Korean Journal of English Language and Linguistics, Vol 22, November 2022, pp. 1236-1252
DOI: 10.15738/kjell.22..202211.1236

KOREAN JOURNAL OF ENGLISH LANGUAGE AND LINGUISTICS

ISSN: 1598-1398 / e-ISSN 2586-7474

http://journal.kasell.or.kr

Phonological Decoding for L2 Vocabulary Learning and Its relations to Learners' Proficiency Levels*

Chaehee Park (Sun Moon University)

Received: September 09, 2022
Revised: November 11, 2022
Accepted: November 30, 2022

Chaehee Park
Professor, Division of Foreign Languages. Sun Moon University
Tel:041-530-2423
chpark@sunmoon.ac.kr

* Data were presented at the conference of the Jungang English language and literature association of Korea in 2022.

ABSTRACT

Park, Chaehee. 2022. Phonological decoding for L2 vocabulary learning and its relations to learners' proficiency levels. *Korean Journal of English Language and Linguistics* 22, 1236-1252.

This study tests whether or not phonological decoding process contributes to vocabulary learning in the framework of task involvement load, and how learners' proficiency levels are related to the effectiveness of the task involving phonological decoding process. Three tasks with different amount of load were created and were given to 62 Korean EFL university students (21 for reading comprehension task, 22 for reading comprehension plus phonological decoding process, and 19 for reading comprehension plus gap-filling task). The participants' performance was tested three times (pre-test, immediate post-test, and delayed post-test). First, it was found that there were significant differences only between the reading group and the fill-in group in the immediate test, but no differences among the three groups in the delayed-test although the phonology group retained initially-learned words best. Second, the results also revealed that there were significant differences between the phonology group and the reading group among the low-level students, indicating that decoding process was conducive to both initial word learning and successful retention of vocabulary knowledge particularly among the low-level learners. It is suggested that phonological decoding process needs to be considered as one of the components of involvement loads for low-level students, and that English instructors create various task materials both entailing students' phonological process for the low-levels and inducing much amount of involvement load for the high-levels.

KEYWORDS

vocabulary learning task, involvement load hypothesis, phonological decoding, second language learning, Korean learners of English

1. Introduction

Creating an effective task material for second language vocabulary learning has been of major interest among L2 language researchers and teachers since vocabulary knowledge is essential in all aspects of second language learning. In the past decade, a great deal of research on word learning and acquisition has been conducted and filed documents on the learning task for better vocabulary learning. One of the growing interests in this area is to develop task materials for vocabulary learning. In regard to vocabulary task, Hulstijn and Laufer (2001) proposed involvement load hypothesis (ILH), which is fundamentally affected by the notion of processing hypothesis. It has been well discussed that the amount of task load is related to learners' levels of processing of the unknown words. They had also discussed three components which could be inherent to the task: *need, search,* and *evaluation.* A number of research has confirmed that the task with much amount of load of these components contributes to a deeper level of processing, leading to more effective vocabulary learning, than the task with less amount of load of the components (Hulstijn and Laufer 2001, Keating 2008, Kim 2011, Kim and Na 2010, Park 2017). However, this meaning-focused task in the framework of ILH did not consider the form of word which is also related to word learning. One thing about the form of word involves the phonology of the word, and the relation between phonology and word learning has been widely discussed in previous studies as well.

In a psycholinguistic perspective, it has been found that phonological decoding process of unknown words contributes to literacy development including vocabulary learning; Bowey (2001) and Metsala (1999) for vocabulary learning, and Hamada and Koda (2010) and Hamada and Park (2011) for word-meaning inferences. Hamada and Koda (2010) reported a correlation between the efficiency of phonological-decoding of real words and word-meaning inference among the Korean L1 students. Regarding L2 word-meaning inference performance, learners were also identified to use strategies such as word-form analogy and morphological analysis which entail phonological process of target words (Hamada 2009). A better L2 word learning improvement was also found when words are presented or clustered in the way of phonological similarity rather than semantically linked (Wilcox and Medina 2013).

Given the effectiveness of both involvement load and phonological decoding process on word learning, it would be worth investigating these two in one experiment since both in either case (form vs. meaning) contribute to learners' vocabulary learning. There is little or no prior research directly testing the effectiveness of L2 learners' phonological activities in relation to the task involvement load and learners' proficiency levels. In this experiment, it is investigated whether or not the phonological decoding process in a task contributes to vocabulary learning and whether Korean EFL learners' proficiency levels and the effectiveness of the task were related or not. On the basis of the results, it would be also discussed whether there is a possibility that phonological decoding process or its awareness could be considered as one of the components of task involvement load.

2. Previous Studies

2.1. Involvement Load Hypothesis

Over the past two decades researchers have documented a large number of studies theorizing and exploring what effective vocabulary-learning task materials are composed of. Laufer and Hulstijn's (2001) study, one of the pioneering works in the vocabulary-learning task material research within the framework of task involvement load, suggested that the construct of task involvement is composed of motivational and cognitive dimensions: *need,*

search and *evaluation* for the target words. These three components, taken together, constitute a construct of involvement of the task, and a few assumptions regarding ILH were suggested in terms of the depth of processing and elaborating learning vocabulary. One of the assumptions suggested by Laufer and Hulstijn (2001) as below.

