KOREAN JOURNAL OF ENGLISH LANGUAGE AND LINGUISTICS

The Effects of Writing Tasks on Vocabulary Learning among Korean EFL Learners and Involvement Load Hypothesis*

Chaehee Park (Sunmoon University)

Park, Chaehee
Associate Professor,
Division of Foreign Languages
Sunmoon University
70, Sunmoon-ro 221 Beon-gil,
Tangjeong-myeon, Aasna-si
Chungnam, Korea, 31460
Tel: 041) 530-2423
Email: chpark@sunmoon.ac.kr

ABSTRACT

Park, Chaehee. 2024. The effects of writing tasks on vocabulary learning among Korean EFL learners and involvement load hypothesis. *Korean Journal of English Language and Linguistics* **24, 554-567.**

The purpose of this study is to test whether the same amount of load of the two tasks contributes to the same amount of vocabulary learning among Korean EFL learners. Twenty nine Korea EFL university students participated in the task experiments (15 for sentence writing task, 14 for composition writing task). The participants' performance was tested three times (pre-test, immediate post-test, and delayed post-test). It was found that both sentence writing and composition writing task were similarly effective for initial word-learning (short-term memory), whereas the latter group was stronger in word-retention (long-term memory) than the former group. The results indicated that the composition writing task requires learners to process target words in a deeper and meaningful way than the sentence writing task. It was suggested that the learners' information process mechanism for unknown words is different in two tasks: somewhat different chunking in an associated context and hierarchical organization of information in two tasks also yielded different word learning (Zou 2017). Pedagogical implications for researchers and educators are discussed within the framework of involvement load hypothesis and information process mechanism.

KEYWORDS

involvement load hypothesis, L2 vocabulary learning, task load, second language learning

1. Introduction

A critical role of vocabulary in L2 literacy development cannot be emphasized too much because it is one of the important predictors of reading comprehension and L2 development (Hamada and Park 2011, 2013, Hu and Nassaji 2014, Nation 2000, 2001, Pulido 2007, 2009). A great deal of research has discussed how vocabulary is learned and what processes are involved during learning. One of the hypotheses regarding vocabulary learning is involvement load hypothesis (ILH, Laufer and Hulstijn 2001), which is fundamentally affected by the notion of processing hypothesis. Laufer and Hulstijn (2001) suggested ILH that holds vocabulary learning is conditional upon task's involvement load. A learning task which induces learners' motivational and cognitive efforts is comprised of components such as *need*, *search* and *evaluation*. The degree of these components vary and can be quantified in number: none 0 (-), moderate 1 (+), and strong 2 (+ +). A few experimental studies testing the effectiveness of these components confirmed that the task with much amount of component load was more effective in vocabulary learning than the task with less amount of component load (Hulstijn and Laufer 2001, Kang and Shin 2019, Keating 2008, Kim 2008, Kim and Na 2010, Park 2017, Park 2020a, Park 2020b, Sung 2013).

In the same vein, one of the assumptions to be tested within the framework of ILH is that two tasks inducing the same quantified amount of involvement load would lead to the same amount of vocabulary learning (Laufer and Hulstijn 2001). But, little research has been so far carried out to test this assumption, even showing inconclusive results: Kim (2008) confirmed two tasks (sentence writing vs. composition writing), assumed to involve the same amount of load (Laufer and Hulstijn 2001) similarly contributed to word learning whereas Zou's (2017) findings suggest the composition writing is more effective for word learning than the sentence writing due to different information processing mechanism. This inconsistent research findings require further investigation on the task load and the degree of learners' involvement in order to provide more thorough insight on vocabulary learning within the framework of ILH. The purpose of this experiment is to test whether the same amount of load of the two tasks (sentence writing vs. composition writing) contributes to the same amount of vocabulary learning among Korean EFL learners.

2. Literature Review

2.1 Involvement Load Hypothesis (ILH)

Laufer and Hulstijn (2001) proposed assumptions on involvement, involvement load, and task-induced involvement load. Regarding the assumption of task involvement, some factors involved in a task were identified as affecting the retention of words. First, they suggested "retention of words when processed incidentally, is conditional upon the following factors in a task: *need*, *search*, and *evaluation*" (Laufer and Hulstijn 2001, p.14). In other words, vocabulary task itself requires learners to follow a given instruction to successfully complete it, during which involvement in the task entails learners' motivational *need* and cognitive *search* and *evaluation*. *Need* component is motivational, non-cognitive dimension of involvement in a task and described as "a desire to comply with the requirements of the task" (Kim and Na 2010, p.187). That being said, it is based on a drive of how strongly learners need to know the meaning of target words to accomplish the given task. The other two components, *search* and *evaluation* are related to learners' cognitive dimensions (word-information processing), which depends on learners' noticing and allocating attention to the form-meaning relationship (Schmidt 1994,

2000). *Search* refers to an attempt learners try to make to find the meaning of an unknown L2 word or word form expressing a concept. It usually involves learners' activity in which they consult a dictionary or a teacher for unknown words. *Evaluation* component requires a "comparison of a given word with other words, a specific meaning of a word with its other meanings, or combining the word with other words in order to assess whether a word does or does not fit its context" (Laufer and Hulstijn 2001, p.14). One example is that in a gap-filling task, learners are required to choose 10 words out of 15, and L1 word meanings are glossed as well. The provided 15 words need to be evaluated against each other and the most suitable 10 ones have to be selected for blanks. Another example is an L2 sentence writing task in which L1 word-meaning is provided. In this case, target words need to be evaluated with learners' additionally chosen words for sentence writing. In order to evaluate target words, it requires a learner "some kind of selective decision based on a criterion of semantic and formal appropriateness of the words and its context" (Laufer and Hulstijn 2001, p.15).

