Involvement Load and Proficiency Level on Long-term Retention

Source	Type III Sum of Squares	*df*	Mean Square	*F*	*Sig.*	Partial Eta Squared(η^2)
Corrected Model	12.25	5	2.45	.868	.504	.024
Intercept	522.51	1	522.51	185.112	.000	.515
Task Type	0.49	2	0.24	.807	.917	.001
Prof. Level	8.35	1	8.35	2.960	.087	.017
Task * Level	4.04	2	2.02	.716	.490	.008
Error	491.15	174	2.82			
Total	1027.00	180				
Corrected Total	503.39	179				

*$p < .05$

As was not in the short-term learning, Table 10 shows that interaction effect ($F = 0.716$, $p > .05$) between task type and proficiency level was not found in long-term retention. Again, this indicates that the effect of task-induced involvement load on the learners' word

recall after 3 weeks was not interrupted by the difference of the participants' proficiency level. According to the data shown under η^2, the strength of proficiency level was stronger (0.017) than that of involvement load (0.001) as was in short-term learning. However, in long-term retention, the strength gap between proficiency level and task type did not lead to a statistical significance ($p = 0.490$): proficiency level did not have a main effect in the learners' long term retention as task type did not.

The fact that the task-induced involvement load and the learners' proficiency level in this study did not interact each other on the two posttests would shed a light on the point that we need to reconsider the function of task-induced involvement to the low-level learners. Hulstijn and Laufer (2001) put forth the ILH without limiting learners' proficiency: they claimed that a more involving task leads to better vocabulary retention 'if learners are capable of performing the task.' Then it is highly likely that now we need to more focus on the learners' condition, more specifically, their proficiency level. In the discussion on their research findings, H. Kim and Y. Na (2010) mentioned that for beginner and intermediate level learners, designing vocabulary tasks in a more elaborated manner is definitely necessary. In the same vein, the present study recommends teachers to consider the students' proficiency level, especially for low-level group when they design vocabulary tasks in their classroom.

All participants of the current study could belong to the same proficiency group (low) largely if their TOEIC scores are only considered. However, this study revealed that they have to be subdivided and treated separately in applying vocabulary tasks since there apparently exist some tasks that one group can do but the other group can not. At least, the difference of proficiency level did work in the learners' short-term learning outcomes in the current study. Thus, teachers have to carefully differentiate novice learners from those who are a bit beyond the novice level before asking to perform a certain task.

V. CONCLUSION

This study purposed to testify the ILH especially focusing on low proficiency level. Also, the current study examined whether the interaction effect between the two factors, proficiency level and task-induced involvement load takes place when the learners are engaged in task performance. In doing so, two posttests immediately and three weeks after the treatment were administered to measure the learners' short-term learning and long-term retention. And then the scores were analyzed by using two sets of two-way ANOVA.

With reference to the research question one, difference of task-induced involvement load affected the learners' short-term learning of target words, but it did not on long-term retention. Likewise, the learners' different proficiency level influenced on their short-term

learning of target words, but did not take effect on long-term retention. As for the research question two, no interaction was found between the two factors, task type and proficiency level: the factor of involvement load did not affect the factor of proficiency level in the learners' vocabulary short-term learning and long-term retention.

The findings of this study evinced that the task-induced involvement load is not influential enough for the learners to recall new words for a longer period of time. Unlike other previous studies (H. Kim & Y. Na, 2010; Y. Kim, 2008), the participants of the present study were all low-level learners. In this sense, the involvement load might have been less crucial in the current study than previous studies.

Ruminating on what this study found, some practical implications can be addressed to EFL teachers. Firstly, tasks carrying 'evaluation' component such as gap-filling or unscrambling sentences could be applicable to the learners who have the basic ability to write individual sentences although their overall English proficiency is low. This may be because 'evaluation' is the qualification or capacity that can be expected from the learners with a certain level of English proficiency.