> **Assumption one:** Retention of words when processed incidentally, is conditional upon the following factors in a task: need, search, and evaluation. (Laufer and Hulstijn 2001: 14)

Three components suggested in assumption one can be entailed in a task, contributing to word learning. Each of these components is suggested to vary in the degree of strength. Surely, we, educators are aware that the highly motivated learners would be more deeply engaged in learning activity than the learners who are less motivated, which also leads to higher chance of success in learning. For this reason, teachers in a language classroom attempt to increase students' intrinsic and external motivation by creating task materials which facilitate their *need* for word learning. The *need* component in the first assumption is based on a drive to comply with the task requirements. It is motivational, non-cognitive component of involvement and the degree of the motivational *need* can be either *moderate* or *strong* depending on how it is imposed. For instance, the *need* component can be either *moderate* (1) when it is imposed by an external agent (e.g., the task requires students to fill in a word in sentence) or *strong* (2) when imposed by a learner him or herself (e.g., a learner wishes to look up a word in an L1-L2 dictionary to accomplish the task). *Moderate* and *strong* in need drive different degree of motivation.

Along with *need* component, Laufer and Hulstijn (2001) also identified two other components which are conducive to elaborate cognitive processing of information for vocabulary learning: *search* and *evaluation*. These two components are more related to learners' cognitive activity involving learners' noticing new words and allocating attention to the form and meaning of new words. *Search* refers to the attempt learners try to make in order to find the meaning of an unknown L2 word in a task. This component 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 can be glossed). *Evaluation* component requires a "comparison of a given word with other words, the 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: 14). The degree of *evaluation* is *moderate* (1) when it entails recognizing differences between words (e.g., deciding which meaning of a target word best fits the context), and *strong* (2) when the task requires learners' making a decision on new words and combining them with known words in original contexts (e.g., sentence writing and composition).

> **Assumption two:** Other factors being equal, teacher/researcher-designed tasks with a higher involvement load will be more effective for vocabulary retention than tasks with a lower involvement load. (Laufer and Hulstijn 2001: 15)

Second assumption was suggested in regard to the combination of the presence and absence of the involvement components, and the amount of load and its relation to word retention. In a classroom learning environment, a vocabulary task can be designed to induce much involvement load of *need*, *search* and *evaluation* for a better vocabulary learning. Table 1, excerpted from Laufer and Hulstijn (2001) illustrates some examples of task-induced involvement load in terms of the degree of each component. A task created focused on the bottom of the table would be more effectively used for vocabulary learning than the top of the table since it entails much more loads which help learners process unknown words deeply.

Table 1. Task-induced Involvement Load

	Task	Status of target words	*N*	*S*	*E*
1	Reading and comprehension questions	Glossed in the text but irrelevant to task	-	-	-
2	Reading and comprehension questions	Glossed in the text and relevant to task	+	-	-
3	Reading and comprehension questions	Not glossed but relevant to task	+	+	+/-
4	Reading and comprehension questions and filling gaps	Relevant to reading comprehension. Listed with glosses at the end of text	+	-	+
5	Writing original sentences	Listed with glosses	+	-	++
6	Writing a composition	Concept selected by the teacher (and provided in L1). L2 learner writer must look up the L2 form	+	+	++
7	Writing a composition	Concepts selected and looked up by L2 learner-writer	++	+	++

N = need; S = search; E = Evaluation; - = absent; + = moderate; ++ = strong

A great deal of research has tested this assumption and confirmed that a higher involvement load leads to deeper processing of the target words than a lower involvement load, resulting in better vocabulary learning (Hulstijn and Laufer 2001, Keating 2008, Kim 2011, Kim and Na 2010, Lee and Kim 2015). One of the initial studies, Hulstjin and Laufer (2001) investigated the effects of involvement load on the retention of 10 English unknown words by two groups of EFL learners: Israeli EFL and Dutch EFL learners. To this end, they designed an experimental study with three vocabulary learning tasks involving different degrees of involvement load; reading only (involvement load 1), reading plus fill-in (involvement load 2), and composition writing (involvement load 3). The results indicated that word retention was related to the amount of task-induced involvement load: it was argued that the composition task yielded the highest because it involved a higher involvement load than the other two tasks. However, the reading plus fill-in was more effective only in the Israeli group not in the Dutch group, indicating partial support for ILH.

Experiments testing ILH with Korean EFL learners were conducted (Lee and Kim 2015, Park 2017), reporting somewhat different results in terms of learners' proficiency levels (Kim and Na 2010, Sung 2013, 2019) and the effectiveness of individual components (Park, Yun and Lee 2019). More specifically, Sung (2013) conducted an experiment with two hundred three Korean EFL high school students on the involvement load and its relation to learners' proficiency levels. Two post-tests results on the three different loads of tasks (reading comprehension of involvement load 1, reading and gap-filling of involvement load 2, and unscrambling sentences of involvement load 3) revealed that the low-level learners did not benefit much better than the higher-level learners in promoting short-term learning and long-term retention, suggesting proficiency level is a significant factor both in the learners' short-term and long-term retention. On the contrary, the low-proficiency learners benefitted much better from the gap-filling task (involvement load 2) than the writing task (involvement load 3) in later word retention (Sung 2019).

Given these mixed results on the task-induced involvement load and proficiency levels, it was worthy of investigating whether or not the degrees of involvement load (*absent*-0, *moderate*-1 and *strong*-2) for each component such as *need*, *search*, and *evaluation* could be equally treated. In other words, when the total involvement loads are the same in two tasks, learners would or would not get similar word gains from the two tasks. To this end, Park et al. (2019) took both two dimensions (motivation and cognition) into account in their experiment and designed two different tasks which are the same in the total involvement load of three, but different in the degree of each component. The results indicated that even in the same amount of task-induced load, the

cognition-focused task contributed to both short-term vocabulary learning and long-term retention more than the motivation-focused task, and even to the low-level English learners (Park 2020). However, despite a little different results of studies, it has been confirmed that ILH works for L2 vocabulary learning in general.