The second assumption regarding involvement load said "other factors being equal, words which are processed with higher involvement load will be retained better than words which are processed with lower involvement" (Laufer and Hulstijn 2001, p.15). This assumes that each component has its own degree of learners' involvement in the task. The degree of *need* component is either *moderate* (1) when imposed by an external agent (e.g., requiring students to fill in a word in sentence) or *strong* (2) when imposed by a learner himself or herself (e.g., looking up a word in an L1-L2 dictionary to accomplish the task). In the case of *search*, the degree is conceptualized as either *present* (1) when learners must seek the meaning of unknown L2 words by consulting a dictionary or teachers, or *absent* (0) when no such effort is required (e.g., the meanings of unknown words are provided). The degree of *evaluation* can be either *moderate* (1) when it entails recognizing differences between words (e.g., deciding which meaning of a target word best fits the context) or *strong* (2) when the task requires learners' evaluation on new words while combining them with known words in original contexts (e.g., sentence writing and composition writing). The total degree of these components in a task can be quantified, and it has been well grounded that the words processed in a writing task (including unscrambling sentences and fill-in task) involving higher involvement load are retained to learners better than the words processed through the reading comprehension task which involves lower involvement load (Hulstijn and Laufer 2001, Keating 2008, Kim 2008, Kim and Na 2010, Park 2017, Sung 2013, 2019), supporting the second assumption.

Table 1. Task- induced Involvement Load

	Task Type	Status of Target Words	Need	Search	Evaluation
1	Reading and Com. question	Glossed in text but irrelevant to task	-	-	-
2	Reading and Com. question	Glossed in text and relevant to task	+	-	-
3	Reading and Com. question	Not glossed but relevant to task	+	+	-
4	Reading and Com. question and filling gaps	Relevant to reading comprehension Listed with glosses at the end of the text	+	-	+
5	Writing original sentences	Listed with glossed	+	-	++
6	Writing a composition	Listed with glossed	+	-	++
7	Writing a composition	Concepts selected(and looked up) by L2 learner-writer	++	+	++

(Adjusted from Laufer and Hulstijn 2001, p.18)

Table 1 illustrates some examples of task-induced involvement proposed by Laufer and Hulstijn (2001). Two symbols (+) and (-) indicate '*presence*' or '*absence*' of each component, and a single (+) shows a moderate degree

and a two plus symbol (+ +) denotes a rather strong degree of involvement. The total number of plus symbol (+) represents the involvement index, from which it is assumed that task seven has the strongest involvement because it has a total of five as the involvement index. In the case of a reading task one where unknown words are glossed for the student but not relevant, the comprehension questions can be answered without reference to these target words because those questions do not induce any *need* to focus on the glossed words, nor any *search* for their meaning. A reading comprehension task with glossed words that are relevant to the questions (task 2) will induce a moderate *need* to look at the glosses, but it does not induce any cognitive efforts such as *search* and *evaluation*. A reading task with no glossed but relevant to the reading questions (task 3) will induce *need* as well as *search* because the meaning of the unknown words is required to answer the questions. When target words are deleted in a reading task and listed at the bottom of the text (task 4), it requires learners to fill the text gaps with the correct words. When adding more words to the list which are not relevant to the text, the task will become more difficult. This filling gap task induces a moderate *need*, no *search* (word meaning is also provided), and a moderate *evaluation* because all the words provided with meaning need to be evaluated against each other in the context. In the same vein, task five (writing original sentences) and task six (writing a composition) require the same degree of involvement index of total three. Both require moderate *need*, no *search* (since word meaning is provided) and *strong evaluation* (since target words are used in learner-generated contexts with other words).

Laufer and Hulstijn' (2001) argued about the effectiveness of a task on a basis of involvement load, saying that "other factors being equal, teacher-designed tasks with a higher involvement load will be more effective for vocabulary retention that tasks with a lower involvement load" (p.17), and similarly, the two tasks that induce the same amount of involvement index should promote similar results of vocabulary acquisition. Empirical studies have been carried out to test these assumptions.

