The Effects of Receptive/Productive Tasks and Sentence Contexts on English Vocabulary Retention and Knowledge of Korean Middle School Students

수용적/생산적 과업과 문장 문맥이 한국 중학교
학생들의 영어 어휘 보유와 지식에 미치는 영향

2016년 2월

서울대학교 대학원

외국어교육과 영어전공

김 이 경

The Effects of Receptive/Productive Tasks and Sentence Contexts on English Vocabulary Retention and Knowledge of Korean Middle School Students

by

EE KYOUNG KIM

A Thesis Submitted to

the Department of Foreign Language Education

in Partial Fulfillment of the Requirements

for the Degree of Master of Arts in Education

At the

Graduate School of Seoul National University

February 2016

The Effects of Receptive/Productive Tasks and Sentence Contexts on English Vocabulary Retention and Knowledge of Korean Middle School Students

수용적/생산적 과업과 문장 문맥이 한국 중학교
학생들의 영어 어휘 보유와 지식에 미치는 영향

지도교수 이 병 민

이 논문을 교육학 석사 학위논문으로 제출함
2016년 2월

서울대학교 대학원
외국어교육과 영어전공
김 이 경

김이경의 석사학위논문을 인준함
2016년 2월

위 원 장 ＿＿＿＿＿＿＿＿＿＿＿＿＿＿

부위원장 ＿＿＿＿＿＿＿＿＿＿＿＿＿＿

위   원 ＿＿＿＿＿＿＿＿＿＿＿＿＿＿

The Effects of Receptive/Productive Tasks and Sentence Contexts on English Vocabulary Retention and Knowledge of Korean Middle School Students

APPROVED BY THESIS COMMITTEE:

_____

Jin-Wan Kim, COMMITTEE CHAIR

_____

Sun-Young Oh

_____

Byungmin Lee

ABSTRACT

The present thesis attempts to investigate the effects of task type (productive versus receptive) and sentence contexts (same versus diverse) on the vocabulary learning of Korean middle school English students in two areas: overall vocabulary learning, and the gain and retention of specific vocabulary knowledge. First, this study will look at the impact that the two variables have on overall vocabulary learning; measured using the sum score of five different tests (recognition, passive word learning, active word learning, and two productive vocabulary use tests: gap-filling and word reordering). Second, the gain and retention of specific vocabulary knowledge measured by the five tests will be compared to verify the impact of the two variables.

The receptive and productive aspects of vocabulary have been derived from the two fundamental communication processes, so both aspects are worth studying. Although many studies agree on the superiority of productive tasks over receptive tasks in vocabulary instruction, the results between these studies have been inconsistent; therefore, more research is needed on the impact of these two tasks.

Moreover, vocabulary tasks in Korea largely depend on receptive vocabulary instruction rather than productive instruction, which goes against the majority of findings from previous research that suggest productive vocabulary instruction is more effective. Context, the other important factor for vocabulary learning, has been a controversial issue in the vocabulary instruction research.

Many studies were conducted to determine whether context should be provided for vocabulary learning but few studies were conducted on how to effectively provide context for vocabulary instruction. In other words, these two factors are significant factors influencing vocabulary learning, but few studies have been conducted to investigate the relationship between these two variables. Therefore, this study attempts to integrate sentence contexts into the types of tasks so that the interactive effect of both variables on vocabulary knowledge development can be examined. Besides, this study attempts to scrutinize the multifaceted features of lexical knowledge, so five different sorts of assessment have been implemented.

In this study, 117 3rd grade middle school students in Korea completed one of four different treatment combinations, each having a different combination of the two task types and two sentence contexts (receptive task and same context - RS, receptive task and diverse context - RD, productive task and same context - PS, productive task and diverse context – PD), and took immediate and one-week delayed post-tests. Each of the two tests was composed of five different tests.

Regarding overall vocabulary learning, the results of this study revealed that task type was a factor that significantly affected vocabulary learning in both immediate word gain and its retention but sentence contexts were not. However, the interaction effect between the two variables was shown in word retention. The same context had a positive effect on the productive task but not on the receptive task. The findings from the individual analysis of the five vocabulary

tests showed similar results regarding word retention with the exception of the two productive use tests. The task effect was substantial, while that of context was not.

Above all, the productive task was statistically shown to have considerable power to help students retain several stages of vocabulary knowledge with the exception of the productive use tests. When combined with the task, sentence contexts had a strong effect on vocabulary learning in passive and active word learning tests. On the other hand, the retention of word knowledge, measured by the productive use of vocabulary tests, was influenced fundamentally by sentence contexts rather than task type. Results and the implications regarding task types and sentence contexts are discussed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1.
INTRODUCTION

This chapter introduces the research by presenting the purpose of the study. Section 1.1 discusses the purpose of the study. Section 1.2 presents the research questions, and Section 1.3 outlines the overall structure of the study.

1.1. The Purpose of the Study

Vocabulary is one of the significant factors in language learning since lexical knowledge is the most fundamental and essential for actual communication. Thus, Wilkins (1972) stated that "without grammar very little can be conveyed, without Lexis nothing can be conveyed" (p.11), representing that vocabulary mostly conveys its meaning in order to comprehend and produce messages.

Although many practitioners and learners agree on the importance of vocabulary instruction and often ascribe communication breakdown to the lack of vocabulary knowledge, vocabulary is one of the most neglected issues in the ESL research field (Zimmerman, 1997). Because of insufficient lexical input in EFL/ESL settings, it is a significant challenge for EFL/ESL learners to possess sufficient lexical knowledge.

Regarding the sufficient amount of vocabulary knowledge needed, Nation (2006) and Schmitt (2008) advocated that English learners have to know about 8,000-9,000 words for reading and 5,000-7,000 words for speaking and listening.

Not surprisingly, many ESL/EFL students fail to reach that vocabulary level without explicit vocabulary instruction (Nation, 2006). This creates a demand for more effective vocabulary instruction in ESL/EFL education settings.

Vocabulary instruction, especially in Korea, depends largely on students and their rote memorization of isolated single words. Most vocabulary tasks employed were mostly receptive-oriented (Kim, 2013). Receptive-centered vocabulary instruction may lead to discrepancies between English learners' comprehension and their production of words. Korean learners of English may have no difficulties retrieving some words for receptive uses such as reading and listening but it is difficult for them to retrieve them for productive purposes such as writing and speaking. Hence, effective vocabulary learning that can provoke both receptive and productive vocabulary knowledge is required.

The efficiency of vocabulary learning can be enhanced when words are provided with definitions and contextual clues and processed at a deeper level (Stahl & Fairbanks, 1986). It is necessary to select, sequence and present vocabulary appropriately and to choose the right tasks that integrate vocabulary knowledge development into communication when designing effective vocabulary instruction (Nation, 2001). Therefore, it is important to guide learners by providing them with appropriate task types and context for more effective and efficient vocabulary instruction.

As to the vocabulary task, its receptive and productive aspects have been explored a lot in the previous studies. The receptive and productive aspects of vocabulary have been derived from the two fundamental communication

processes: comprehension and production (Nation, 2001). Based on input and output process of communication, receptive task and productive task contribute considerably to vocabulary learning (An & Min, 2011; Shintani, 2011; Stahl & Fairbanks, 1986). Moreover, a word is presented in a relevant context in the process of communication which implies that context can be a more useful tool for language learning, especially vocabulary learning (Sternberg, 1987). For these reasons, the effects of task types and context on vocabulary learning are important factors to be explored

Many researchers have studied how differing task types, receptive and productive vocabulary instruction, involve lexical knowledge development. Although the majority of research agreed on the priority of productive tasks over receptive tasks in vocabulary instruction (Hulstijn & Laufer, 2001; Kim, 2013; Son, 2007; Pichette, De Serres, & Lafontaine, 2011; Webb, 2005), there were some studies opposed to this result (Barcroft, 2004). That is, the efficacy of receptive versus productive tasks on language learners' vocabulary learning has been open to debate (Webb, 2005). Moreover, vocabulary tasks in Korea depend largely on receptive vocabulary instruction rather than productive instruction, which required the productive vocabulary teaching. For that reason, exploring the effects of both task types, receptive and productive, on vocabulary knowledge gain and retention may provide valuable data that can enhance current vocabulary instruction in Korea.

Sentence contexts, the other important factor for vocabulary learning, have been a controversial issue in the field of vocabulary instruction as well. A lot of

researchers suggested the positive effect of sentence contexts on vocabulary learning by simulating schema and providing sufficient cognitive cues so as to help reinforce word retention (An & Min, 2014; J. R. Anderson, 1990; Bolger, Balass, Landen, & Perfetti, 2008; Schouten-van Parreren, 1989). Other researchers questioned the positive impact of sentence contexts since they tend to increase the cognitive load (File & Adams, 2010; Mondria & Wit-de Boer, 1991) and scatter learners' attention with too many cues (Hu & Nassaji, 2012; Nation & Coady, 1988).

Many studies were conducted to determine whether sentence contexts should be provided for vocabulary learning, but few studies were conducted how to provide the sentence contexts for vocabulary learning. Only An and Min (2014), Bolger and Balass et al. (2008) and Sternberg's (1987) studies dealt with the effect of sentence contexts. That is, they compared the differential effect between the diverse contexts and the same context on vocabulary learning. Although both studies proved the benefits of the diverse sentence contexts, the sentence contexts were only given through the receptive tasks and the number of experiments was small. Thus, further investigation of the role of the sentence contexts in vocabulary learning is required.

When a word was given with its sentence context without its definition, learners guess its meaning from the sentence. However, context guessing can be influenced by other variables (Nagy, 1995) and can lead to a false grasp of the word definition. Thus, this present study provides a definition of the target vocabulary which was regarded as an important factor for effective vocabulary

instruction (for example, An & Min, 2014; Bolger et al.; 2008; Stahl & Fairbanks, 1986). In addition, as suggested by Mondria & Wit-de Boer (1991), a single sentence context per a target word can diminish learners' cognitive load and prevent their attention from scattering due to excessive cues (Kim, 2013; Pichette et al., 2011), so this study employed a sentence context per target vocabulary for word learning tasks.

Overall, previous studies on receptive and productive vocabulary learning merely focused on comparing the effects of the two tasks. This study attempts to integrate sentence contexts into the types of tasks so that the interactive effect of both variables on vocabulary knowledge development can be examined.