2.2. Phonology and Word Learning

An effective task for vocabulary learning is not limited to the framework of IL. Why the task material should be just focused on the meaning of word in a context, ignoring the form and/or phonology of a word. With regard to the vocabulary learning, Ellis (1994) claims that the perceptual aspects of new words, i.e., acquiring their phonetic and phonological features, are learned implicitly as a result of frequent exposure. In a psycholinguistic perspective, regarding the relation of phonological awareness or phonological activity to vocabulary learning, it has been found that phonological process of words or the presentation of words in a form of phonological similarity is one of the fascinating factors or skills enhancing vocabulary learning (Bowey 2001, Hamada and Koda 2010, Hamada and Park 2011, Metsala 1999, Wilcox and Medina 2013). Metsala (1999) and Bowey (2001) found that both measures of phonological awareness and non-word repetition predicted vocabulary development among young learners. Research in a cognitive psychology also found that containing repetitive aspects (such as certain consonant-vowel combinations) of words in verbal information facilitates word information retrieval. For instance, Wilcox and Medina (2013) investigated the semantic and phonological clustering effects among L2 Spanish learners. The results showed that the semantically clustered words were more difficult to learn than the words presented in the form of phonological similarity.

A critical role of phonology of a word was also found in word-meaning inference. Word-learning through lexical inference strategies also entails learners' phonological activity for successful inference (Hamada and Koda 2010, Hamada and Park 2011, Hu and Nassaji 2014). For instance, Hamada and Koda (2010) conducted an experiment with both Korean ESL students and Chinese ESL students. They examined the relationship between phonological decoding and word-meaning inference, and found the efficiency only from Korean ESL students whose L1 is alphabetic, but not Chinse ESL students. This implies that the experience of phonological process in L1 would also contribute to L2 word phonology process and its meaning inference.

Given the association of phonological processing and its vocabulary learning and that phonology of a word plays an important role for L2 learners to acquire target vocabulary, it was worth investigating whether or not phonological decoding process in the framework of involvement load could benefit ESL students. As discussed, so far the components of IL were categorized into *need*, *search*, and *evaluation* focused only on the meaning of target word in the reading context, but learners at least have to see the form of the target word in order to infer the meaning of the word or understand the context. However, ILH did not consider how much the learners' recognition of form contributes to the word learning. Park (2019) found in his phonological decoding task which considered as a form recognition activity, it was most beneficial in retaining learned vocabulary, suggesting that phonological decoding should be included as one of the components for ILH. However, the number of subjects participated in the experiment was relatively small and the participants were homogeneous in their proficiency levels. Whether learners' proficiency levels and phonological decoding are related to vocabulary learning is worthy of our attention for the creation of more effective task materials. The current experiment reexamines Park's (2019) study with more number of participants and various learners' proficiency levels, and extends its results to document the effectiveness of phonological decoding in the framework of ILH. Research questions are as follows.

Research Question 1

Does the phonological decoding process in a task involvement load affect the initial and retention of vocabulary learning among Korean EFL students?

Research Question 2

If so, does the effect of the phonological decoding on vocabulary learning vary depending on learners' proficiency levels?

3. Research Method

3.1 Participants

The total participants were 62 Korean EFL students who are currently majoring in English in an undergraduate degree program. Participants were divided into two groups based on the mock TOEIC test (50 questions of reading comprehension): 33 students of low-level group and 29 students of high-level group[1]. The mean TOEIC scores of the low-level group were 20.00 (*SD* = 3.17), and 33.24 (*SD* = 4.57) for the high-level group respectively. The low-level students were randomly divided into three groups (11 of reading group, 13 of phonology group and 9 of fill-in group), and the mean scores of each low-level group were 20.39 (*SD* = 2.93), 19.75 (*SD* = 2.83), and 19.88 (*SD* = 4.16) respectively. One-way ANOVA showed the mean differences between the three low-level groups were not significant (p = .931), indicating that the participants of these groups can be considered to have similar English proficiency levels. Each group was given tasks: reading, phonology, and fill-in task respectively. The high-level students were also randomly divided into three groups (10 of reading group, 9 of phonology group and 10 of fill-in group), and the mean scores of each high-level group were 33.02 (*SD* = 5.33), 33.00 (*SD* = 4.58), and 33.70 (*SD* = 4.19) respectively. One-way ANOVA showed the mean differences between these groups were not significant (p = .916), indicating that the participants of these three high-level groups can be considered to have the same English proficiency levels. Each group was also given tasks: reading, phonology, and fill-in task respectively.

3.2 Task Materials and Target Vocabulary

Based on the previous studies (Hulstijn and Laufer 2001, Keating 2008, Kim 2001), three task materials with the different amount of involvement loads were created: reading comprehension task, phonology task, and fill-in task). For the reading material, the original reading text "Digital Heaven" was adapted from the English learning website. The length and complexity of the original text were adjusted to students' proficiency levels. Three groups in each level were assigned to tasks with different involvement load. For the reading group (hereafter, reading), the reading text was provided with marginal glosses. Students in this group were required to read the text and answer the accompanying questions. Participants needed to know the meanings of target words in order to answer the questions correctly (moderate need +, no search -, and no evaluation -, total involvement load 1).