2.2 Experimental Studies Testing ILH

A great deal of research has confirmed ILH, the learners who deal with the task with much amount of task load were successful in vocabulary learning than the learners dealing with the task with less amount of task load (Hulstijn and Laufer 2001, Keating 2008, Kim 2008, Kim and Na 2010, Lee and Kim 2015, Park 2017). Initially, Hulstijn and Laufer (2001) conducted an experiment testing whether the three vocabulary learning tasks with different amount of involvement load (reading comprehension, comprehension plus filling in target words, and composition writing with target words) would result in different amount of vocabulary learning. Two groups of advanced university learners participated in the tasks: Israeli (Hebrew L1) and Netherlands (Dutch L1), and the results of the Hebrew-English experiment showed that the target words processed with higher involvement load (composition writing) were retained better than the target words processed with lower involvement load (reading comprehension), fully supporting the ILH. However, the results from the Dutch-English learners partially supported the involvement load hypothesis; the composition writing group (task load 3) yielded higher retention than the reading comprehension group (task load 1) and the comprehension plus filling group (task load 2), but showing no differences between the comprehension group and the comprehension plus filling group. Following this study, many studies tested the ILH and confirmed the overall effectiveness of task involvement load in vocabulary learning (Keating 2008, Kim 2008, Kim and Na 2010, Lee and Kim 2015, Park 2019, 2020b), and also suggested an alternative explanation regarding the degree of involvement index according to cognition and motivation. For instance, Park et al. (2019) conducted an experiment with two tasks which involve the same amount of load but different in the distribution of components, and found the cognitive-focused task (writing task) was more effective for vocabulary learning than the motivation-focused task (reading task).

The effectiveness of writing task on word learning has been found in many other studies. For instance, Teng and Zhang (2023) reported that sentence writing task during the process of multimedia input of lexical items was effective on word learning. In their empirical studies on learning vocabulary through sentence writing task and multimedia input (definition), 235 Chinese EFL learners participated in three different task conditions (definition + sentence writing, definition + background information + sentence writing, definition + background information + video + sentence writing). It was found that the sentence writing task with target word information with video was the most effective for learning lexical items. It has been well established that the writing task of either sentence-level or composition-level is more effective than the reading comprehension task since the former usually induces stronger involvement than the latter. However, in regard to comparison of two different writing tasks (sentence writing vs composition writing), there has been so far inconclusive results. Kim (2008) conducted two cognition-focused tasks (sentence writing vs. writing composition) to ESL learners and found the two tasks similarly contributed to vocabulary learning, supporting the assumption that the two tasks with the same amount of involvement index promote similar results of vocabulary learning (Laufer and Hulstijn 2001). However, Zou (2017) proposed composition writing is the best for vocabulary learning. She developed three tasks (cloze exercise, sentence writing, and composition writing) to examine the effectiveness of each task on vocabulary learning. 147 Chines EFL students participated in the tasks and it was found that the sentence writing task and the composition writing task were much more effective for vocabulary learning compared to the cloze exercise task, which supports ILH. But the differences between two writing tasks were also significant, suggesting the composition writing is much more effective for vocabulary learning than the sentence writing. This inconclusive research findings require further examination on the effectiveness of writing tasks on vocabulary learning. Little research comparing the effectiveness of these two writing tasks, particularly in the Korean EFL context has been carried out, and it is necessary to document more research findings to provide more thorough insight on vocabulary learning within the framework of ILH. Following research questions were posed:

1. Do the tasks of sentence writing and composition writing which involve the same amount of load contribute to the similar initial L2 vocabulary learning?
2. Do the tasks of sentence writing and composition writing which involve the same amount of load contribute to the similar retention of L2 vocabulary learning?

3. The Study

3.1 Participants

The participants in this study were 29 EFL undergraduate students who enrolled in courses for English major at the mid-sized university located in Korea. They were all studying English studies as either their major or double-major at the university. They were given a mock TOEIC test which only includes 50 questions of reading comprehension part. They were divided into two groups based on their TOEIC scores (sentence writing group vs. composition writing group). The mean scores of TOEIC were 25.26 (SD=7.21) for the sentence writing group and 24.50 (SD=7.49) for the composition group. Participants' knowledge of the target words was indicated 3+ (3, 4, or 5) in their vocabulary test (detailed descriptions about measurement of vocabulary knowledge are provided in section 3.2.3 vocabulary tests). The mean of pre-test in each group was 0.40 (SD=0.63) in the sentence writing group, and 0.50 (SD=0.85) for the composition writing group (Table 2).

Table 2. Mean Scores of TOEIC and Pretest for Each Group

Group (N)	N	TOEIC		Pre-test (3+)	
		M(50)	SD	M(7)	SD
Sentence writing	15	25.26	7.21	0.40	0.63
Composition writing	14	24.50	7.49	0.50	0.85

Table 3. Results of Independent t-test between Two Groups

Tests	N	Sentence Writing vs. Composition Writing		
		t	df	p
TOEIC	29	.281	27	.781
Pre-test (3+)	29	-.360	27	.722

Independent T-tests were conducted to see mean differences in both TOEIC and a pre-test. In the TOEIC test, mean differences between two groups were not significant for the sentence writing group (M=25.26, SD=7.21) and the composition writing group (M=24.50, SD=7.49) conditions; T(27)=.281, p=.781. It indicated that two groups' English proficiencies were almost the same. In the pre-test, the meaning differences between the two groups were also not significant for the sentence writing group (M=0.40, SD=0.63) and the composition group (M=0.50, SD=0.85) conditions; T(27)=-.360, p=.722, indicating that two groups' target vocabulary knowledge before the tests (immediate and delayed test) was also almost the same (Table 3).