It can be another significant issue to define lexical knowledge because of its multifaceted feature (Laufer & Nation, 1999), which needs to be reflected properly in studies on vocabulary instruction. In Korea, however, the vocabulary tests were usually limited to simple tests requiring 1:1 translations of context-excluded word items. This assessment measure is not sufficient enough to measure multifaceted vocabulary knowledge. The issue calls for a more comprehensive assessment.

Lexical processing needs to be investigated further to comprehend what it means to "know" a word and to further discover the constructs of vocabulary knowledge (Nation, 2001). In the endeavor to evaluate ESL learners' vocabulary knowledge, Paribakht and Wesche (1993) created a Vocabulary Knowledge Scale (VKS). The VKS requires learners to self-report their knowledge with five levels of word recognition, ranging from passive word knowledge to its

composition, showing that their vocabulary develops from partial to full knowledge (Nation, 2001).

This study extracted stage 1 to 3 of the VKS for the word recognition test and stage 4 for the passive word learning test. Because of the huge gap in difficulty level between stage 4 (retrieval of a target word) to stage 5 (free writing with a target word), learners may fail to prove their productive use knowledge of words (Bolger et al., 2008). The negative results of stage 5 can be triggered by the lack of their language proficiency, not by their vocabulary knowledge deficiency.

Therefore, assessing the learners' productive use knowledge of words by the VKS had a limitation, especially in ESL/EFL learners with low language skills. Specifically, most Korean middle school students are not familiar with composition, so it is difficult to assess their productive use of word knowledge through free-writing. This present study employed two productive use tests, gap-filling and word reordering, in order to compensate the limitation of the VKS. In addition, the VKS omitted the retrieval of word items from its meaning, called active word knowledge; this study also added the active word learning test.

According to An and Min (2011), the context-included tests led to a significant difference from the context-excluded tests, so both types of assessments are required for an in-depth understanding of lexical knowledge. To understand the overall depth of vocabulary knowledge and its development, five test items were employed to evaluate word recognition, passive and active word knowledge, productive word use in proper context and grammar.

To summarize, the impact of task types (productive versus receptive) and sentence contexts (diverse versus same) on vocabulary learning are controversial issues. Furthermore, little research that demonstrates the influence of the two variables on lexical knowledge development considering various factors has been conducted. As a result, the present study investigates those interventions related to vocabulary instruction in the Korean EFL classroom setting to observe how those factors contribute to Korean middle school English learners' lexical knowledge development.

1.2. Research Questions

The focus of the present study is to investigate the effects of receptive/productive task and sentence contexts on the vocabulary learning of Korean middle school English learners from the following two perspectives. First, this study looks into the impact of task types (productive versus receptive) and sentence contexts (same versus diverse) on the overall vocabulary learning of Korean middle school English learners. Here, the overall learning refers to the sum of five different test scores. Second, the gain and retention of specific vocabulary knowledge measured by five different tests are compared regarding task types and sentence contexts. The participants in this study completed one of four treatments with different task and context combinations and took the five sorts of the immediate and delayed post-tests. Every experimental process was

thoroughly developed and administrated to answer the following research questions.

1. How do the type of task (receptive versus productive) and sentence contexts (diverse versus same) influence Korean middle school English learners' overall immediate vocabulary learning and its retention?

2. How do these two factors influence Korean middle school English learners' vocabulary knowledge measured by five different vocabulary tests?

1.3. Organization of the Thesis

The present study consists of five chapters. Chapter 1 introduces the purpose of the study and presents the research questions. Chapter 2 provides an overview of the literature review on vocabulary knowledge and the effect of task and context on vocabulary learning. In Chapter 3, the methodology of this study is described regarding the research design, the participants, the procedure, the instruments, the treatment, the assessment, and the data analysis. Chapter 4 presents the results and discusses the research findings. Finally, Chapter 5 concludes the research with a summary of the significant findings and shows the implications of the present study and the suggestions for further research.

CHAPTER 2.
LITERATURE REVIEW

The current chapter presents the literature overviews about the effect of task type—receptive and productive—and context on vocabulary knowledge development. Section 2.1 discusses vocabulary knowledge in specifying its three components: receptive and productive aspects, breadth and depth, and context of vocabulary use. Section 2.2 details the main issues involved in this study—receptive versus productive task and context that calls for comprehending vocabulary knowledge development

2.1. Vocabulary Knowledge

The issue of "knowing" vocabulary had been demonstrated and debated among a large number of previous studies. Bachman and Palmer (1996) stated that vocabulary knowledge is the ability to use general and concrete words precisely with the appropriate contexts. However, word knowledge is a multifaceted construct (Laufer & Nation, 1999) that calls for proper reflection in vocabulary acquisition research. Therefore, many researchers tried to demonstrate the vocabulary knowledge construction (Chapelle, 1994; S. M. Gass, Behney, & Plonsky, 2013; Henriksen, 1999; Laufer & Nation, 1999; Read, 2000).

Chapelle (1994) divided vocabulary ability into three components: the

context of vocabulary, fundamental procedures of vocabulary knowledge, and metacognitive strategies of vocabulary use. On the other hand, some researchers (Henriksen, 1999; Read, 2000) defined vocabulary knowledge with three different aspects: "partial–precise knowledge," "depth of knowledge," and "receptive–productive control," during its gradual development stages (p. 304). Partial–precise knowledge refers to the progressive vocabulary development. As mentioned in Read's (2000) study, breadth and depth of knowledge are the quantity and quality of learners' vocabulary knowledge. Receptive and productive control of vocabulary knowledge was related to its comprehension and production. Previous studies, such as Gass et al. (2013), suggested similar components of lexical knowledge such as "production and reception," "knowledge and control," and "breadth and depth".

All things considered, three exemplary components of vocabulary knowledge were frequently cited: reception and production, vocabulary breadth and depth, and the context of vocabulary use, which are discussed in the following sections in more detail.

2.1.1. Reception and Production

The receptive and productive aspects of vocabulary have been doubtlessly regarded as to exist derived from the two fundamental communication processes: comprehension and production (Nation, 2001). However, there is no

clear-cut way of distinguishing between receptive and productive aspects in word knowledge. Rather, vocabulary knowledge was considered to be gradually developed from receptive to productive phases (Laufer & Paribakht, 1998; Melka, 1997). It is complicated to conceptualize which part of the continuum is occupied by the receptive aspect or that of the productive, and even more intricate to put the absolute threshold where vocabulary is developed from receptive to productive phases (Read, 2000). However, the segregation to put vocabulary knowledge on either stage can be practical (Melka, 1997).

Many researchers coined their definitions in a bid to delineate the term "receptive" and "productive" (Gass et al., 2013; Henriksen, 1999; Meara, 2009; Nation, 2001). Nation (2001), for example, described the "receptive" phase as the process to receive language input and to comprehend its meaning through listening or reading. The "productive" phase, on the other hand, is the procedure to generate language output and deliver a particular message through speaking or writing. The terms "receptive" and "productive" from previous studies were used to entail the receptive and productive facets of language processes and the use of receptive and productive language skills.

Also, the terms "receptive" and "productive" are often described as the corresponding terms, "active" and "passive," which are related to one another. Meara (1990), for instance, delineated "active vocabulary" can be activated through word association. "Passive vocabulary," in contrast, can only be triggered by a receptive stimulus such as reading and listening (Meara, 1990).

The division between the receptive and productive aspects of vocabulary

knowledge entails different facets of vocabulary knowledge such as lexical procedures, language skills, and word associations (Gass et al., 2013). In other words, the distinction between the two aspects is a complex mixture of several aspects of vocabulary knowledge (Henriksen, 1999). However, reception and production themselves, as the primary domains of vocabulary knowledge (Melka, 1997), will be observed in this study to explore the development of vocabulary knowledge.

2.1.2. Vocabulary Breadth and Depth

According to previous research, "breadth" and "depth" of vocabulary knowledge were key issues in language development (S. M. Gass et al., 2013). Milton (2009) defined the "breadth" as a learner's vocabulary size and "depth" as the quality of the learner's lexical knowledge. That is, breadth of word knowledge shows how many words someone knows and depth refers to what they known about those words (Milton, 2009).

Previous research addressed that degrees of knowledge (Melka, 1997; Paribakht & Wesche, 1993; Wesche & Paribakht, 1996), and word association (Meara, 2009; Read, 2000) were eloquently related to breadth and depth of lexical knowledge.

Regarding "breadth" of vocabulary knowledge, several studies were conducted. Goulden, Nation, and Read's (1990) study, for example, indicated

that English native speakers know about 20,000-word families on average. Nation (2006) concluded that English learners are required to know about 8,000-9,000 word families for reading and 6,000-7,000 for speaking. Schmitt (2008) reached a similar conclusion that language learners have to know about 8,000-9,000 word families for reading, and 5,000-7,000 for speaking and listening. Specifically, for written or oral communication, at least 98-99% of vocabulary should be possessed by English language learners (Hsueh-Chao & Nation, 2000).

Doubtlessly, not many students can reach this stage (Nation, 2006). Based on those findings, vocabulary should be strategically selected for vocabulary instruction, especially in EFL/ESL settings, to achieve vocabulary knowledge in a more effective and efficient way.

As part of an endeavor to gauge vocabulary breadth and depth and its receptive and productive aspects representing word knowledge development, Paribakht and Wesche (1993) created a Vocabulary Knowledge Scale (VKS) to evaluate EFL/ESL learners' vocabulary knowledge. The VKS asks learners to self-report their knowledge of each word by responding to the following statements

(1) I have never seen this word.

(2) I have seen this word before, but I don't know what it means.

(3) I have seen this word before, and I think it means _____.

   (Synonym or translation)

(4) I know this word. It means _____. (Synonym or translation)

(5) I can use this word in a sentence.

(Paribakht & Wesche, 1993, p. 15)

This assessment shows that a learner's vocabulary develops from partial to full knowledge through the "semantization" process (Nation, 2001) including the specific nature of this development from word recognition to its productive use in context.

Considering the lexical developmental stages mentioned above, vocabulary breadth and depth should be considered in the vocabulary learning processes (Milton, 2009; Nation, 2013; Nation & Gu, 2007).

2.1.3. Context of Vocabulary Use

A large number of studies detailed earlier asserted that the context in which a word is used makes up a significant part of the lexical ability. (Bachman & Palmer, 1996; Chapelle, 1994; Martinez, 2010).