For the second group, the same task with the reading group was given with one more activity. In addition to the

[1] Proficiency levels in this study were used only for a comparison purpose between two groups. TOEIC scores do not represent participants' real linguistic performance.

reading comprehension questions, participants in this group were required to count the number of syllable of each target word (hereafter, phonology: moderate need +, no search -, and phonological decoding &, involvement load 1+&). Before the task, this group had received an instruction on how to count the number of syllable of a word, and they practiced counting with nontarget words. This instruction was given two times for a week-period before the experiment (one before a week and second on the experiment date). It was assumed that phonological decoding process, extracting a word's phonological information, occurs while counting the number of syllable of each word (phonological decoding process).

Regarding the fill-in group, students were required to read the text and answer the accompanying questions as in the reading and phonology group, but the target words were deleted from the texts and they were asked to select the appropriate one from a list of glosses as they read the text (hereafter, fill-in: moderate need +, no search -, and moderate evaluation +, involvement load 2).

3.3 Target Vocabulary

Target words used for learning were *protrude, diviner, resurgence, decay, perennial, insurmountable, descendant, halt, hindrance, and envisage.* Paribakht and Wesche's (1993) Vocabulary Knowledge Scale (VKS) was adapted to measure participants' vocabulary knowledge: pre (pretest), initial (immediate test), and the retention of words (delayed test) because this measurement is appropriate to track the early development of specific word knowledge (Read 2000, Wesche and Paribakht 1996). The level of vocabulary knowledge was measured in five scales below.

1. I've never seen this word before.
2. I've seen this word before, but I don't know what it means.
3. I've seen this word before, and I think it means _____.
4. I know this word and it means _____
5. I know this word and I can use this word in a sentence.
_____.

For a word to be counted as learning in the VKS, the participants were required to either provide a correct English synonym or a Korean translation (score 3 or 4) or write an English sentence using target word (score 5). Thus, a score of 3, 4, or 5 was considered as participants' knowledge for analysis with no differences between those numbers.

3.4. Procedures

The experiment was conducted on three separate days over two week period during the class time. First, all students of the phonology group had received an instruction about English word syllable since it was assumed that students might not be able to count the number of English word syllable. They became aware of difficulty identifying syllable in English words compared to words in Korean. The purpose of this instruction was to make the participants experience phonology-decoding process. The other two groups received an instruction about reading comprehension task and the fill-in the task. On the date of reading task, each group had 10 minutes for pre-vocabulary test, then 40 minutes for treatments (reading, phonology, and fill-in), and 10 minutes for an immediate-vocabulary test. One week later, each group had another 10 minutes for the delayed-vocabulary test.

3.5. Data Analysis

For the first research question, data were reviewed as a whole in each group (reading, phonology, fill-in), and calculated using SPSS statistical analysis program in order to verify the efficiency of the tasks. Then, for the second research question, data were divided into two levels based on participants' TOEIC scores in each group and analyzed using SPSS program. One-way ANOVA was conducted to see whether or not the effects in each group exist. The dependent variable for each research question was the scores on the immediate and delayed-tests and the independent variable was task types (reading, phonology, fill-in). Post-hoc comparisons were also conducted to locate the significant differences among the variables when the significant main effects were found in ANOVA.

4. Results

Table 1 provides descriptive analysis of TOEIC, pretest, immediate test, and post test of each group. One-way ANOVA showed the mean differences of TOEIC scores of the three groups were not significantly different (p = .912), indicating that the participants of the three groups can be considered to have the same English proficiency levels. It also provides learners' vocabulary knowledge in each test (pre, immediate, and post). Overall, the participants showed target vocabulary knowledge of less than 0.47 out of 10 in each task (pretest), and their vocabulary gains became increased after treatment in each group. Overall, it showed that phonology group (m = 4.18) performed much better than reading group (m = 3.33) but, not better than the fill-in group (m = 4.94) in the immediate. However, the phonology group retained the vocabulary (m =2.77) more than the reading group (m = 2.09) and the fill-in group (m = 2.10).

Table 1. Descriptive Analysis of TOEIC, Pre-test, Immediate-test, and Delayed-test

Groups	*N*	TOEIC		Pre		Immediate		Delayed	
		M(50)	*SD*	*M*(50)	*SD*	*M*(50)	*SD*	*M*(50)	*SD*
Reading	21	19.39	2.93	.47	.60	3.33	2.19	2.09	2.23
Phonology	22	18.75	2.83	.45	.96	4.18	1.5	2.77	1.82
Fill-in	19	18.88	4.16	.42	.69	4.94	2.14	2.10	1.66
Total	62	19.00	3.17	.45	.76	4.12	2.02	2.37	1.92

Table 2. Effects of Groups on Vocabulary Learning

Tests	Groups	*SS*	*df*	*MS*	*F*	*p*
Immediate Test	Between Groups	26.081	2	13.040	3.421	.039
	Within Groups	224.887	59	3.812		
	Total	250.968	61			
Delayed Test	Between Groups	24.256	2	3.212	.864	.427
	Within Groups	69.622	59	3.720		
	Total	93.879	61			

Table 3. Post-hoc Test Results on the Immediate-test

Group	Inter-dependent variable	Mean difference	*SE*	*Sig.*
Reading	Phonology	-.848	.595	.160
	Fill-in	-1.614	.618	.011
Phonology	Reading	-	-	-
	Fill-in	- 765	611	216
Fill-in	Reading	-	-	-
	Phonology	-	-	-

Table 4. Descriptive Statistics of TOEIC Scores in terms of Proficiency Levels

Proficiency Levels	Groups	*N*	TOEIC	
			M(50)	*SD*
Low	Reading	11	20.39	2.93
	Phonology	13	19.75	2.83
	Fill-in	9	19.88	4.16
	Total	33	20.00	3.17
High	Reading	10	33.02	5.33
	Phonology	9	33.00	4.58
	Fill-in	10	33.70	4.19
	Total	29	33.24	4.57

One-way ANOVA was conducted and the effects of group on the scores revealed that there was significant difference in the immediate-test [$F(2, 59) = 3.421$, $p = .039$] but not in the delayed-test [$F(2, 59) = .864$, $p = .427$] (Table 2). Post hoc comparison showed significant mean differences between the reading task and the fill-in task: [$d = -1.614$, $p = .011$], but not in both the phonology task and the fill-in task [$d = -.765$, $p = .216$], and the phonology task and the reading task [$d = .848$, $p = .160$] (Table 3).