3.2 Materials

3.2.1 Tasks

Following Hulstijn and Laufer (2001), two different writing tasks were operationalized: sentence writing task and composition writing task. For the sentence writing task, students were given a worksheet which contains 10 target words with their meanings and they were asked to write an original sentence for each of the 10 target words. The students in a composition writing group were provided a list of 10 target words with their English meaning and they were asked to write a three-paragraph essay about their daily life, using all 10 target words. It was required to use all of the target words in their composition. Hulstijn and Lafuer (2001) hypothesized that the involvement index of both tasks was 3, including a moderate *need*, no *search*, and a strong *evaluation* because the L2 word forms and their meanings were already provided, and new words must be evaluated against suitable collocations in a learner-generated context.

3.2.2 Target vocabulary

A total 10 target words were selected according to the following criteria. First, three undergraduate students who are in the same year (grade) at the university, but do not participate in the experiment, were asked to choose unknown words from a list of 40 words selected from the reading material (reading explorer 3). The current researcher assumed that three students who have similar English proficiency levels with other students who participated in the experiment were good enough for initial target words choice. Based on their selections, 10 target

words which are assumed to be unfamiliar to the participants were selected for the study. Second, the researcher (also the instructor of the course) confirmed that the chosen 10 target words would likely be unknown to the participants in the current study. The target words are as follows: *distort, fiercely, constrain, drawback, replenish, revere, stubborn, disastrous, evade* and *compel*.

3.2.3 Vocabulary tests

To measure participants' vocabulary learning, the current study included two vocabulary tests: an immediate posttest (immediately after the treatment for initial learning) and a delayed posttest (one week later for retention of vocabulary). Paribakht and Wesche's (1993) Vocabulary Knowledge Scale (VKS) was adapted to measure both participants' initial vocabulary learning and retention of new vocabulary knowledge. VKS uses a 5-point scale to measure both perceived and demonstrated knowledge of specific words in written form. Each score indicates the following level of vocabulary knowledge: 1 (not familiar), 2 (familiar but meaning is not known), 3 (correct synonym or translation is given), 4 (word is used with semantic appropriateness in a sentence), and 5 (word is used with grammatical and semantic appropriateness in a sentence). The VKS was regarded as a valid instrument to measure participants' vocabulary knowledge for this study because it was designed to track the early development of specific word knowledge.

3.2.4 Procedure

The experiment was conducted on two separate days over a week period. The pretest, treatment and immediate posttest were carried out by the researcher on the same day, and the delayed posttest was administered a week later. The participants were informed of a general concept, not the details of the study. They were not told that they would take vocabulary test after the task in order to avoid their attention to the target words, which would facilitate incidental vocabulary learning to occur. Each student was given either sentence writing task or composition task. Time on task was approximately 50 minutes for each group: 10 minutes for a pre-test, then 50 minutes for the task, 10 minutes for an immediate posttest, and another 10 minutes for a delayed posttest a week later.

Before the task, they were checked about their knowledge of target vocabulary. They were asked to mark the words they had already known in order to check for their pre-knowledge of the target words. After the completion of the vocabulary task, the worksheets were collected and an immediate posttest was carried out to measure their initial vocabulary learning. One week later, the participants also had the delayed posttest to measure their retention of vocabulary knowledge, which had the same format but different from the immediate posttest in the order of the test items. The researcher made sure that all students did not practice the target words in class between the immediate and delayed posttests, and the data collected from the students who indicated more than three words as previously known in a pretest were eliminated from the analysis.

3.2.5 Analysis

Data were reviewed as a whole in each group (sentence writing and composition writing), and calculated using SPSS statistical analysis program in order to verify the efficiency of the tasks. Dependent variable for each research question was the scores on the immediate and delayed posttests, and independent variable was task types (sentence writing and composition). An independent t-test comparing mean scores of two groups in both immediate posttest and delayed posttest was carried out to see whether or not the significant effects in each group exist.

3.3 Results

   Table 4 shows the descriptive statistics for immediate posttest, and delayed posttest of each group. Overall, the participants showed prior target vocabulary knowledge less than 0.40 in sentence writing and 0.50 in composition writing out of 10. Vocabulary gains became increased after treatment in each group. As table 4 shows, there was little difference between the sentence writing and composition writing in the immediate posttest. For example, the sentence writing group remembered the words of 6.93 and the composition group remembered the target words of 6.92, showing no much differences in the initial word gain. However, the mean differences became larger in the delayed posttest which was conducted a week later. The sentence writing group retained target words only of 3.73 and the composition group retained target words of 6.50 (Figure 1).