Chapelle (1994), for instance, regarded the context in which vocabulary is used as one of the three major components of vocabulary knowledge. Bachman and Palmer (1996) considered lexical knowledge as the knowledge of words and their appropriate use in the appropriate context. They believed that the development of vocabulary knowledge calls for the ability to use vocabulary in the right context as well as its incremental gain (Bachman & Palmer, 1996). Martinez (2010) also pointed out that the acquisition of vocabulary knowledge is

the procedure during which a learner deliberates a target word, retrieves its lexical information and uses it in a proper context.

On the whole, it is beneficial to measure diverse aspects of lexical knowledge in order to deal with the complexity of vocabulary learning development. This compensates for a single component of vocabulary knowledge which would hardly capture the dynamic aspects of vocabulary (Gass et al., 2013).

2.2. Research Issues in Vocabulary Instruction

Nation (2001) proposed that effective vocabulary instruction demands decisions to select, sequence and present vocabulary while choosing appropriate tasks in order to integrate lexical progress into communication. As an effort to discover effective vocabulary instruction, research has been conducted on various issues such as vocabulary knowledge development (Anderson & Freebody, 1981; Koda, 1989; Read, 2000), the relationship between vocabulary knowledge and language proficiency (Koda, 1989; Qian, 2002), word frequency (Balota & Chumbley, 1984; Eckerth & Tavakoli, 2012; Hu, 2013; McKeown, Beck, Omanson, & Pople, 1985; Rayner & Duffy, 1986; Rott, 2007), explicit versus implicit learning (Berry & Broadbent, 1987), incidental versus intentional learning (Brown, Waring, & Donkaewbua, 2008; Day, Omura, & Hiramatsu, 1992; Ghabanchi & Ayoubi, 2012; Heidari-Shahreza & Tavakoli, 2012;

Hemmati & Asmawi; Joe, 1998; Laufer & Hulstijn, 2001; Laufer & Rozovski-Roitblat, 2011; Paribakht & Wesche, 1999; Pellicer-Sánchez & Schmitt, 2010; Song & Sardegna, 2014; Srichamnong, 2008; Webb, 2008, 2012), vocabulary assessment (Bachman & Palmer, 1996; Laufer & Nation, 1999; Schmitt, Schmitt, & Clapham, 2001), vocabulary learning strategies, task effect on vocabulary learning (An & Min, 2011; Bolger et al., 2008; Hulstijn & Laufer, 2001; Kim Ji, 2014; S. S. Kim, 2013; Pichette et al., 2011; Ryoo, 2009) and the effect of context on vocabulary learning (An & Min, 2014; Bainbridge, Lewandowsky, & Kirsner, 1993; Bensoussan & Laufer, 1984; Bolger et al., 2008; Carroll & Drum, 1982; Mondria & Wit-de Boer, 1991; Nagy, 1995; Nagy, Anderson, & Herman, 1987; Nagy, Herman, & Anderson, 1985; Schatz & Baldwin, 1986; Stallman, 1991; Sternberg, 1987; Webb, 2008). Among these issues, this study will mainly examine the effects of the task type and context on vocabulary learning.

2.2.1. Effects of Receptive and Productive Tasks on Vocabulary Learning

The two task types, receptive and productive, are commonly assumed to reflect input and output of communication in a number of previous studies (Amiryousefi & Kassaian, 2010; An & Min, 2011; Bao, 2015; De La Fuente, 2002; Folse, 2006; Hazrat, 2015; Jeon & Shin, 2011; S. Y. Kim & Lee, 2008; Laufer, 1998; Lee, 2003; Llach, 2009; Melka, 1997; Mondria & Wiersma,

2004a; Waring, 1997; Webb, 2005). There has been a consensus that both receptive and productive tasks, based on input and output process of communication, contribute considerably to vocabulary learning (An & Min, 2011; Shintani, 2011; Stahl & Fairbanks, 1986).

Among several definitions of receptive and productive vocabulary learning, this study adopted Mondria and Wiersma's (2004) terminology, as follows:

(1) Receptive vocabulary learning is to learn the meaning of an L2 word. Learning a word is going from L2 to L1.

(2) Productive vocabulary learning is to express a concept using an L2 word. Learning a word is going from L1 to L2. (p. 38)

A large number of studies were conducted to discover the efficacy of receptive and productive tasks on learners' vocabulary learning, but the results were rather contradictory (Barcroft, 2004; Choi, 2007; Griffin & Harley, 1996; Hulstijn & Laufer, 2001; S. S. Kim, 2013; Laufer & Hulstijn, 2001; Son, 2007; Waring, 1997; Webb, 2005).

Most research has proved the superiority of the productive task over the receptive task on either immediate vocabulary gain (Pichette et al., 2011), vocabulary retention (Hulstijn & Laufer, 2001) or both (Kim, 2013; Son, 2007; Webb, 2005).

Hulstijn and Laufer (2001) conducted research about EFL students' incidental short-term and long-term vocabulary retention after three different tasks: one productive task (free writing) and two receptive tasks (reading with fill-in and reading only) with various task involvement loads. As predicted,

retention was higher in the productive task compared to the two receptive tasks. It was highest in the composition, lower in the fill-in-the-blank task with reading, and lowest in reading only.

Webb (2005) discovered how Japanese EFL students learned target vocabulary using three glossed sentences and a sentence composition task. Five elements of vocabulary knowledge—"orthography, syntax, association, grammatical functions, and meaning and form" (p. 33)—were assessed. With the sufficient amount of time for task completion, the productive task was more effective for vocabulary gain as well and its retention.

Son (2007) examined Korean university students' immediate vocabulary gain and its retention by comparing one productive task and two receptive tasks with differential task loads and the combination of all three tasks. Corresponding to Hulstijn & Laufer's (2001) research, among a single task, the composition task resulted in the highest scores in immediate and delayed post-tests. However, unlike other results, there was no significant difference between two repetitive tasks with differing involvement loads. This study only proved the differential impact between different task types, productive and receptive, rather than those of involvement loads.

Pichette et al. (2011) investigated the relative effect of reading and writing sentences for ESL French learners' incidental vocabulary learning. Different from the results of the Son (2007) and Webb (2005), which implied the superiority of productive task over the receptive on immediate and delayed tests, immediate recall scores showed superior recall for writing tasks over reading

tasks while delayed recall scores demonstrated no differences between them over time.

Compared to the research results that confirmed the dominance of productive task effect over that of receptive on overall vocabulary knowledge gain and retention, the results of Griffin and Harley, (1996) and Waring's (1997) research proposed that the vocabulary task types are widely influenced by the types of vocabulary knowledge. In other words, the receptive task made learners gain more receptive vocabulary knowledge, whereas the productive task led students to learn more productive vocabulary knowledge.

Some studies even proposed the dominance of receptive tasks over the productive task in vocabulary learning and retention. The results of Webb's (2005) first experiment, within the same limited amount of time, showed that the receptive task was superior to the productive one. Although, as time passed, the superiority of the receptive task disappeared, receptive vocabulary tasks still make up an important part of vocabulary learning, which was shown in Choi's (2007) study. Choi's study partially replicated Webb's (2005) study. Choi (2007) showed that receptive tasks yielded better gains in both receptive and productive vocabulary.

The majority of previous research agreed on the positive effects of productive tasks, whether it is partial or full, on overall vocabulary learning, or at least on productive vocabulary learning. Barcroft's (2004) research, however, showed the opposite results. Barcroft (2004) compared the effects of writing new sentence including target words with those of word-picture repetition on L2

Spanish learners' vocabulary learning. The research findings showed a strong negative effect from the productive task, suggesting that this task can inhibit learning word forms during the initial phases of L2 vocabulary acquisition

Although a large number of studies were conducted, the effect of receptive versus productive tasks on language learners' vocabulary learning is not conclusive (Webb, 2005). Despite the inconsistency of the research results, vocabulary tasks tend to be conducted receptively rather than productively, especially in EFL settings (Kim & Lee, 2008; Webb, 2005). According to Kim and Lee (2008), Korean EFL vocabulary instruction has mainly been conducted using receptive tasks rather than productive ones. This receptive-centered vocabulary instruction would hamper students' output production.

Therefore, investigating the effects of two types of task–receptive and productive–may provide important implications to improve current vocabulary instruction in Korea. In that sense, this study aims to compare the main effects of the differing tasks in order to provide meaningful information to compensate for the inconsistency of the previous research.

2.2.2. Effects of Context on Vocabulary Learning

Vocabulary knowledge is the ability to use general and specific word items in its precise context (Bachman & Palmer, 1996). In authentic communication, a word is generally presented with relevant context, whether it

is written or said. That is, understanding context can be more useful for language learning, especially vocabulary learning (Sternberg, 1987). However, previous studies have revealed rather controversial results on the effect of context.

The effects of context on vocabulary instruction have been consistently investigated through reading the research. Many studies agreed on the positive impact of diverse contextual information on vocabulary learning (An & Min, 2014; J. R. Anderson, 1990; Bolger et al., 2008; Schouten-van Parreren, 1989; Sternberg, 1987; Webb, 2008). Nagy, Herman, and Anderson (1985)' study showed that context lead to small but statistically reliable gains in word knowledge. The incidental learning from context through learners' reading resulted in a substantial vocabulary development during the school years. Schouten-van Parreren's (1989) experiment about comprehension and retention of vocabulary in texts revealed that context with an appropriate level of difficulty is beneficial to vocabulary learning. This study supposed that reading the same words in various context sentences would provide plentiful references to retrieve word meaning.

In line with the previous studies, Webb (2008) also noted a positive effect of context on vocabulary learning after comparing the effects of different context types. He divided Japanese EFL learners into two groups: one with more contextual clues and the other with less contextual clues and made them learn target vocabulary through reading. His research proved the superiority of the more informed context group over the other on the retrieval of vocabulary meaning, but not on that of the retrieval of its form. Rather, the number of

encounters had a greater effect on retrieving the forms of words. Based on his results, Webb calculated that sentence contexts would affect different features of vocabulary knowledge in different ways, which calls for assessing various dimensions of vocabulary knowledge.

However, there have been few studies conducted that examine the sentence contexts (diverse versus same) except for Sternberg (1987), Bolger et al. (2008) and An and Min's (2014) study. Sternberg (1987) demonstrated the possible effect of context variation on vocabulary learning. According to his research, a proper level of context variable helps learners get an overall understanding of the meanings of given words. He found that repetition of the same context sentence alone could not lead to the same favorable result as repetition of multipe context sentences.