For the second research question, each task group was reclassified into two levels based on participants' TOEIC scores (low vs. high), and data were reexamined in terms of task types and proficiency levels. Table 4 provides the mean scores of each low-level group were 20.39 ($SD = 2.93$), 19.75 ($SD = 2.83$), and 19.88 ($SD = 4.16$) respectively. One-way ANOVA showed the mean differences between the three low-level groups were not significantly different ($p = .931$), indicating that the participants of the three low-level groups can be considered to have similar English proficiency levels. Each group was given the reading task, the phonology task, and the fill-in task respectively.

The high-level students were also divided into three groups based on their TOEIC scores, and the mean scores of each high-level group were 33.02 ($SD = 5.33$), 33.00 ($SD = 4.58$), and 33.70 ($SD = 4.19$) respectively. One-way ANOVA showed the mean differences between the three low-level groups were not significantly different ($p = .916$), indicating that the participants of the three high-level groups can also be considered to have similar English proficiency levels. Each group was given the reading task, the phonology task, and the fill-in task respectively.

Table 5 represents learners' vocabulary knowledge in each test. Overall, the participants showed target vocabulary knowledge of less than 0.66 out of 10 in each task (pretest), and their vocabulary gains became increased after treatment in each group as well. First, in the low-level groups, it showed that the phonology group performed much better than the reading group and the fill-in group in the immediate and delayed test. More

specifically, the phonology group showed the best performance in the immediate test (m = 4.3) and this highest score was being kept in the delayed-test, indicating the phonology group retained the word most (m = 2.61) and followed by the fill-in group (m = 1.33).

Table 5. Descriptive Analysis of Pre-test, Immediate-test and Delayed-test

Proficiency Levels	Groups	*N*	Pre-test (3+)		Immediate (3+)		Delayed (3+)	
			M(10)	*SD*	*M*(10)	*SD*	*M*(10)	*SD*
Low	Reading	11	0.45	0.52	2.27	1.10	0.63	0.92
	Phonology	13	0.30	0.48	4.30	1.31	2.61	1.85
	Fill-in	9	0.33	0.70	3.77	1.30	1.33	1.58
High	Reading	10	0.50	0.70	3.70	1.05	1.90	1.10
	Phonology	9	0.66	1.41	5.32	1.39	2.33	1.00
	Fill-in	10	0.50	0.70	5.70	1.82	4.2	1.87

Second, in the high-level groups, it showed that both the fill-in group and the phonology group performed much better than the reading group, and there is no much difference between the phonology group (m = 5.32) and the fill-in group (m = 5.7) in their short-term memory (mean differences 0.38). However, this difference becomes larger in the retention of vocabulary (delayed test). The fill-in group (m = 4.2) tends to retain much more vocabulary than the phonology group (m = 2.33) and the reading group (m = 1.9).

Table 6. Effects of Groups on Vocabulary Learning in the Low-level

Tests	Groups	*SS*	*df*	*MS*	*F*	*p*
Immediate Test	Between Groups	25.736	2	12.868	8.301	.001
	Within Groups	46.507	30	1.550		
	Total	77.242	32			
Delayed Test	Between Groups	24.256	2	12.128	5.226	.011
	Within Groups	69.622	30	2.321		
	Total	93.879	32			

One-way ANOVA was conducted for the immediate-test and the delayed-test of the low-level group (Table 6). The effects of group on the scores revealed that there were significant differences in both the immediate-test [$F(2, 30) = 8.301$, $p = .001$] and the delayed-test [$F(2, 30) = 5.226$, $p = .011$]. In order to examine the significant mean differences among the low-level groups, post hoc comparison was conducted for the immediate-test (Table 7). There were significant mean differences between the reading task and the phonology task: [$d = -2.034$, $p = .001$], and the reading task and the fill-in task: [$d = -1.505$, $p = .035$], But it did not show any significant mean differences between the phonology task and the fill-in task [$d = .529$, $p = 1.00$]. Post hoc comparison was also conducted for the delayed-test (Table 8). There were significant mean differences between the reading task and the phonology task: [$d = -1.979$, $p = .01$]. But it did not show any significant mean differences between the reading task and the fill-in task [$d = -.696$, $p = .951$], and between the phonology task and the fill-in task: [$d = 1.282$, $p = .185$]

Table 7. Post-hoc Test Results on the Immediate-test in the Low-level

Group	Inter-dependent variable	Mean difference	*SE*	*Sig.*
Reading	Phonology	-2.034	.510	.001
	Fill-in	-1.505	.559	.035
Phonology	Reading	-	-	-
	Fill-in	.529	.539	1.00
Fill-in	Reading	-	-	-
	Phonology	-	-	-

Table 8. Post-hoc Test Results on the Delayed-test in the Low-level

Group	Inter-dependent variable	Mean difference	*SE*	*Sig.*
Reading	Phonology	-1.979	.624	.010
	Fill-in	-.696	.684	.951
Phonology	Reading	-	-	-
	Fill-in	1.282	.660	.185
Fill-in	Reading	-	-	-
	Phonology	-	-	-