Secondly, for the learners who are not able to construct a single sentence, familiar tasks with a small amount of involvement load like reading exercises along with words list could be beneficial. Novice learners may not be capable to handle tasks imposed high level of cognitive load (Paas & Van Merriënboer, 1994). Thus, less anxiety and less involving tasks are strongly recommended for novice learners.

Thirdly, to facilitate better retention on new words, what students learned from task performance has to be retrieved and repeated at regular intervals. Here, finding out the best time interval falls to each individual teacher's share. It holds true that the Involvement Load Hypothesis is an insightful and functional suggestion, but it still remains controversial on how we ensure the designated degree of a certain component. For example, we are not sure of whether 'moderate (1)' or 'strong (2)' of 'evaluation' component actually generates the estimated involvement load as much as it was literally set. Therefore, it is necessary to design optimum tasks considering the learners' proficiency level or capacity, wisely adopting the fundamental contention that Hulstijn and Laufer (2001) posit.

Since the participants of the current study were recruited based on convenience sampling not random basis, findings and issues discussed so far may not be applied for all learners in general. Another potential limitation is that in this study there might be the probability for the participants to learn or repeat the target words outside the classroom for a three-week period. Therefore, it is expected that further research can redeem these shortcomings.

REFERENCES

Akbarian, I. (2010). The relationship between vocabulary size and depth for ESP/EAP learners. *System, 38*(3), 391-401.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 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*(3), 286-294.

Ellis, R. (1994). Implicit/explicit knowledge and language pedagogy. *TESOL Quarterly, 28*(1), 166-172.

Folse, K. S. (2004). *Vocabulary myths*. Ann Arbor: University of Michigan Press.

Folse, K. S. (2006). The effect of type of written exercise on L2 vocabulary retention. *TESOL Quarterly, 40*(2), 273-293.

Gu, Y., & Johnson, R. (1996). Vocabulary learning strategies and language learning outcomes. *Language Learning, 46*(4), 643-679.

Hahn, Hyeryeong. (2014). Korean EFL learners' lexical development: Evidence from a word association test. *Modern English Education, 15*(1), 27-48.

Huckin, T., & Coady, J. (1999). Incidental vocabulary acquisition in a second language: A review. *Studies in Second Language Acquisition, 21*(2), 181-193.

Hulstijn, J. H. (2003). Incidental and intentional learning. In C. J. Doughty & H. M. Long (Eds.), *Handbook of second language research* (pp. 349-381). Malden, MA: Blackwell.

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

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

Kim, Hyungsun, & Na, Yoonhee. (2010). Vocabulary learning and task-induced involvement. *Korean Journal of Applied Linguistics, 26*(4), 183-211.

Kim, Jieyoung. (2015). The effects of task-induced involvement and learners' English proficiency on English vocabulary learning. *English Language Teaching, 27*(3), 125-149.

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

Kojig-Sabo, I., & Lightbown, P. M. (1999). Students' approaches to vocabulary learning and their relationship to success. *The Modern Language Journal, 83*(2), 176-192.

Laufer, B. (2003). Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence. *Canadian Modern Language Review, 59*(4), 567-587.

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

Lee, Seongwon, & Min, Miok. (2006). A study on Korean high school students' use of vocabulary learning strategies. *English Teaching, 61*(2), 115-137.

Lu, Minhui. (2013). Effects of four vocabulary exercises on facilitating learning vocabulary meaning, form, and use. *TESOL Quarterly, 47*(1), 167-176.

Nagy, W. E., & Herman, P. A. (1987). Breadth and depth of vocabulary knowledge: Implication of acquisition and instruction. In M. G. McKeown & M. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 19-35). Hillsdale, NJ: Erlbaum.

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

Paas, F., & Van Merriënboer, J. (1994). Variability of worked examples and transfer of geometrical problem solving skills: A cognitive load approach. *Journal of Educational Psychology, 86*(1), 122-133.