Bolger et al. (2008) explored the effect of sentence contexts and use of definitional context on vocabulary learning. A group who repeatedly encountered target words in the same sentences and another group who met the same target words but in different sentences without definition were compared. The result showed that the multiple-context group had higher scores on comprehension of word meanings and on judging whether a newly given word was proper in context. Thus, they argued that the degree of sentence contexts has a significant influence on the learner's vocabulary learning, primarily performed in a receptive way.

An and Min (2014) examined two EFL Korean learner groups with different sentence contexts which practiced target vocabulary through the receptive task.

However, each of these groups had a different sentence contexts ; diverse context or the same context. Participants repeatedly practiced the target words in the given sentences through reading and listening. This research demonstrated that sentence contexts has a statistically meaningful influence on developing vocabulary knowledge as the diverse context group's test results showed better mean scores than those of the single context group.

In contrast, some researchers (Herman et al., 1987; Jenkins, Pany, & Schreck, 1978; Lawson & Hogben, 1996) doubted the effect of context on vocabulary gain and suggested that its efficacy is rather negligible when compared to direct vocabulary instruction. Jenkins, Stein, & Wysocki (1984) argued the rather modest effects of the context variable on vocabulary acquisition are due to the redundancy of cues in the text. Each word presented plenty of contextual clues that did not need to be understood receptively, so learners did not pay selective attention to each word item (Jenkins, Stein, & Wysocki, 1984).

Mondria and Wit-de Boer (1991) reported that guessing the meanings of words through diverse contexts did not show statistically meaningful effects on word gain. Rather, they proposed a negative correlation between contextual guessing and retrieving the meaning of words. This research discovered that the easier it is for learners to guess the meaning of words from their context, the faster they tend to forget them.

Corresponding to the previous research result, Nation and Coady (1988) explained the negative correlation between context and word learning. According to Nation and Coady, language learners seldom focus on the

meanings of individual words when too many contextual clues were given since they can readily comprehend the general messages from context. This may hamper retention of the target words in the end.

File & Adams (2010) compared three ESL university learners groups that focused on taking isolated vocabulary instruction without context sentences, integrated teaching, and incidentally learning vocabulary through context. The group learning words without context sentences achieved better mean scores on vocabulary tests than those with context sentences. Although some words were incidentally learned through reading, the number of words learned was much fewer than expected. They believed that the cognitive load of comprehending context might have hinder vocabulary learning.

Although the role of context on vocabulary learning has been studied a lot, they were mostly conducted in receptive learning settings such as reading and listening, so in most cases, context richness had a significant role in understanding the meanings of target words in receptive settings (Bolger et al., 2008). This calls for research exploring the effect of context on vocabulary learning in productive learning settings.

Also, to compensate for the misleading contextual information of a target word, which was mentioned in previous research, that cast doubt on context effect, both proper definitions and contexts for new words (Stahl, 1986) should be provided for effective vocabulary instruction (Bolger et al., 2008).   .

CHAPTER 3.
METHODOLOGY

This current chapter presents the methods used in this study. Section 3.1 introduces the research design. Section 3.2 discusses the participants. The procedures of the study are described in Section 3.3. Section 3.4 provides details on the instruments regarding the target words and sample sentences used. The treatment of the receptive task group and productive task group are explained in Section 3.5. The word learning assessment methods and their scoring procedures are described in Section 3.6. Section 3.7 describes the data analysis.

3.1. Research Design

A multifactorial design with no control group was implemented in this study. When it comes to the vocabulary treatment, task groups had the significant superiority over control groups in previous studies (An & Min, 2011; Stahl, 1986). Moreover, this study aims to compare differences across four treatment groups, depending on their task types and contexts.

The independent variables (2) were task types (receptive versus productive) and sentence contexts (same versus diverse). The dependent variables (5) were five types of vocabulary tests: the recognition test, the receptive translation test, the passive/active word learning test, and the two productive vocabulary use tests: gap-filling, and word reordering.

3.2. Participants

This study was conducted from June to July 2015. All participants (N=128) were third-grade middle school students from one co-educational middle school (M), located in Sinrim-dong, Gwanak-gu district in Seoul. Most of the learners have had at least five and half years of English education: three years in elementary school, and two and half years in middle school. Only data from students who signed a consent form were used in this study. Four intact classes were chosen based on the mean scores of English mid-term and final exams. These exams had been administrated in the target school during the 1st semester of 2015. Table 3.1 shows the mean scores and the standard deviations of the English scores of the four participating classes. One-way analysis of variance (ANOVA) confirmed the homogeneity of the participating classes ($p = .525$) (F $= .749$, $p > .05$).

Table 3.1

Descriptive Statistics of 4 Participating Classes

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
C 1	29	72.655	24.0071	4.4580	63.523	81.787	18.0	97.5
C 2	30	72.717	22.5773	4.1220	64.286	81.147	23.5	98.0
C 3	28	72.661	22.8559	4.3193	63.798	81.523	18.0	100.0
C4	30	72.717	18.8001	3.4324	65.697	79.737	19.0	95.0
Total	117	72.688	21.8290	2.0181	68.691	76.685	18.0	100.0

Note. The maximum test score was 100; C = Class

One of four treatments with a different combination of task types and contexts (RD, RS, PD, PS) was randomly assigned to each of the four homogeneous classes; RD to Class 1 (32 students), RS to Class 2 (32 students), PD to Class 3 (32 students), and PS to Class 4 (32 students).

To ascertain the effects of each treatment, students who identified more than two items as pre-known words or as re-encountered words during the tests (see Appendix 5) were also excluded from the data analysis (N=11). As a result, 117 students (29 for RD, 30 for RS, 28 for PD, 30 for PS) were selected for data analysis.

3.3. Procedure

The whole process was composed of three sessions; the task, immediate test, and delayed test. The research was conducted in the classes at middle school M under the guidance of the researcher with the help of one English teacher of Korean nationality. Before the experiment, the students were informed of the purpose of this study and then read and signed the consent form (see Appendix 1). Then, an orientation session was held in which the details about the process of the experiment were explained to participants. After the orientation session, the researcher asked students to practice the eight target words in class, which involved one of the following: learning words through a receptive task with diverse context sentences (RD), a receptive task with the same context sentences (RS), a productive task with diverse context sentences (PD), and a productive task with the same context sentences (PS).

Right after the task was done, the students' learning of the target words was measured by five different types of vocabulary tests: recognition, passive word learning, active word learning, and two productive word uses: gap-filling and word reordering. A second test was administered one week after. The tests were conducted in the following order: the active word learning test, the recognition test, the passive word learning test, then the two productive use tests, gap-filling and word reordering. There were three different versions per test that presented items in a random order to avoid any fixed-order effects (Puff, 1982).

3.4. Instruments

This section details the instruments (the target words and sample sentences) implemented in the present study.

3.4.1. Target Words

All The target words in the present study were eight words that were unknown to participants. Students unrelated to the participants of this study selected a total of thirty candidate target words (6 verbs, 5 adjectives, 19 nouns) from a list of Lv 1000 and Lv 2000 words in the Academic Word List (Coxhead, 2000). The students who selected these words have a similar level of English, regarding the mean scores of English mid-term and final exams, as the target students but were excluded from the study in order to make four treatment groups. To ensure the target students' absence of knowledge on the selected target words, students that already knew the meaning of more than two words were excluded from data analysis.

The researcher consulted with the teacher, who had taught the target students for two and half years, and selected eight target words from the thirty candidate words. Words were selected whose meanings did not overlap. Six nouns and two verbs were selected as target words to balance out the students' use of the words in context. Table 3.2 presents the selected words for the experiment.

Table 3.2

List of the Target Words

sacrifice	희생; 희생물
inquire	(…에게) 묻다
wander	거닐다, 돌아다니다
dispose	배치하다, 배열하다
firm	회사, 사무소
recognize	~ 알아보다[알다]
charity	자선[구호] 단체
proclaim	선언[선포]하다

3.4.2. Sample Sentences

In the research, two groups were presented with the target words using one of two different context conditions: diverse context sentences or the same context sentences. Students in the same context group were given only one sample sentence per target word during the task and practiced it three times whereas those in the diverse context group received three different sentences each time they were given a target word (see Appendix 2 for receptive task groups and Appendix 3 for productive task groups).

This sort of grouping aimed to save time to evaluate the quality and quantity of context clues, and represent sentence contexts in a more practical way. This methodology followed previous studies with a similar purpose (Bolger et al., 2008; Gass et al., 1999). The sample sentences were taken from various resources, including the Academic Word List (Coxhead, 2000), Naver online

concordances, and Your Dictionary web resources. The sentences were modified to balance the difficulty of context cues and the length of sentences.

Because different types of context cues are inclined to influence students' word learning differently (Drum & Konopak, 1987), the sample sentences should be chosen with great care to counterbalance the differential cognitive load required to acquire target vocabulary. In this study, the sentence contexts take focus, rather than the contextual clues. Therefore, each sample sentence was reviewed cautiously in order to avoid unexpected learning effects from other variables besides sentence contexts and repeated task.

Sample English sentences were used for all groups. The sentences were translated into Korean as L1 samples for the productive task. As mentioned earlier, the diverse context group was given three sentences per target word while the same context group was assigned a sentence that was repeated three times.

3.5. Treatment

During the task, the participants were given eight words that were unknown to them in the form of handouts and on screen projections. They were asked to practice using them in sentences and check their answers on the screen. A set of eight words were shown three times across two sessions.

3.5.1. The Receptive Task Groups

In the receptive task groups, the students practiced the target words in receptive ways. Students were guided to listen to and read the given sentences and translate L2 sentences (English) into L1 sentences (Korean) (see Appendix 2). At first, the learners were shown the form and sound of a target word. Then, the meaning of the word was given to them. In the following stage, the students were given a sample sentence that included the target word. They were given time to read the sentence, and they were advised to listen and repeat the sentence together. Subsequently, they were asked to translate the given L2 sentence into their L1 equivalent and then check the suggested answer on the screen.

Following the procedure stated above, the receptive groups carried out the treatment task using two different context conditions, as follows:

(1) Receptive Task + Same Context (RS): The students practiced the target words, found in the same sentences, three times each by reading and translating them (L2   L1).

(2) Receptive Task + Diverse Context (RD): The students practiced the target words, found in three different sentences, by reading and translating them (L2   L1).