One-way ANOVA was conducted for the immediate-test and the delayed-test of the high-level group as well. The effects of group on the scores revealed that there were significant differences in both the immediate-test [$F(2, 26) = 6.344$, $p = .006$] and the delayed-test [$F(2, 26) = 7.630$, $p = .002$] (Table 9). Post hoc comparison was conducted for the immediate-test (Table 10). There were significant mean differences between the reading task and the phonology task: [$d = -1.622$, $p = .016$], and between the reading task and the fill-in task: [$d = -2.000$, $p = .015$], But it did not show any significant mean differences between the phonology task and the fill-in task [$d = .388$, $p = 1.00$]. Post hoc comparison was conducted for the delayed-test (Table 11). There were significant mean differences between the reading task and the fill-in task: [$d = -2.30$, $p = .003$], But it did not show any significant mean differences between the reading task and the phonology task [$d = -.4333$, $p = 1.00$], and between the phonology task and the fill-in task: [$d = -1.866$, $p = .222$].

Table 9. Effects of Groups on Vocabulary Learning in the High-level

Tests	Groups	*SS*	*df*	*MS*	*F*	*p*
Immediate Test	Between Groups	27.210	2	13.605	6.344	.006
	Within Groups	55.756	26	2.144		
	Total	82.966	28			
Delayed Test	Between Groups	29.638	2	14.819	7.630	.002
	Within Groups	50.500	26	1.942		
	Total	80.138	28			

Table 10. Post-hoc Test Results on the Immediate-test in the High-level

Group	Inter-dependent variable	Mean difference	*SE*	*Sig.*
Reading	Phonology	-1.622	.672	.016
	Fill-in	-2.000	.654	.015
Phonology	Reading	-	-	-
	Fill-in	.0388	.672	1.00
Fill-in	Reading	-	-	-
	Phonology	-	-	-

Table 11. Post-hoc Test Results on the Delayed-test in the High-level

Group	Inter-dependent variable	Mean difference	*SE*	*Sig.*
Reading	Phonology	-.4333	.640	1.00
	Fill-in	-2.300	.623	.003
Phonology	Reading	-	-	-
	Fill-in	-1.866	.640	.222
Fill-in	Reading	-	-	-
	Phonology	-	-	-

5. Discussion

This study investigated whether the phonological decoding process in the framework of ILH improves vocabulary learning. To examine the effectiveness of phonological decoding process, three tasks were created (reading, phonological decoding, and fill-in) and given to Korean EFL students with similar proficiency levels. Before discussing research questions directly, one of the findings is worthy of getting our attention for discussion. It seemed that the fill-in task was most effective in the immediate-test, but became less effective in word retention (delayed-test). It also showed significant mean differences only in the immediate-test between the reading task (involvement load 1) and the fill-in task (involvement load 2). This finding supports ILH in which the task with higher involvement load leads to more vocabulary learning than the task with lower involvement load (Hebrew-English experiment in Hulstijn and Laufer 2001, Kim and Na 2010, Park 2020), but it also contradicts findings from studies (Dutch-English experiment in Hulstijn and Laufer 2001, Kim 2011, Lee and Kim 2015, Park 2017) in which the fill-in task was not proven to be more effective than the reading only task. While these inconsistent research results require further research in order to more accurately evaluate the effectiveness of gap-filling activity, one of the possible explanations is that the amount of load to which the degree (moderate and strong) of each individual component (need, search, and evaluation) contributing to the overall involvement load might not be the same (Park, Yun and Lee 2019). More specifically, with the results of this research and previous findings, learners' activity of gap-filling with target words during reading might not induce moderate cognitive process (load 1) from learners. Various experiments testing the degree of each component would be necessary to review the task types and the amount of their induced-load suggested by Hulstijn and Laufer (2001).

Next, the first research question asked whether the phonological decoding process in a task affects initial and retention of vocabulary learning. It was found that although the phonological decoding group did not show main effects in relation to the reading group and the fill-in group both in the immediate and the delayed test, the mean

scores of each group showed that participants doing phonological decoding task demonstrated a better initial word learning (*m* = 4.18) than the reading group (*m* = 3.33), and best retention of vocabulary (*m* = 2.77) compared to other two groups: the reading task group (*m* = 2.09) and the fill-in group (*m* = 2.19). This order in the amount of word learning among the tasks in the immediate test and the delayed test is the same as in that of Park's (2019) study even though the main effects of phonological decoding group in relation to the reading group and the fill-in group are different. These findings still emphasize the importance of phonological decoding process in the task. Park's (2019) study showed significant mean differences between the phonological decoding group and the reading group, and between the phonological decoding group and the fill-in group. Despite the replication of the same task with same material, the differences in the results suggest taking participants' language skills into consideration for analysis in order to get more thoughtful insight on the effectiveness of tasks in terms of learners' proficiency levels. The participants' proficiency level in the Park's (2019) study was a little higher (m = 21.38 out of 30) than that of the participants in this study (m = 19.00 out of 30), and this leads to analysis of data based on learners' proficiency levels in this study.

The second research question asked whether the effect of the phonological decoding on vocabulary learning vary depending on learners' proficiency levels. As for the immediate- test vocabulary gains in the low-level, the mean of the phonology group (m = 4.30) was higher than the reading group (*m* = 2.77) and the fill in group (*m* = 3.77), and its significance was found between the reading group and the phonology group, and between the reading group and the fill-in group. However, there were no significant mean differences between the phonology group and the fill-in group. One of the things worthy of our attention is that the mean score of the phonology group was the highest in the immediate-test, and it has been preserved in the delayed-test as well; the mean scores in the delayed test were 2.60 for the phonology group, 1.33 for the fill-in group and 0.63 for the reading group. Figure 1 provides mean differences of each task visually in both the immediate-test and the delayed-test.