Park, Juyeon, & Oh, Junil. (2015). Effects of task-induced vocabulary learning involvement on vocabulary learning by Korean university students. *English Language Teaching, 27*(3), 175-191.

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

Selfidge, O. G., & Neisser, U. (1960). Pattern recognition by machine. *Scientific American, 203*, 60-68.

Sung, HyunMi. (2013). Task-induced involvement load in Korean EFL incidental vocabulary learning. *The Journal of Studies in Language, 29*(2), 269-296.

Sutherland, N. S. (1968). Outlines of a theory of visual pattern recognition in animals and man. *Proceedings of the Royal Society of London B: Biological Sciences, 171*(1024), 297-317.

Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in Second Language Acquisition*, *19*(3), 287-307.

APPENDIX 1

Task 1 Reading Comprehension

아래의 지문을 읽고 질문에 답하시오.

Natura is a fast-growing Brazilian cosmetics and toiletries company which started as a small laboratory in a garage in Sãn Paulo 37 years ago. Today, it is trying to go international. *Natura*

picked France as the first country outside Latin America to try out its ideas. "We wanted to choose a sophisticated market where people want things to be very good," says Alessandro Carlucci, *Natura*'s Chief Executive. "We wanted a test that was tough."

Mr Carlucci says the experiment was successful. In five years, he says, *Natura* wants to have 'at least' 10 percent of its revenues coming from outside Brazil. Apart from Brazil and its small office in France, *Natura* currently sells its cosmetics in a few other South American countries.

The company promotes itself as an ethical company that works with growers who harvest their products in an environmentally sensitive way. It also uses biodegradable packaging.

- toiletries (명사) 세면도구류
- laboratory (명사) 실험실
- sophisticated (형용사) 세련된, 교양 있는
- experiment (명사) 실험
- revenue (명사) 수익, 매출
- cosmetics (명사) 화장품
- ethical (형용사) 윤리적인
- harvest (동사, 명사) 농작물을 수확하다, 수확
- environmentally (부사) 환경(보호)적으로
- biodegradable (형용사) 자연분해성의

1. Natura에 관한 설명 중 옳지 않은 것은?

① 화장품과 세면도구류를 판매하는 회사이다.
② 프랑스에서 처음 시작된 회사이다.
③ 브라질 이외 지역의 매출을 계획한다.
④ 자연친화적인 포장재를 사용한다.

2. 위의 글에 나오는 다음의 단어들 중 의미적으로 가장 연관이 없어 보이는 것은?

① ethical ② environmentally ③ sophisticated ④ biodegrabable

3. 위의 글의 내용으로 보아 현재 Natura 제품이 판매되는 지역이 아닌 곳은?

① Brazil ② France ③ a few countries in Latin America ④ South Africa

4. 위의 글에서 언급되지 않은 내용은?

① Natura의 건립역사 ② Natura의 주요제품
③ Natura의 기업정신 ④ Natura의 현재매출

5. 위 글의 성격과 가장 가까운 것은?

① persuasive ② informative ③ biographic ④ poetic

APPENDIX 2

Task 2 Reading Comprehension & Gap-filling

다음 지문의 문맥과 내용을 이해하여 1) 아래에 주어진 단어들로 빈칸을 채우고
2) 질문에 답하시오.

Natura is a fast-growing Brazilian cosmetics and _____ company which started as a small _____ in a garage in Sān Paulo 37 years ago. Today, it is trying to go international. *Natura* picked France as the first country outside Latin America to try out its ideas. "We wanted to choose a _____ market where people want things to be very good," says Alessandro Carlucci, *Natura*'s Chief Executive. "We wanted a test that was tough."

Mr Carlucci says the _____ was successful. In five years, he says, *Natura* wants to have 'at least' 10 percent of its _____ coming from outside Brazil. Apart from Brazil and its small office in France, *Natura* currently sells its _____in a few other South American countries.