3.5.2. The Productive Task Groups

In the productive task groups, the students completed the productive tasks by repeating the target words in sentences while writing and speaking (see Appendix 3). The learners were given the L1 meaning of the target word. Then, the target word for the given meaning was shown to the students on screen. They were asked to speak aloud all together. The researcher gave them time to write it down. An L1 sentence was given, and the learners were asked to translate it into the L2 sentence using the target word. After that, students were advised to check the suggested answer on the screen. They are asked to speak the sentence aloud together.

Following the procedure stated above, the productive groups carried out the treatment task using two different context conditions, as follows:

(1) Productive Task + Same Context (PS): The students practiced each target word three times using the same context sentences by writing an L2 sentence with the target word corresponding to the given L1 sentence.

(2) Productive Task + Diverse Context (PD): The students practiced the target words through the three different sentences by translating a given L1 sentence into an L2 sentence.   .

3.6. Assessment

To assess the students' knowledge of the target words, five types of vocabulary test were employed in this study—active word learning test, recognition test, passive word learning test, and two productive use tests: gap-filling, word ordering. These tests aimed to assess different developmental stages of the students' vocabulary knowledge. The vocabulary tests were intended to efficiently evaluate the students' ability to recognize the target word, retrieve its form and meaning, and use them in proper context (Nation & Gu, 2007; Yamashita, 2003). The tests weres also intended to evaluate the students' ability reorder the target word with good syntactic knowledge (Zwarts & Dras, 2007) with regards to the five specific stages of vocabulary knowledge: recognizing new words, getting their form and meanings, and using them in proper context and with appropriate grammar. The assessments were based on and revised from the developmental stages suggested in the VKS (Paribakht & Wesche, 1997) were also designed to better indicate the development of vocabulary knowledge.

This study was conducted to gauge exactly what students learn through vocabulary treatments by evaluating various parts of vocabulary knowledge development. Each vocabulary test has three different versions, in which the order of the target words was arranged in a different way to minimize the effect of repeated task at each time point. Each vocabulary test is described in detail in the following sections.

3.6.1. Active Word Learning Test

The active word learning test assesses the students' knowledge of the form of a target word associated with its meaning. In the test, the meaning of the target word was given, and the students were asked to retrieve the equivalent form. The active word learning test follows the format used in previous studies (Mondria & Wiersma, 2004; Webb, 2005, 2007, 2008). The active word learning test was the first test conducted. Appendix 4 and Appendix 5 illustrate sample active word learning test items for immediate and for delayed test each.

3.6.2. Recognition Test

The recognition test was designed to observe the gradual development of learners' vocabulary knowledge. Unlike other performance-based knowledge tests to assess learners' word knowledge through a given task, the recognition test adopted a self-report format to report the students' level of understanding of the words based on a scale. This arrangement was intended to measure the initial development of students' understanding of word form and meaning. It is expected that this test will serve as an appropriate tool to capture even partial or small progress in learners' knowledge, as in previous studies (Dale, 1965; Read, 2000; Paribakht & Wesche, 1997). The test intends to measure a learner's overall understanding of the target word on a modified version of the VKS

(Paribakht & Wesche, 1997), having learners answer on a four-point Likert-type scale of vocabulary knowledge (see Appendix 2 for immediate and Appendix 5 for delayed). The recognition test consists of eight test items.

3.6.3. Passive Word Learning Test

The passive word learning test measures the students' capability to retrieve the meaning of the target word, which contains the initial stage of vocabulary knowledge when a learner encounters a word (Nation & Gu, 2007). In contrast to the self-reported scaled recognition test, which has a similar purpose, this test aims to measure performance in a more direct way by letting them write down the meaning of the given word. Like the active word learning test, this test format follows that of previous studies (Webb, 2005, 2007, 2008) in which students translated the given L2 word without context into its L1 equivalent.

The recognition test and the passive word learning test were done concurrently to follow the revised format of the VKS from previous studies (Weinfurt, 2000). Appendix 2 and Appendix 5 illustrate a sample of the recognition test and the passive word learning test item.

3.6.4. Two Productive Use Tests: Gap-Filling and Word Reordering

The two productive word use tests measure how to retrieve an appropriate word in the given context (see Appendix 6 for immediate and Appendix 10 for delayed) and how to rearrange the given word clusters with appropriate syntactic and semantic meanings (see Appendix 7 for immediate and Appendix 11 for delayed). In the first test, incomplete sentences were given to students who had to complete them by filling in the proper words in the right context. Previous studies used this type of test as a retrieval cue to measure a learner's productive vocabulary knowledge (Laufer, 1998; Laufer & Rozovski-Roitblat, 2011). They adopted this test so as to evaluate learners' vocabulary use in context and monitor their vocabulary knowledge development that simple translation tests could not offer.

In the second test, a cluster of words, including the target word, were provided to the students to rearrange for meaning using proper grammar. The present study adopted this test from Zwarts & Darts's (2007) research to assess students' grammatical knowledge and observe aspects of vocabulary knowledge development that the direct translation test could not provide.

Some researchers criticize this sort of tests for not reflecting authentic aspects of the production procedure (Milton, 2009). They argue that a more genuine and suitable approach would employ essays to directly measure the students' productive vocabulary knowledge. Regardless of this criticism,

considering the English level of most Korean middle school students, it is almost impossible to make them write essays fluently. The practicality and feasibility of this test method cannot be ignored, and the present study has adopted it as one of the tests used.

3.6.5. Scoring

Two independent raters conducted scoring; the researcher of the present study, with one year of English teaching experience in high school, and another teacher, who has three years of experience teaching English in middle school. In particular, for the productive tests, scores for items showing disagreement were confirmed with a native English teacher. Excluding the self-reported recognition test, Pearson's r was calculated to check inter-rater reliability. The attained values were and 0.987 for the active word learning test, 0.979 for the passive word learning test, 0.999 for the Gap-Filling test and 0.986 for the word reordering test. Due to the very high inter-rater reliability, one of the rater's scores was randomly selected and included in the data analysis. The recognition test is a type of self-reporting measurement. The students' answers to the test items were scored according to the criteria shown in Table 3.3. These criteria are from a modified version of the VKS (Paribakht & Wesche, 1993)

Table 3.3

The Scoring Criteria for the Recognition Test

Score Knowledge Scale Description

0	I have never seen this word before, and I don't know this word at all.
1	I have seen this word before, but I don't know what it means.
2	I have seen this word before, and I think I partially know the meaning of the word.
3	I have seen this word before, and I know the meaning of this word.

The rest of the vocabulary tests are performance-based tests in which the students show their word knowledge by completing given test items. When scoring these tests, one aspect of the students' learning was considered: knowledge of the target form or meaning scored by asking learners to provide the form or meaning of the target word. A maximum of three points were assigned based on the following criteria in Table 3.4.

Table 3.4

Scoring Criteria for Performance-based Tests

Points	Criteria	Points by Criteria
Form/Meaning	correct	3
(3 points)	partially correct	1.5
	wrong	0

For form/meaning correctness, the points given to each answer ranged from 1.5 to 3, graded as shown in Table 3.4. Three points were given for the right answer and one and a half points for a partial or near-right answer. Awarding partial points makes it possible to be more sensitive to incomplete but still meaningful knowledge (Waring & Takaki, 2003).

Recognition Test

In the recognition test, students reported their vocabulary knowledge themselves according to a given scale, shown in Table 3.4. The score for each test item ranges from 0 to 3, so the maximum score for the full test is 24 (8×3).

Passive Word Learning Test

Two Korean English teachers scored the passive word learning tests according to the rubric for performance-based tests (see Table 3.4). Partial points were given to answers including a meaning semantically close to the right

answer, a decision made by the two English experts. The score for each test item ranges from 0 to 3, so the maximum score for the full test is 24 (8×3).

Active Word Learning Test

The active word learning tests scored by two Korean EFL teachers with assistance from one native-English-speaking teacher consistent with the rubric for performance-based tests (see Table 3.4). Partial points were given for spelling errors that did not distort the sounds of words. In particular, they got partial scores with 1) more than half of correct syllables or letters, 2) the correct consonant clusters with wrong vowels or 3) the correct vowels with reverted consonants. For example, a student who misspelled wander as wonder was given 1.5 points for the answer. The score for each test item ranges from 0 to 3, so the maximum score for the full test is 24 (8×3).

Productive Use Tests

Two Korean EFL teachers mentioned above scored the productive use tests. No partial points were awarded for the gap-filling test. In the word reordering test, partial points were given only if students put the target word in the right position but had a mistake in placing the remaining parts. The maximum total score for all types of tests used in this study was 24 (8×3) respectively.

3.7. Data Analysis

To adequately explore the research questions, analysis was conducted on each of the four main study groups to investigate the effects of task types and sentence contexts on the sum of five vocabulary test scores and individual test scores. Statistical analysis was implemented using SPSS for Windows (v. 22.0) to verify the research questions; how task types and sentence contexts involve vocabulary gains and retention while vocabulary knowledge development.

First, a set of two-way analysis of variance (ANOVA) was employed with task types and sentence contexts as independent variables and the total vocabulary test scores as a dependent variable. Univariate between-group analysis was followed. Second, a set of two-way Multivariate Analysis of Variance (MANOVA) was employed with task types and sentence contexts as independent variables and the five types of vocabulary tests as dependent variables. Univariate between-group analysis was followed.

CHAPTER 4.
RESULTS AND DISCUSSION

This chapter describes the results of the statistical analysis of the test scores and discusses the findings. Section 4.1 reports the sum of the five test scores and a discussion of the immediate post-test and delayed post-test depending on the assigned task type and sentence contexts. The effect of these two factors on the specific outcomes of each test item and their discussion of the immediate post-test and the delayed post-test are described in Section 4.2.

4.1. The Effects of Task Type and Sentence Contexts on the Overall Immediate Vocabulary Learning and Retention

To investigate the impact of the task type and sentence contexts on general vocabulary gains and its retention, the participating students were divided into one of four treatment groups. Each group was assigned one of two task types and one of two contexts (RS, RD, PS, PD) and all groups took the immediate post-test and delayed test.

To analyze the effects of the task type and sentence contexts on the overall vocabulary learning, the sum of five vocabulary test scores were analyzed. Table 4.1 summarizes the overall means and the standard deviations of the test scores,

according to the four groupings with a combination of different task types and contexts. The data is then represented in chart format in Figure 4.1. .