Table 4. Descriptive Statistics for Pre, Immediate and Delayed Tests

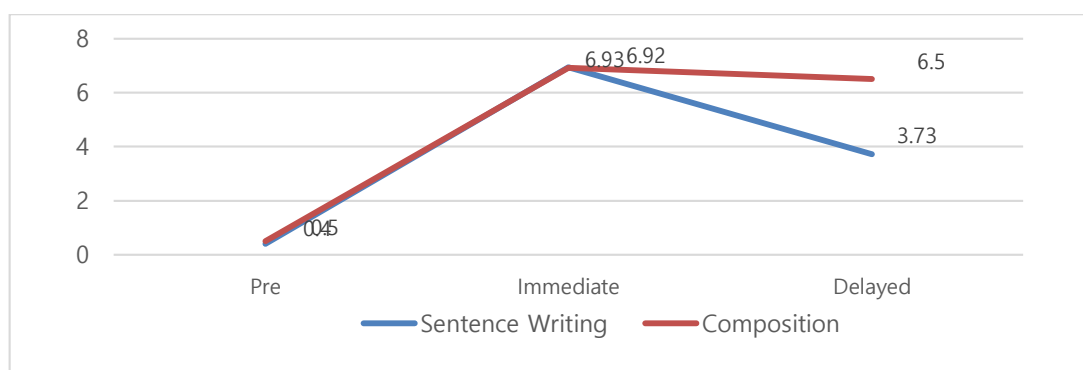| Group (N)	N	Pre-test (3+)		Immediate Test(3+)		Delayed Test (3+)	
		M(10)	SD	M(10)	SD	M(10)	SD
Sentence writing	15	0.40	0.63	6.93	2.52	3.73	2.08
Composition writing	14	0.50	0.85	6.92	2.46	6.50	2.65

Figure 1. A Comparison of Sentence Writing and Composition Writing

Table 5 Test of Equality Variances

Test types	Levene's test for equality variances	
	F	Sig.
Immediate test	.001	.982
Delayed test	2.639	.116

Table 6. Results of Independent t-test for the Immediate-test

Tests	N	Mean	SD	t	df	p
Sentence writing	15	6.93	2.52	.005	27	.996
Composition	14	6.92	2.46			

Table 7. Results of Independent t-test for the Delayed-test

Tests	N	Mean	SD	t	df	p
Sentence writing	15	3.73	2.08	-3.133	27	.004*
Composition	14	6.50	2.65			

Before comparing the means of the two groups, Levene's test was performed to assume the equal variance between the groups. Levene's test was not significant (p=.982 for the immediate test, and p=.116 for the delayed test), assuming that the variance between the groups are the same (Table 5). An independent t-test conducted for the immediate posttest shows that meaning differences between the two groups were not significant for the sentence writing group (M=6.93, SD=2.52) and the composition writing group (M=6.92, SD=2.46) conditions; T(27)=.005, p=.996, indicating that both the sentence writing group and the composition writing group gained similar word learning in their initial vocabulary test (Table 7).

An independent T-test conducted for the delayed posttest shows that the mean differences between the two groups were significant for the sentence writing group (M=3.73, SD=2.08) and the composition group (M=6.50, SD=2.65) conditions; T(27)=-3.133, p=.004 indicating the composition writing group had retained more vocabulary than the sentence writing group (Table 7).

4. Discussion

The research questions were to test the assumption suggested by Hulstijn and Laufer (2001) that the same amount of involvement load in tasks would facilitate the same amount of word learning. The experiment was conducted to examine whether the two tasks of sentence writing and composition writing which are assumed to involve the same amount of load equally contribute to the similar initial vocabulary learning and retention of them for a week. The answer "yes" in the initial word learning (the first research question) but "no" in the retention of word (the second research question). The results partially support ILH showing that only the initial word learning is the same in both tasks whereas the word retention was stronger in the composition writing group rather than the sentence writing group. This represents writing a composition seems to require learners a deeper cognitive processing than writing an original sentence writing for word learning, which is also conducive to long-term memory. A few points regarding the results of this experiment and previous studies on task involvement load can be discussed.

Little research investigating the benefit of writing practice for word learning within the framework of involvement load hypothesis (Hulstijn and Laufer 2001) has been carried out, and there have been inconclusive results on the experiment of writing practice tasks. Kim (2008) found that the sentence writing and the composition writing task were equally effective for both initial word learning and its retention, but on the other hand, Zou's (2017) experiment with Chinese students on two writing tasks (sentence writing and composition writing) showed that the composition writing task was more beneficial to the students than the sentence writing task in both initial word learning and its retention. The results of current study are in line with Kim's study (2008) in that the initial word learning was the same in both tasks, and with Zou's (2017) study in that word retention (long-term learning) was more effective in composition task rather than in sentence writing task. Three studies showing different results on the same tasks are worthy of our attention to get more clear insight on the process of target words within the framework of involvement load hypothesis, and Kim's (2008) argument that the tasks carried out within the self-generated context would induce similar degree of involvement resulting in the same degree of word learning is

also necessary to be reconsidered in different perspectives.