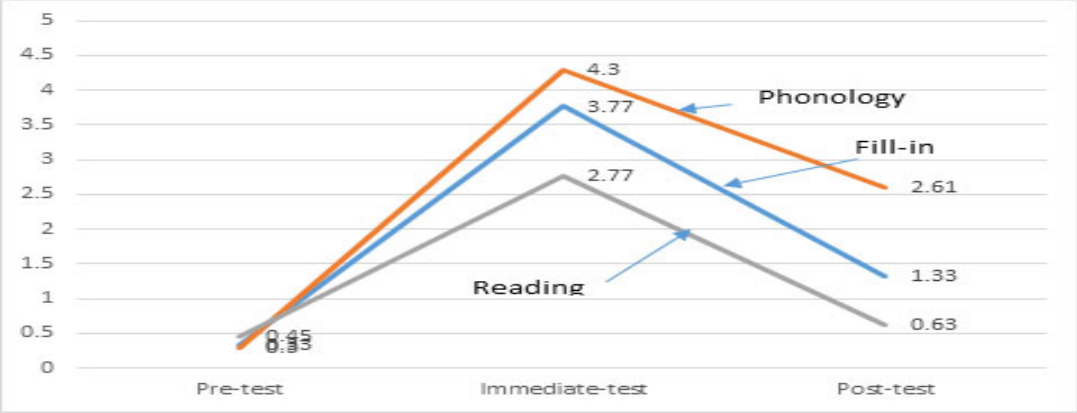

Figure 1. Mean Differences between Immediate-test and Delayed-test of Low-level

The significant mean differences were also found between the reading group and the phonology group in both immediate-test and the delayed-test, which indicates that the phonological decoding process was helpful for the low-level learners to both retrieve meanings of unknown word and retain learned words in a long-term memory. These results are in accordance with previous findings in that phonological decoding process of target words facilitates the retention of vocabulary knowledge more than the reading comprehension task (Park 2019). Park (2019) found significant mean differences only in the delayed-test whereas the current study indicates significances both in the immediate-test and the delayed-test. It is assumed that phonological decoding process would effectively

work in the low-level group.

This tendency was not found in the high-level groups in which the significant differences between the reading group and the phonological decoding group were found only in the immediate-test, not in the post-test. Learners in the phonology group were not able to retain unknown words as much as the reading group and the fill-in group. Even though the mean score (m = 2.33) is higher than the reading group (m = 1.90), it did not lead to statistical differences, and also no differences were found between the fill-in group (m = 4.2) and the phonology group (m = 2.33). Given the findings of this study along with the results of park's (2019) study, one point needs to be addressed in regard to including the phonology decoding activity into one of the involvement loads. Despite the main effects of the phonological decoding task, it does not override the effect of the fill-in task which has involvement load of 2, particularly in the high-level learners. Thus, it is assumed that the involvement load of phonological decoding process would be between 1 and 2, or could be 0.5 effectively working for the low-level students. Although phonology decoding activity is not directly related to learners' activity for word-meaning in the framework of ILH, it still needs to be incorporated into one of the components because learners under any circumstance have to process the form of target word for word learning. Then, it would be necessary to include phonological decoding activities into the reading assignments for word learning for the low-level students to be able to acquire word-meaning of unknown word.

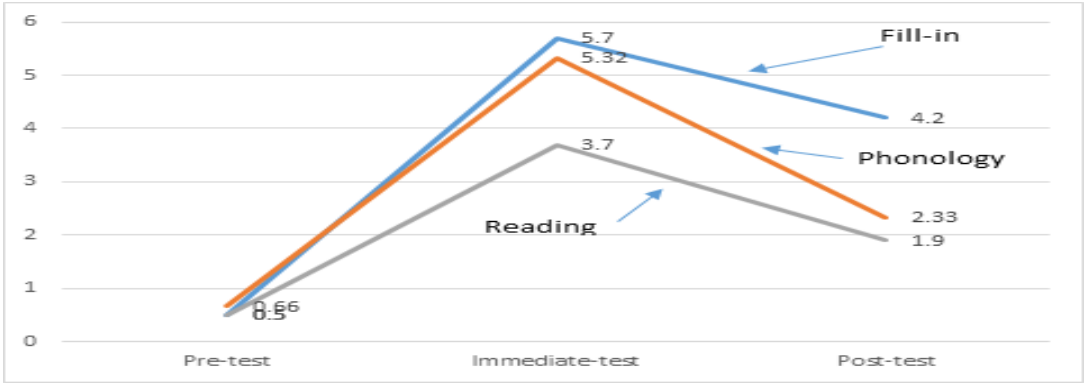

Figure 2. Mean Differences between Immediate-test and Delayed-test of High-level

It seems that phonological decoding process of words is one of the effective learning methods for literacy development in both L1 and L2. Effectiveness of phonological activities including phonology training, phonological awareness, and phonological clustering was found for L2 reading development, word-learning, and word-meaning inference (Chen 2013, Hamada and Koda 2010, Li and Chen 2016, Wilcox and Medina 2013). For example, Chen's (2013) investigation on the phonological knowledge of L2 revealed a correlation between phonological awareness and reading development among the lower proficiency learners. She conducted three tasks of phonological awareness, prosodic skills and reading comprehension to clarify the relationship among them. Non-word reading task which is one of the phonological awareness tasks indicated playing a major role in L2 learners' reading comprehension.