Table 4.1

Descriptive Statistics of Immediate Test

Task	Context	Mean	Std. Deviation	N
	Same	76.6500	27.04216	30
Receptive	Diverse	75.2759	30.73928	29
	Total	75.9746	28.67542	59
	Same	97.2333	26.29925	30
Productive	Diverse	85.7143	31.34794	28
	Total	91.6724	29.17349	58
	Same	86.9417	28.40985	60
Total	Diverse	80.4035	31.20846	57
	Total	83.7564	29.85768	117

Note. The scores are the sum of five vocabulary test scores; the maximum score is 120.

Estimated Marginal Means of Whole Test Score

Figure 4.1

Overall Test Scores by Task and Context on Immediate Test

As shown in Table 4.1 and Figure 4.1, the productive groups had significant superiority over the receptive groups for vocabulary learning in the immediate tests. In sentence contexts, however, although the same context groups showed slightly higher mean scores than the diverse context ones. No observable difference was noticed between the two different context groups. Regarding Figure 4.1, no interaction between the two variables was represented. Therefore, in the immediate test, the PS and PD groups were ahead of RD and RS groups but differences between the same task groups were marginal. In particular, the mean scores of the two receptive groups were practically the same.

In order to verify the statistically significant differences between each variable, a set of two-way analysis of variance (ANOVA) were conducted. As seen in Table 4.2, below, the task type, not the context, had significant main effects on the initial learning of the target vocabulary.

Table 4.2

Effects of Task and Context on the Immediate Test

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	7031.443	1	7031.443	8.430	.004	.069
Context	1214.597	1	1214.597	1.456	.230	.013
Task * Context	751.983	1	751.983	.902	.344	.008
Error	94254.949	113	834.115			

*p<.05

In the immediate test score analysis, the primary impact of the task was shown, $F (1, 115) = 8.430$, $p = .004$, $\eta2= .069$, but the statistically meaningful impact of sentence contexts was not, $F (1, 115) = 1.425$, $p = .235$, $\eta2= .230$. There was no significant interaction effect between the task and the context in

the immediate test, F (1, 115) = .902, p = .344,η2= .008.

According to the result of the immediate test, the productive groups had are remarkable superiority over the receptive groups for vocabulary learning in the immediate tests. In sentence contexts, however, no observable difference was noticed between the two different context groups. Thus, the task type, and not sentence contexts, influenced immediate word gain. To be specific, the productive task is more beneficial for immediate vocabulary learning than the receptive task but whether the task is provided in the same context or in diverse contexts does not meaningfully affect immediate vocabulary learning.

The productive groups outperformed the receptive groups in vocabulary retention as well, as shown below in Table 4.3 and Figure 4.2. Compared to the immediate word gain, the productive task proved more beneficial when it comes to maintaining words. No observable difference was shown between the two different context groups, but the interaction between the two variables is shown in Figure 4.2. This means that sentence contexts may affect vocabulary retention differently depending on the type of the task.

Table 4.3

Descriptive Statistics of Delayed Test

Task	Context	Mean	Std. Deviation	N
	Same	52.2667	23.91246	30
Receptive	Diverse	56.1724	33.13912	29
	Total	54.1864	28.63472	59
	Same	89.2500	28.73054	30
Productive	Diverse	72.3393	31.44780	28
	Total	81.0862	31.00129	58
	Same	70.7583	32.16397	60
Total	Diverse	64.1140	33.05164	57
	Total	67.5214	32.62934	117

Note. The scores are the sum of five vocabulary test scores; the maximum score is 120

Figure 4.2

Overall Test Scores by Task and Context on Delayed Test

Looking at the delayed test scores, the PS group was still ahead of others, followed by PD (see Figure 4.3 and Table 4.3). The same context affected vocabulary learning more positively than diverse contexts during the productive task session. The context effect, however, were reversed when students did the receptive task; the RD outperformed the RS for word retention. Overall, the PS task had more durability of overall vocabulary knowledge than the PD task. On the other hand, the RD task had stronger durability than the RS task in the delayed posttest. That is, the sentence contexts conjugating the task type may not sufficiently influence immediate word gain but may influence its retention

Estimated Marginal Means of Vocabulary_Test

*cf. TC= Task type * Context*

Figure 4.3

Overall Test Scores by TC Immediate and Delayed Test

Table 4.4

Effects of Task and Context on the Delayed Test

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	20640.565	1	20640.565	23.807	.000	.174
Context	1235.749	1	1235.749	1.425	.235	.012
Task * Context	3166.106	1	3166.106	3.652	.059	.031
Error	97971.906	113	867.008			

*p<.05

The statistical significance of the test result differences was checked through a set of two-way analysis of variance(ANOVA).The task type had the main effects: F (2, 114) = 30.52, p = 0.000,η2 =.174, but context did not. The p-value of the interaction effect was .059, which did not meet the statistically meaningful level (p<.05), but showed a certain power of the sentence contexts on learners' word retention.

To sum up, task type was a factor that significantly affected vocabulary learning, and this ultimately shows support for previous studies (Hulstijn & Laufer, 2001; Kim, 2013; Webb, 2005) that argued the superiority of the productive task over the receptive task on vocabulary knowledge development.

The productive task group produced higher overall scores than the receptive task group in both the immediate and the one-week delayed test. In the delayed test, the overall test score of the productive task groups decreased much less than those of the receptive task groups, which indicates that productive tasks have greater potential to help retain word knowledge.

Sentence contexts, on the other hand, were not a statically significant factor affecting vocabulary learning throughout the experiment. Even though no statistically meaningful differences were found between the groups, the descriptive statistics show that, when compared to the diverse context groups, the same context groups had a slightly higher mean score overall. This result is in contradiction to those of Bolger et al.'s (2008) and An and Min's (2014) study, which proposed the superiority of diverse contexts over the same context in vocabulary learning.

Regarding word retention, the interaction effect between two variables was shown. In delayed test, like the immediate test, the PS group was ahead of other groups. However, the scores of receptive groups significantly decreased. The result of the RS group, in particular, saw a very large decrease, so this group demonstrated the lowest ability to retain word knowledge. That is, the same context had a positive effect on the productive task but not on the receptive task.

The findings from overall test scores revealed that, depending on the assigned task type, effective context might differ. This suggests that the productive task was more demanding to the learners so that it made them focus more on the vocabulary itself when they repeatedly wrote the same sentence

rather than writing down different sentences. The receptive task, on the other hand, was less challenging, so students could benefit from several context sentences, with their focus on the target vocabulary itself.

4.2. The Effects of Task Type and Sentence contexts on the Immediate Learning and Retention of Specific Vocabulary Knowledge

In this section, the scores of the recognition, passive word learning, active word learning, gap-filling and word reordering tests are treated as five dependent variables and analyzed to investigate whether the task type and sentence contexts affect them differently.

The results of each vocabulary test are explained in greater detail in Sections 4.2.1 to 4.2.5.

4.2.1. Recognition Test

Table 4.5, Table 4.6 and Figure 4.4 represent the descriptive statistics of the recognition test results in the immediate and delayed post-test. The productive task groups showed slightly higher mean scores than the receptive groups on the immediate test, but the difference was negligible (see Table 4.5). Concerning the

sentence contexts, the difference between the same context group and diverse context group was marginal, and the mean scores of the groups were practically the same. Regarding the combination of the two factors (see Table 4.5 and Figure 4.4), even though the PS group had slightly higher mean scores than the other groups, no observable difference was noticed among the four different treatment groups for immediate word gain. Overall, the effect of task type and sentence contexts on the immediate gain of word recognition was subsidiary.

Table 4.5

Descriptive Statistics of the Immediate Recognition Test

Task	Context	Mean	Std. Deviation	N
	Same	19.2000	5.47345	30
Receptive	Diverse	19.5172	5.77322	29
	Total	19.3559	5.57631	59
	Same	21.3333	3.57514	30
Productive	Diverse	20.0357	5.70563	28
	Total	20.7069	4.72770	58
	Same	20.2667	4.70797	60
Total	Diverse	19.7719	5.69466	57
	Total	20.0256	5.19526	117

Note. The maximum score for each test is 24.

Figure 4.4

Recognition Test Scores by TC on Immediate and Delayed Tests

As shown below in Table 4.6, the productive task groups were more capable of recognizing the target words than the receptive groups. Concerning sentence contexts, the difference between the delayed test scores of the same context groups and those of diverse context groups was trivial. As represented in Table 4.6 and Figure 4.4, the scores of the PS and PD groups outperformed RD and RS groups, but no observable difference was noticed among the same task groups. That is, the task type affects retaining word recognition knowledge, not the sentence contexts

Table 4.6

Descriptive Statistics of the Delayed Recognition Test

Task	Context	Mean	Std. Deviation	N
	Same	15.7000	5.71839	30
Receptive	Diverse	16.7586	6.68503	29
	Total	16.2203	6.18136	59
	Same	20.5667	3.77545	30
Productive	Diverse	18.5000	5.88469	28
	Total	19.5690	4.97401	58
	Same	18.1333	5.39449	60
Total	Diverse	17.6140	6.30973	57
	Total	17.8803	5.83858	117

Note. The maximum score for each test is 24.

The statistical significance of the test result differences was checked through a set of two-way MANOVA tests, shown below in Table 4.7 and Table 4.8. The results of the recognition test scores in Table 4.7 revealed that neither task nor context, served as a between-subjects variable, had statistically evocative influence over the immediate recognition test scores (Task, $F(5,109) = 1.903$, p $= .171$, $\eta2= .017$; Context, $F(5,109) = .260$, p $= .611$, $\eta2= .002$).

However, in the delayed test, shown in Table 4.8, the effect of task type was statistically meaningful; $F(5,109) = 10.161$, p $= .002$, $\eta2= .083$, but still no

meaningful effect of context was observed; $F(5,109) = .236$, $p = .628$, $\eta2= .002$. As to interaction, there was no significant interaction effect in both tests (immediate, $F(5,109) = .706$, $p = .403$, $\eta2= .006$; delayed, $F(5,109) = 2.273$, $p = .134$, $\eta2= .020$).

Table 4.7

Effect of Task and Context on Immediate Recognition Test

Effect	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	51.380	1	51.380	1.903	.171	.017
Context	7.023	1	7.023	.260	.611	.002
Task * Context	19.054	1	19.054	.706	.403	.006

p<.o5

Table 4.8

Effect of Task and Context on Delayed Recognition Test

Effect	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	319.049	1	319.049	10.161	.002	.083
Context	7.425	1	7.425	.236	.628	.002
Task * Context	71.366	1	71.366	2.273	.134	.020

p<.o5

Because word recognition ability is the initial stage of vocabulary knowledge, according to the VKS (Paribakht & Wesche, 1993), no differences were observed depending on difference in task types and context. As for retention for this knowledge, however, the productive task had more durability than the receptive one (Hulstijin & Laufer, 2001).