One of the reasons that the composition writing task was more effective on a long-term memory than the sentence writing task is due to the degree of information process. It is commonly believed that the composition writing task is more difficult than the sentence writing task, inducing holistically deeper cognitive process of target words (Kim 2008), and the overall difficulty of a task attributes to the degree of processing target words as well as maintaining the coherence of context in which target words are accordingly used (Zou 2017). Then, how the information-process mechanism during the tasks, not just the degree of process, occurs needs to be discussed in order to provide more tangible explanation on the effectiveness of composition task. In addition to the degree of information-processing for word learning, Zou (2017) brought up the notion of 'chunking' and 'hierarchical organization' as information-processing mechanism. First, the notion of 'chunking,' as a unifying information-processing mechanism, was initially proposed by Chase and Simon (1973) through which the information being processed is grouped into meaningful units and becomes much easier to memorize (Hintzman 1978). Chunking plays an important role increasing leaners' abilities of extracting information from the context as well as determining the effectiveness of information memorization activities (Gobet et al. 2001). Zou (2017) reported the differences of information-processing mechanism in regard to using chunking in each writing task. For instance, in a sentence writing task, when creating a sentence using a target word, sentences are rather spontaneously generated with no relations between the target words or there might not be any 'chunking processes' between target words. The example sentences from the sentence writing task in Figure 2 do not show any chunking process or relations between target words and there is even no specific context grouping 10 target words. In other words, participants doing a sentence level task are only required to create contexts for a separate and independent individual target words rather than contexts that connect a target word with previously generated contexts.

I really <u>revere</u> god 2 years ago.
My mom <u>compel</u> me to study every day.
The famous singer <u>evade</u> her guilty for popularity.
My sister <u>replenish</u> water in empty vase.
Japan <u>distort</u> an important history long time ago.
The children are <u>stubborn</u> when they want something.
After the war, the world is <u>disastrous</u>.
This test's <u>drawback</u> is that I never study.
Tigers change <u>fiercely</u> when I touch their babies.
The teacher <u>constrain</u> us to use phone in class.

Figure 2. An Example of Sentence Writing Written by One Female Participant

In contrast, learners who are required a composition writing in a task need to conceive a rather organized scenario in which every target word needs to be related with at least one of the others to create a story through associated contexts. Writing a story with using the target words in a consistent and structured context would require a learner more complex and deep information-processing mechanism, which also helps learners retain the learned words more effectively. In other words, hierarchical organization which involves various relations between units and sub-units has been found to be an effective mechanism for retrieving memory of target information (Anderson 2010). Participants doing composition writing in our study also made use of hierarchical organization. Overall, they wrote a three-paragraph essay about their daily life. Figure 3 shows the composition written by a male participant, in which systematically developed three chunks that connect all target words were found, and the essay

has been hierarchically constructed through the organization of the three chunks. At the beginning of the writing, a participant found his feelings on the writing assignment which bases the story (chunk 1), then the chunk 2 and chunk 3 are connected to build a story about one of his neighbors (Figure 4). These three chunks which are connected enable him to compose an essay in an organized hierarchy.

In addition to the hierarchical organization of chunks, it was also found that the participants doing composition task tend to repeat the target words in their essay. For instance, in Figure 3, the target words *drawback* and *revere* were repeated twice (*he is nice but he has a lot of <u>drawbacks.</u>*, *this is good point of my neighborhood but also <u>drawback</u>.*, *I can respect him but I can't <u>revere</u> him.*, *no one in my neighbor <u>reveres</u> him.,.*). When creating a story, it would be sometimes necessary to repeat the same words two or three times. But, in the sentence writing task, it is not necessary to repeat a word. The repeating of target words to create hierarchically organized story also contributes to long-term word memory.

Lastly, although there was 'the use' of chunking in the sentence writing group, it might not lead to hierarchical organization of information, which also might be only effective in learning vocabulary for a short-term. However, it is assumed that the composition writing task entails structuring of information on the target words through hierarchical organization within the consistent context, and contributes to retention of learned words for a long-time. Thus, in short-term memory, target words can be maintained at any level as long as they are cognitively processed, but in long-term memory, target words are most likely to be remembered when they are processed at a deep and meaningful way and chunked in an associated context with hierarchical organization. It concludes that simple cognitive stimulation on the target words will not facilitate long-term recall, but, elaboration at deeper levels of cognition will promote long-term recall.

Well, I start this essay because someone <u>compels</u> me to write this one. Actually I wanted to <u>evade</u> this situation but I won't. That's cheating. Alight then, I want to introduce my neighborhood. He is nice but he has a lot of <u>drawbacks</u>. He has <u>disastrous</u> mind. First, he is <u>distorted</u>. Let me give you an example. When I say "You look happy today." to him and he start to think "Do I look happy? I am just walking on the street. But he said I look happy. He is mocking me with something.' After that, he starts to fight against me <u>fiercely</u>. I started to relax me, but it didn't work. He is so <u>stubborn</u> that I couldn't have a deep breath. Second he <u>constrains</u> himself to make it perfect, like, he make a jar full of water. If someone drinks water which he made, then he starts to empty that jar and <u>replenish</u> it once again. He is doing that every day. It looks perfect, but also exhausted, So this is good point of my neighborhood but also <u>drawback</u>. I can respect him but I can't <u>revere</u> him. Well actually no one in my neighbor <u>reveres</u> him.