The importance of phonology of L2 was also found in the L2 word-meaning inference strategies, particularly for students of alphabetic L1 learning English L2 which is also alphabetic language. Hamada and Koda (2010) conducted an experiment to clarify precisely whether L2 decoding efficiency contributes to word-meaning inference to different degrees among the learners who are different in their L1 backgrounds (Korean alphabetic language vs. Chinese logographic language). The relationship between phonological decoding efficiency and

meaning-inference strategies was found stronger among Korean ESL learners than among Chinese ESL learners. Given the importance of phonology in word learning particularly for low-level learners, it also seems that low-level learners rely more on the local strategies for successful word-meaning inference, which involves phonological decoding process. Hamada and Park (2013) reported low-level learners frequently use local strategies for accurate meaning inference compared to high-level learners. On one hand, only local strategies including phonological activity might be in their control due to their low linguistic proficiency level. On the other hand, those strategies are something we teacher can encourage low-level learners to employ for word-learning until they reach a certain point where they are able to grasp the overall context of reading. Give the results of this study along with the aforementioned studies, phonological decoding process or its awareness is quite important for low-level learners, but it was never considered as one of the components in the framework of ILH. It is strongly suggested that phonological decoding needs to be considered as one of the dimensional components of involvement loads, and it should be included when creating task materials particularly for low-level students.

Educational implications can be drawn from these findings. Inasmuch as learners' proficiency levels are one of the points affecting the effectiveness of word-learning task materials, the educators need to carefully consider learners' proficiency levels when creating task materials. First, educators in instructional environment would need to take phonological decoding activities into account for low-level learners when creating word-learning task materials. Since it could be one of the limited strategies learners can utilize within their proficiency levels, we teachers need to carefully include a few phonological activities in the task which facilitate word-learning process. Second, educators need to create higher involvement load tasks to assist high-level learners for their long-term retention of target words. For instance, tasks inducing high-levels of evaluation from learners such as writing an essay or sentence writing which induce two points in the involvement load should be considered as more effective word-learning task material.

6. Conclusion and limitations

This study provides an account of ILH which focuses on the contribution of phonological decoding process in initial word learning and its retention. The results provide an evidence supporting the effectiveness of phonological decoding of the participating Korean EFL learners in the involvement load task. In particular, it was effective among the low-level participants where the phonological group was best in both initial word-learning and its retention for a long-term memory. Regarding the components of involvement load, the results also suggest a new perspective in the framework of ILH by supporting the importance of phonological decoding process. It is suggested that the phonological decoding activity in a vocabulary learning task could entail some involvement loads, leading to more word gains than the reading only task. To be more specific, it seems that for the low-level learners, phonological decoding process can be regarded as one of the components of involvement load for both short-term memory and long-term retention, and for the high-level learners, it would be an effective involvement load only for short-term learning.

A critical issue has been identified for further research. Learners' proficiency levels which were one of the analysis points were not measured consistently in all studies mentioned. For instance, various measures have been used to evaluate participants' linguistic proficiency levels; TOEFL scores were used in Kim (2011) and Hulstijn and Laufer (2001), a citywide achievement test in Sung (2019), and mock-TOEIC scores with a small number of questions in this study. Considering these various measurements for the placement of learners' levels, proficiency levels mentioned in each study should be regarded as relative compared to the other groups in the study. Thus, it

would be hard to confirm the results in terms of learners' proficiency levels since we do not know what the low-level and the high-level exactly means. Further studies should take this inconsistent measurement of proficiency levels into consideration to create more widely-used standardized test for proficiency measurement. It is strongly recommended that at least for Korean EFL learners, TOEIC scores can be standardized in measuring learners' level since it has been widely used among Korean college students. Then, the results based on the TOEIC scores would provide more realistic approach to EFL instruction in Korea.

It is also necessary to add more phonological decoding activities in the task to obtain more conclusive results. In this study, counting the syllable of the target word was only used for the experiment, but it might not be enough to facilitate phonological decoding process. Further research should incorporate various phonological decoding activities with more consistent and various learners' levels into analysis in order to confirm the results of this study.

References

Bowey, J. 2001. Nonword repetition and young children's receptive vocabulary: A longitudinal study. *Applied Psycholinguistics* 22(3), 441-469.

Chen, H-C. 2013. The roles of phonological knowledge in L2 lower achievers' reading development. *The Journal of Asia TEFL* 10(2), 1-34.

Hamada, M. 2009. Development of word-meaning inference while reading. *System* 37(4), 447-460.

Hamada, M. and K. Koda. 2010. The role of phonological decoding in second language word-meaning inference. *Applied Linguistics* 31, 513-531.

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.

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.

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. 2011. The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language Learning* 61(1), 100-140.

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.

Li, Y. and S. Chen. 2016. Relative effectiveness of phonological and morphological awareness training on L2 word reading in EFL children. *System* 60, 93-104.

Metsala, J. 1999. Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology* 91(1), 3-19

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

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* 30(2), 203-226.

Read, J. 2000. *Assessing Vocabulary*. 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.

Wesche, M. and S. Paribakht. 1996. Assessing second language vocabulary knowledge: Depth vs. breadth. *Canadian Modern Language Review* 53, 13-39.

Wilcox, A. and A. Medina. 2013. Effects of semantic and phonological clustering on L2 vocabulary acquisition among novice learners. *System* 41, 1056-1069.

Examples in: English
Applicable Languages: English
Applicable level: Tertiary