4.2.2. Passive Word Learning Test

Table 4.9, Table 4.10 and Figure 4.5 show the descriptive statistics of the passive word learning test scores in the immediate test and the delayed test.

Table 4.9

Descriptive Statistics of the Immediate Passive Word Learning Test

Task	Context	Mean	Std. Deviation	N
Receptive	Same	17.5000	7.35199	30
	Diverse	17.4828	7.00444	29
	Total	17.4915	7.12118	59
Productive	Same	19.8000	5.37812	30
	Diverse	17.4643	7.07359	28
	Total	18.6724	6.30897	58
Total	Same	18.6500	6.49074	60
	Diverse	17.4737	6.97535	57
	Total	18.0769	6.72797	117

Note. The maximum score for each test is 24.

Regarding the task type, the productive groups received slightly higher passive word learning test scores than the receptive groups on the immediate test, but the difference was marginal. Sentence contexts also showed minimal difference between the same context groups and the diverse context groups. Table 4.9 and Figure 4.5 show that the PS group was a little ahead of others and no apparent mean differences across the remaining three groups were shown. That is, task type and sentence contexts do not affect immediate retrieval of word meaning.

Figure 4.5

Passive Word Learning Test Scores by TC on Immediate and Delayed Tests

Table 4.10

Descriptive Statistics of the Delayed Passive Word Learning Test

Task	Context	Mean	Std. Deviation	N
Receptive	Same	10.8167	6.12300	30
	Diverse	12.5690	8.35563	29
	Total	11.6780	7.29594	59
Productive	Same	18.4000	6.28956	30
	Diverse	15.0000	6.90411	28
	Total	16.7586	6.75592	58
Total	Same	14.6083	7.24516	60
	Diverse	13.7632	7.70674	57
	Total	14.1966	7.45329	117

Note. The maximum score for each test is 24.

In the delayed test, the outcomes turned out to be different. The productive task groups outperformed the receptive groups with considerable gaps. Meanwhile, sentence contexts did not represent meaningful differences among the two different context groups. In the delayed post-test shown in Table 4.10 and Figure 4.5, the PS group kept ahead of all of the other groups. Unlike the immediate posttest, there were apparent mean differences across the remaining groups. The mean score of the PS group decreased much less than other groups. Receptive task groups, however, especially the RS group, showed a sharper

decline amongst the four groups.

When univariate analyses were conducted to probe whether the differences between groups were statistically meaningful, the results showed that neither task type $(F(5,109) = .838, p = .362, \eta2 = .007)$ nor context $(F(5,109) = .891, p = .347, \eta2 = .008)$ had a significant main effect on the initial word learning. This data is shown below in Table 4.11.

Table 4.11

Effect of Task and Context on Immediate Passive Word Learning Test

Effect	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	38.033	1	38.033	.838	.362	.007
Context	40.452	1	40.452	.891	.347	.008
Task * Context	39.275	1	39.275	.865	.354	.008
p<.o5						

In the delayed test score analysis, the main effect of the task was shown, $F(5,109) = 15.120, p = .000, \eta2 = .118$, but no statistically meaningful effect of sentence contexts was shown, $F(5,109) = 19.837, p = .524, \eta2 = .004$. These results are shown below in Table 4.12. As for interaction, there was an

interaction effect but only in the delayed test ($F(5,109) = 4.002$, $p = .048$, $\eta2= .034$). That is, the sentence contexts worked differently corresponding to the types of tasks. The same context functioned positively on the productive task while it worked negatively on the receptive task. Thus, the receptive task with the same context had a weak power to retain vocabulary knowledge, especially inferring the meaning of the target word.

Table 4.12

Effect of Task and Context on Delayed Passive Word Learning Test

Effect	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	732.755	1	732.755	15.120	.000	.118
Context	19.837	1	19.837	.409	.524	.004
Task * Context	193.961	1	193.961	4.002	.048	.034

p<.o5

Contrary to Waring's (1997) argument that the receptive task outperformed the productive one when assessing receptive vocabulary knowledge, there was no difference in passive word knowledge gain through two different types of the task. Furthermore, regarding retention of passive word knowledge, the

productive task showed more potential than the receptive task. The results can be supported by the level of vocabulary knowledge suggested by the VKS (Paribakht & Wesche, 1997). The productive task is dealing with deeper word process and had a better impact on retention of receptive word knowledge

4.2.3. Active Word Learning Test

Table 4.13, Table 4.14 and Figure 4.6 display the descriptive statistics of active word learning test scores in immediate and delayed post-tests. The productive groups showed a considerably higher mean score than the receptive groups for both the immediate word learning (Table 4.13) and its retention (Table 4.14) in the active word learning. This test aimed at retrieving word forms, which were included in productive vocabulary knowledge, so it is doubtless that the productive task was more beneficial than the receptive task in immediate word knowledge gain and its retention. Regarding sentence contexts, no observable difference was noticed between the two different context groups in the immediate post-test.

The results of the active word learning tests among four treatment groups are displayed in Table 4.13, Table 4.14 and Figure 4.6. The PS group consistently earned the highest mean score, and the PD, RD and RS groups followed in either the immediate or the delayed posttest. This proved that there were adverse effects of sentence contexts on task types over the mean score of the active word

learning test. Even though learners repeatedly practiced the given task, the PS significantly outperformed the other groups, while, the RS group received the lowest active word learning test score.

Table 4.13

Descriptive Statistics of the Immediate Active Word Learning Test

Task	Context	Mean	Std. Deviation	N
Receptive	Same	6.3000	4.44623	30
	Diverse	9.9310	7.97174	29
	Total	8.0847	6.62685	59
Productive	Same	19.6000	6.06346	30
	Diverse	16.9286	7.95673	28
	Total	18.3103	7.10685	58
Total	Same	12.9500	8.52996	60
	Diverse	13.3684	8.64608	57
	Total	13.1538	8.55219	117

Note. The maximum score for each test is 24.

Table 4.14

Descriptive Statistics of Delayed Active Word Learning Test

Task	Context	Mean	Std. Deviation	N
Receptive	Same	2.0500	2.34649	30
	Diverse	6.5690	6.97153	29
	Total	4.2712	5.60428	59
Productive	Same	17.2500	6.71687	30
	Diverse	13.7679	8.43656	28
	Total	15.5690	7.72978	58
Total	Same	9.6500	9.14446	60
	Diverse	10.1053	8.47356	57
	Total	9.8718	8.78899	117

Note. The maximum score for each test is 24.

Figure 4.6

Recognition Test Scores by TC on Immediate and Delayed Tests

The statistical significance of differences in the test results were checked through a set of two-way multivariate analysis of variance (MANOVA). The task was found to have statistically significant impacts on both the immediate test ($(5,109) = 66.330$, $p = .000$, $\eta 2 = .370$), as shown in Table 4.15 and the delayed test ($(5,109) = 87.195$, $p = .000$, $\eta 2 = .436$) as represented in Table 4.16. In terms of context, on the other hand, there was no observable impact on both tests. As to interaction, there was interaction effect in both tests (immediate, $F(5,109) = 6.395$, $p = .013$, $\eta 2 = .054$; delayed, $F(5,109) = 11.126$, $p = .001$, $\eta 2 = .090$).

Table 4.15

Effect of Task and Context on

the Immediate Active Word Learning Test

Effect	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	3010.221	1	3010.221	66.330	.000	.370
Context	6.728	1	6.728	.148	.701	.001
Task * Context	290.223	1	290.223	6.395	.013	.054

p<.o5

Table 4.16

Effect of Task and Context on

the Delayed Active Word Learning Test

Effect	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	3665.766	1	3665.766	87.195	.000	.436
Context	7.855	1	7.855	.187	.666	.002
Task * Context	467.748	1	467.748	11.126	.001	.090

p<.o5

Overall, regarding the form retrieval of target words, repeating several context sentences seems to be more efficient with the productive task, in comparison to repeating the same context sentences with the receptive task. As the RS group was the least effective, repeatedly reading new words in the same context sentences does not seem to be a practical way to learn the word forms. Consequently, the results suggest that the productive task was more challenging for the students so it made them focus more on vocabulary spelling itself when repeatedly writing the same sentence rather than writing down different sentences. The receptive task, on the other hand, was less demanding, so using several context sentences helped them retrieve vocabulary from its corresponding meaning.

4.2.4. Two Productive Use Tests

This section shows the results and discussion of the two productive use tests conducted in this study. The result and discussion of the gap-filling test and word reordering test are described.

4.2.4.1. Gap-Filling Test

The descriptive statistics of the independent variables for the gap-filling test were shown below in Table 4.17 and Table 4.18. When it comes to either the

task type or the sentence contexts, no observable differences were shown between the two groups per each variable regarding immediate word gain (see Table 4.17). The gap of mean scores between groups was not apparent on the immediate test.

As shown in Figure 4.7, the PS group achieved the highest average score, followed by the RS, PD and RD groups on the immediate test. The difference between the RS and PD groups was trivial and the mean scores of the two groups were the same. Neither the type of the task nor the sentence contexts meaningfully affected the immediate use of words in the proper context sentence.

Table 4.17

Descriptive Statistics of the Immediate Gap-Filling Test

Task	Context	Mean	Std. Deviation	N
Receptive	Same	14.4500	7.70395	30
	Diverse	12.5172	8.63804	29
	Total	13.5000	8.16373	59
Productive	Same	17.1000	7.31248	30
	Diverse	14.1429	8.75051	28
	Total	15.6724	8.10542	58
Total	Same	15.7750	7.56577	60
	Diverse	13.3158	8.65439	57
	Total	14.5769	8.17287	117

Note. The maximum score for each test is 24.

Table 4.18

Descriptive Statistics of the Delayed Gap-Filling Test

Task	Context	Mean	Std. Deviation	N
Receptive	Same	9.1500	6.34381	30
	Diverse	8.0690	8.14130	29
	Total	8.6186	7.23994	59
Productive	Same	15.8833	7.37152	30
	Diverse	10.7143	7.84978	28
	Total	13.3879	7.97641	58
Total	Same	12.5167	7.61687	60
	Diverse	9.3684	8.03926	57
	Total	10.9829	7.95026	117

Note. The maximum score for each test is 24.