Figure 3. An Example of Composition Writing Written by One Male Participant
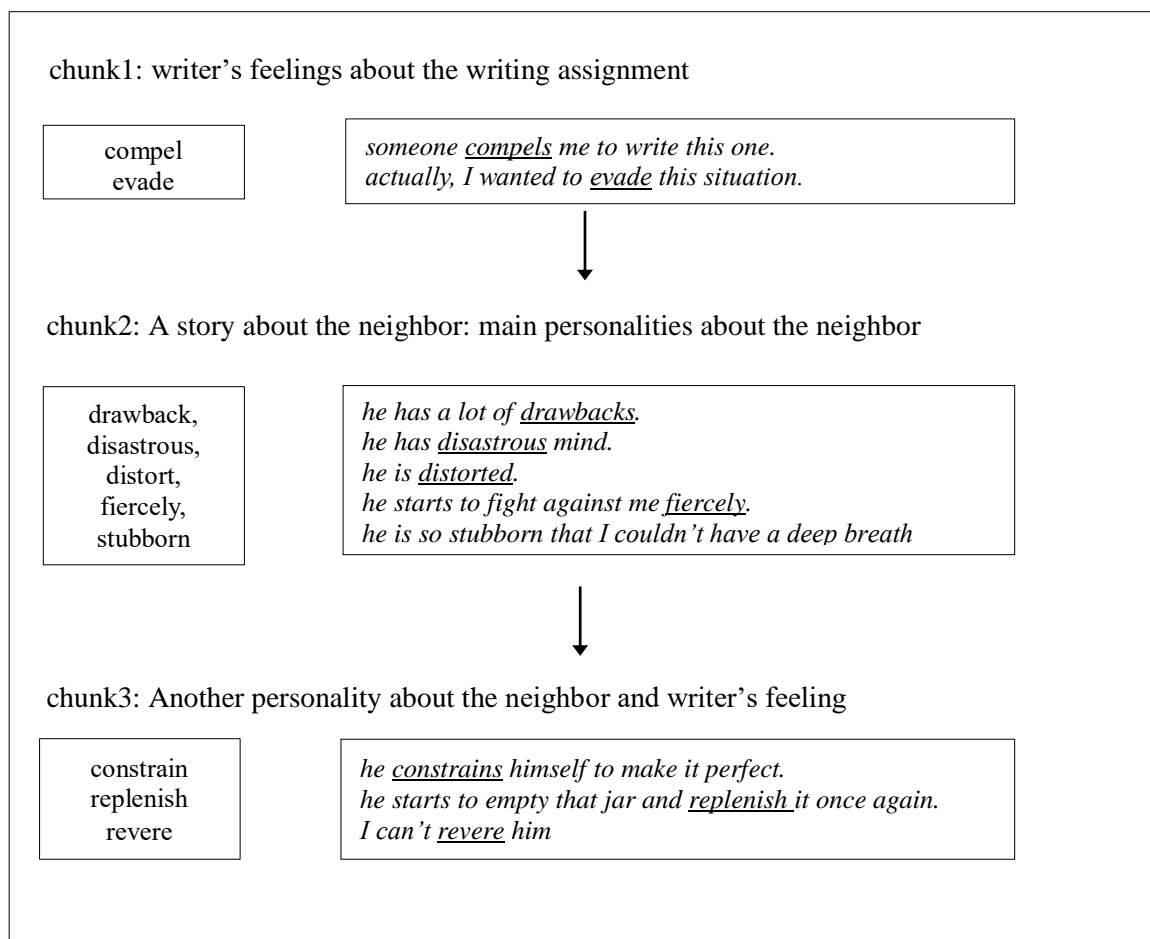
chunk1: writer's feelings about the writing assignment

| compel
evade | *someone <u>compels</u> me to write this one.*
actually, I wanted to <u>evade</u> this situation. |

chunk2: A story about the neighbor: main personalities about the neighbor

| drawback,
disastrous,
distort,
fiercely,
stubborn | *he has a lot of <u>drawbacks</u>.*
he has <u>disastrous</u> mind.
he is <u>distorted</u>.
he starts to fight against me <u>fiercely</u>.
he is so stubborn that I couldn't have a deep breath |

chunk3: Another personality about the neighbor and writer's feeling

| constrain
replenish
revere | *he <u>constrains</u> himself to make it perfect.*
he starts to empty that jar and <u>replenish</u> it once again.
I can't <u>revere</u> him |

Figure 4. Hierarchical organization of composition writing

5. Conclusion and Limitations

The main purpose of this experiment was to examine whether the same degree of involvement load (sentence writing vs. composition writing) results in a similar degree of vocabulary learning by Korean EFL learners. The experiment showed that the same degree of involvement load in two different tasks did similarly contribute to a short-term memory, but not to a long-term memory. This result was in contrast with the previous study (Kim 2008), in which the same theoretical level of task-induced involvement load has similar effects both on the initial word learning and retention of target words. But, this study implies that despite the same degree of involvement load, the composition writing requires a learner to be involved in a deeper word processing level than the sentence writing task, contributing to more consistent chunking process for unknown words, and more hierarchical organization of information, which also results in more effective retrieving memory for target words. The results partially supported the ILH, suggesting that the differential effects of the writing practice on a long-term memory should receive an attention from researchers and educators in order to design more effective task materials. It is also suggested that the involvement degree of the task should be refined in relation to the framework of learners'

memory capacity (short-term vs. long-term).