The company promotes itself as an _____ company that works with growers who _____ their products in an _____ sensitive way. It also uses _____ packaging.

- laboratory (명사) 실험실
- environmentally (부사) 환경 보호적으로
- biodegradable (형용사) 자연분해성의
- cosmetics (명사) 화장품
- harvest (동사, 명사) 농작물을 수확하다, 수확
- revenue (명사) 수익, 매출
- ethical (형용사) 윤리적인
- sophisticated (형용사) 세련된, 교양 있는
- toiletries (명사) 세면도구류
- experiment (명사) 실험

1. Natura에 관한 설명 중 <u>옳지 않은 것은?</u>

① 화장품과 세면도구류를 판매하는 회사이다.
② 프랑스에서 처음 시작된 회사이다.
③ 브라질 이외 지역의 매출을 계획한다.
④ 자연친화적인 포장재를 사용한다.

2. 위의 글에 나오는 다음의 단어들 중 <u>의미적으로 가장 연관이 없어 보이는 것은?</u>

① ethical ② environmentally ③ sophisticated ④ biodegrabable

3. 위의 글의 내용으로 보아 현재 Natura 제품이 판매되는 지역이 <u>아닌 곳은?</u>

① Brazil ② France ③ a few countries in Latin America ④ South Africa

4. 위의 글에서 언급되지 않은 내용은?

① Natura의 건립역사 ② Natura의 주요제품
③ Natura의 기업정신 ④ Natura의 현재 매출

5. 위 글의 성격과 가장 가까운 것은?

① persuasive ② informative ③ biographic ④ poetic

APPENDIX 3

Task 3 Unscrambling Sentences

※ 우리말 해석기준으로 상자 안의 단어들을 순서대로 배열하여 맞는 영어 문장을 써 보세요.

1) laboratory (명) : 실험실

•그의 실험실은 다양한 화학물질로 가득 차 있다.

is	full	of	various	chemicals	(his)	laboratory

→His_____

2) revenue (명) : 수익, 매출

•그 회사의 연간 매출은 매우 높다.

annual	revenue	(the)	company's	high	very	is

→The_____

3) environmentally (부) : 환경적으로

•그의 집은 환경적으로 좋은 곳에 위치해 있다.

in	located	environmentally	(his)	house	is	nice	a	area

→His_____

4) ethical (형) : 윤리적인

•나는 이것이 윤리적인 문제라고 생각한다.

issue	think	(I)	this	an	ethical	is

→I_____

5) biodegradable (형) : 자연분해성의

•우리 공장은 오직 자연 분해성의 제품들을 만든다.

only	manufactures	factory	(our)	products	biodegradable

→Our_____

6) sophisticated (형) : 세련된, 교양 있는

•그녀는 잘 생기고 세련된 남자들을 좋아한다.

handsome	sophisticated	and	likes	(she)	men

→She_____

7) cosmetics (명) : 화장품

• 그녀는 화장품 사는데 많은 돈을 쓴다.

Spend	money	much	buying	in	(she)	cosmetics

→She_____

8) toiletries (명) : 세면도구

•그 호텔은 세면도구를 제공하지 않는다.

provide	toiletries	hotel	(the)	does not

→The_____

9) harvest (동) : (농작물을) 수확하다

•그들은 일 년에 두 번 수확한다.

a	year	twice	harvest	(they)

→They_____

10) experiment (명) : 실험

•그들의 실험은 완전히 성공적이었다.

successful	experiment	(their)	was	completely

→Their_____

Examples in: English
Applicable Languages: English
Applicable Levels: Secondary/Tertiary

HyunMi Sung
Institute of General Education
Kunsan National University
558, Deahak-ro, Kunsan, Jeonbuk, 54150, Korea
Tel: 010-2354-3369
E-mail: hmsung@kunsan.ac.kr

Received 13 September 2016
Revised 23 October 2016
Accepted 11 November 2016