The results of the current study along with the findings from previous studies (Hulstijn and Laufer 2001, Keating 2008, Kim 2008, Park et al 2019, Sung 2013) provide English educators with some implications regarding vocabulary teaching. It is suggested that English teachers create reading tasks for vocabulary learning which induce learners' cognitive *evaluation* rather than motivational *need* since learners' effort for *evaluating target words* in context is much more effective for vocabulary learning than the *need to know* the meaning of the target words. Particularly, creating a task requiring learners' deeper levels of processing load (e.g. composition writing) would be a benefit to those who learning English as a foreign language (EFL) because learners in those learning environments would not have many chances to encounter English vocabulary outside the classroom. Future studies examining the extent to which individual components contribute to both a short-term memory and a long-term memory will help define weighting of the three components (*need*, *search*, *evaluation*) more precisely.

References

Anderson, R. 2010. *Cognitive Psychology and its Implications*. 7th edition. New York: Worth.

Chase, G. and A. Simon. 1973. Perception in chess. *Cognitive Psychology 4*, 55-81.

Gobet, F., R. Lane and S. Croker. 2001. Chunking mechanisms in human learning. *Trends in Cognitive Science 5*, 236-243.

Hamada, M. and C. Park. 2011. Word-meaning inference: A longitudinal investigation of inference accuracy and strategy use. *Asian EFL Journal 13*(4), 10-32.

Hamada, M. and C. Park. 2013. The role of think-aloud and metacognitive strategies in L2 meaning-inference during reading. *JALT Journal 35*(1), 101-125.

Hintzman, L. 1978. *The Psychology of Learning and Memory*. San Francisco, CA: Freeman.

Hu, H.-C. and H. Nassaji. 2014. Lexical inferencing strategies: The case of successful versus less successful inferencers. *System 45*, 27-38.

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

Kang, H. and Y. Shin. 2019. The involvement load hypothesis and vocabulary retention: The interplay of task types and involvement index on L2 vocabulary learning. *Foreign Language Education 26*(4), 89-111.

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

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

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

Laufer, B. and J. Hulstijn. 2001. Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics 22*(1), 1-26.

Lee, Y.-K. and J. Y.. Kim. 2015. Effects of task-induced involvement on EFL learners' vocabulary learning. *The Journal of Linguistic Science 72*, 297-318.

Nation, I. S. P. 2000. Learning vocabulary in lexical sets: Dangers and guidelines. *TESL Journal 9*(2), 6-10.

Nation, I. S. P. 2001. *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.

Paribakht, S. and M. Wesche. 1993. Reading comprehension and second language development in a

comprehension-based ESL program. *TESL Canada Journal 11*(1), 9-29.

Park, C. 2019. Effects of phonological decoding process on L2 vocabulary learning among Korean EFL learners. *The Jungang Journal of English Language and Literature 61*(4), 263-283.

Park, C. 2020a. Vocabulary learning: two dimensions of involvement load and learners' proficiency level. *Studies in Linguistics 56*, 217-237.

Park, C. 2020b. Effects of phonological decoding process on L2 vocabulary learning among Korean EFL learners. *The Jungang Journal of English Language and Literature 61*(4), 263-283.

Park, C., S. K. Yun and Y. Lee. 2019. Task involvement load and its effectiveness: Motivational vs. cognitive dimension. *Studies in Linguistics 52*, 305-322.

Park, H. K. 2017. Effects of task-induced involvement on EFL adult learners' vocabulary learning. *English 21* (2), 203-226.

Pulido, D. 2007. The relationship between text comprehension and second language incidental vocabulary acquisition: A matter of topic familiarity? *Language Learning 57*, 155-199.

Pulido, D. 2009. Vocabulary processing and acquisition through reading: Evidence for the rich getting richer. In Z.-H. Han and N. Anderson (Eds.), *Second Language Reading Research and Instruction: Crossing the Boundaries* (pp. 65-82). Mahwah, NJ: Lawrence Erlbaum.

Schmidt, R. 1994. Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. Ellis (Ed.), *Implicit and Explicit Learning of Languages* (pp. 165-209). London: Academic Press.

Schmidt. N. 2000. *Vocabulary in Language Teaching.* Cambridge: Cambridge University Press.

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

Sung, H. 2019. A Study on vocabulary learning through gap-filling tasks by low-proficiency learners. *Journal of the Korean English Education Society 18*(4), 1-25.

Teng, M. F. and D. Zhang. 2023. Vocabulary learning in a foreign language: Multimedia input, sentence-writing task, and their combination. *Applied Linguistics Review*, published online.

Zou, D. 2017. Vocabulary acquisition through close exercises, sentence-writing and composition-writing: Extending the evaluation component of the involvement load hypothesis. *Language Teaching Research 21*(1), 54-75.

Examples in: English
Applicable Languages: English
Applicable Level: Tertiary