Figure 4.7

Gap-Filling Test Scores by TC on Immediate and Delayed Tests

The In the delayed test, in contrast, there was a considerable effect of both the type of the task and the variable of the context on word retention (see Table 4.18). The test scores of the productive groups significantly exceeded those of the receptive groups. This means the productive group showed more statistically meaningful durability than the receptive group for word knowledge regarding its productive use in the proper context. It implies that the receptive task made the context information of vocabulary harder to recall from memory after one week. In terms of sentence contexts, the same context group outperformed the diverse context group on both tests. The gap between groups was apparent in delayed

post-test.

To be brief, the productive task using the same context sentences was more efficient at helping students retrieve word forms and use them in context in comparison to other treatment. In mastering target words for use in relevant contexts, it seems that practicing the words productively through writing activities repeated in the same sentence is useful in the overall learning process.

Table 4.19 and Table 4.20 show the univariate results of a two-way MANOVA in the immediate test and delayed test. In regard to task type and sentence contexts, the differences were not statistically significant (Task, $F(5,109) = 2.032$, p = .157, $\eta2 = .018$; Context, $F(5,109) = 2.658$, p = .106, $\eta2 = .023$) in the immediate test. In the delayed test, however, both task type and sentence contexts had statistically meaningful effects on the test scores (Task, $F(5,109) = 11.596$, p = .001, $\eta2 = .093$; Context, $F(5,109) = 5.150$, p = .025, $\eta2 = .044$).

Table 4.19

Effect of Task and Context on Immediate Gap-Filling Test

Effect	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	133.570	1	133.570	2.032	.157	.018
Context	174.708	1	174.708	2.658	.106	.023
Task * Context	7.667	1	7.667	.117	.733	.001

*p<.o5

Table 4.20

Effect of Task and Context on Delayed Gap-Filling Test

Effect	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	642.677	1	642.677	11.596	.001	.093
Context	285.419	1	285.419	5.150	.025	.044
Task * Context	122.106	1	122.106	2.203	.141	.019

*p<.o5

Regarding the VKS (Paribakht & Wesche, 1997), the productive use of vocabulary was the most difficult part of learning vocabulary knowledge, so it is more helpful to conduct a deeper level task, which refers to the productive task. The results about context effect in this study yielded conflicting results from An and Min's (2014) previous study. The same context groups, and not the diverse context groups, had statistically meaningful effect on vocabulary knowledge regarding its contextual use. It is because students may focus on its contextual use better when the same context sentences were used repeatedly rather than when the diverse context sentences, using too many cues, were given (Hu & Nation, 2012; Nation & Coady, 1988).

4.2.4.2. Word Reordering Test

Table 4.21 and Table 4.22 show the descriptive statistics of the task type and context for the word reordering tests for immediate word gain and its retention. Regarding the task type, the productive group showed slightly higher test scores than the receptive group but the gap between the two task groups was minimal on the immediate test (see Table 4.21). When it comes to sentence contexts, however, the same context group considerably outperformed the diverse context group for immediate gain of the grammatical use of vocabulary.

Table 4.21

Descriptive Statistics of Immediate Word Reordering Test

Task	Context	Mean	Std. Deviation	N
Receptive	Same	19.2000	5.54853	30
	Diverse	15.8276	7.05372	29
	Total	17.5424	6.50417	59
Productive	Same	19.4000	6.24003	30
	Diverse	17.1429	6.22399	28
	Total	18.3103	6.28129	58
Total	Same	19.3000	5.85503	60
	Diverse	16.4737	6.63286	57
	Total	17.9231	6.37871	117

Note. The maximum score for each test is 24.

Table 4.22

Descriptive Statistics of Delayed Word Reordering Test

Task	Context	Mean	Std. Deviation	N
Receptive	Same	14.5500	6.58152	30
	Diverse	12.2069	7.60347	29
	Total	13.3983	7.13888	59
Productive	Same	17.1500	6.84200	30
	Diverse	14.3571	6.90928	28
	Total	15.8017	6.95786	58
Total	Same	15.8500	6.78377	60
	Diverse	13.2632	7.28692	57
	Total	14.5897	7.12227	117

Note. The maximum score for each test is 24.

In the immediate test, the mean scores of the PS and RS groups and those of the PD and RD groups were practically the same (see Figure 4.8). The outcome shows that the context, not the task type, has a significant effect on the test scores on the immediate test. Both task groups with the same context performed better than those with diverse context.

As for retention of lexical knowledge measured by the word reordering test, the productive group showed relatively higher test scores than the receptive group (see Table 4.22). Regarding the sentence contexts, the same context

groups significantly surpassed the diverse context groups for retention of the productive use of vocabulary

Figure 4.8

Word Reordering Test Scores by TC on Immediate and Delayed Tests

Figure 4.8 shows that the test scores of the receptive groups declined sharply on the delayed post-test. The PS group was still ahead of the other three groups, but the gap between the PS and RS groups significantly expanded. Although the mean score of the RS group was much higher than that of the PD group in the immediate test, there were no visible differences in the mean scores between the two groups on the delayed test.

The univariate results of the word reordering post-tests shown in Table 4.23 prove that the differences between groups were statistically significant in regard to sentence contexts $(F(5,109) = 5.865, p = .017, \eta2= .049)$, but not task type $(F(5,109) = .425, p = .516, \eta2= .004)$, as a between-subjects variable on the immediate test. In addition, there was no significant interaction effect between the task and the context in the immediate test $(F(5,109) = .230, p = .632, \eta2= .002)$.

Table 4.23

Effect of Task and Context on Immediate Word Reordering Test

Effect	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	16.776	1	16.776	.425	.516	.004
Context	231.558	1	231.558	5.865	.017	.049
Task * Context	9.088	1	9.088	.230	.632	.002

p<.o5

Although, in the delayed post-test, task type $(F(5,109) = 3.374, p = .069, \eta2= .029)$ had a statistically meaningful effect on test scores, but it had a very limited effect (see Table 4.24). Sentence contexts $(F(5,109) = 3.944, p = .049,$

η2= .034), on the other hand, had a statistically meaningful effect on test scores. In addition, there was no significant interaction effect between the task and the context in the delayed test (F(5,109) = .030, p = .862, η2= .000).

Table 4.24

Effect of Task and Context on Delayed Word Reordering Test

Effect	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	164.871	1	164.871	3.374	.069	.029
Context	192.733	1	192.733	3.944	.049	.034
Task * Context	1.478	1	1.478	.030	.862	.000

p<.o5

According to the VKS (Paribakht & Wesche, 1997), the productive use of vocabulary knowledge was the most difficult stage, so it is more helpful to conduct the productive task rather than the receptive task. This study, however, failed to prove the positive effect of the productive task in either word gain or its retention. However, comparing the effects of task on word gain, those on its retention was much more dominant. In terms of context effect, the same context groups had a statistically meaningful effect on vocabulary knowledge regarding

its use in both immediate and delayed tests, which is opposed to An and Min's (2014) findings. It is because students may focus on its productive use better when they repeatedly read the same context rather than read the different sentences all the time.

CHAPTER 5.
CONCLUSION

This chapter is composed of three sections. Section 5.1 summarizes the findings of the present study. The pedagogical implications of this study regarding English vocabulary education are discussed in Section 5.2. Finally, Section 5.3 describes the limitations of the present study and makes suggestions for the further research

5.1. Major Findings

This study investigated how task type (receptive versus productive) and sentence contexts (the same context versus diverse contexts) contribute to lexical knowledge development of Korean middle school students.

The first research question looked into the effect of receptive versus productive task and sentence contexts on overall vocabulary learning and retention. The impact of the each task and sentence contexts on the five specific components of vocabulary knowledge development were investigated in the second research question.

Task type was a factor that significantly affected vocabulary learning, and this ultimately show support for the previous studies (Hulstijn & Laufer, 2001; Kim, 2013; Webb, 2005) that argued the superiority of the productive task over the receptive task for developing vocabulary knowledge. The productive task

group produced higher overall scores compared to the receptive task group in both the immediate and the one-week delayed test. In the delayed test, the overall test scores of the productive task group decreased much less than those of the receptive task group, which indicates that productive task had more durability to retain word knowledge (Webb, 2005).

Sentence contexts themselves were not a statistically significant factor affecting vocabulary learning throughout this experiment. However, they played a crucial role when interacting with the type of task, especially in word retention. In other words, the more effective context might differ in relation to the assigned task types in this study. The same context groups were positively affected on the productive task but not on the receptive task. Thus, the PS group showed predominance in word gain and a much more statistically significant power in word retention among the four treatment groups. The RS group, on the other hand, revealed the lowest ability to retain word knowledge.

Since the productive task was more demanding for the students, it made them focus more on vocabulary itself when repeatedly writing the same sentence rather than writing down different sentences. The receptive task, on the other hand, was less challenging, so students could benefit from several context sentences and focus on the target vocabulary itself. However, this study only investigated the gain and retention of target words. Regarding additional word gain, the multiple context groups may have superiority over the same context groups.

The findings from the individual analysis of the five vocabulary tests

demonstrated that the productive task had significant superiority over the receptive task in vocabulary learning. In particular, the productive task had statistically considerable power to retain several phases of vocabulary knowledge, with the exception of the word reordering test. The context itself did not have much influence on the lexical knowledge development from word recognition to passive and active word knowledge. When combined with the task, the sentence contexts had a strong effect on vocabulary learning in passive and active word learning tests, especially in word form and meaning extraction. In line with the overall findings, the same context groups, and not diverse contexts, had a statistically meaningful effect on vocabulary knowledge. Thus, the PS group scored the highest among the four treatment groups in the five types of tests respectively. On the other hand, with the exception of the two productive use tests, the RS group scored the lowest. The results explain the cross effect between task type and context for word recognition and retrieval of word meaning and form.

The retention of word knowledge measured by the productive use of vocabulary tests, on the other hand, was influenced fundamentally by sentence contexts rather than task type. Conflicting with results from a previous study (An & Min, 2014), the same context groups had a statistically meaningful effect on retention of the vocabulary knowledge regarding its contextual use. Using the same context sentence may help students focus more on the words contextual use compared with the diverse context sentences that demands higher cognitive loads.

In general, considering the effect of task type and context on overall

vocabulary learning, the task effect was substantial, while that of context was not. Examining the two variables together, however, shows that the effect of context was different from the task types, especially for retrieving word and meaning connection. The productive task, when completed within the same context, was always ahead of other treatment groups. The receptive task, when completed within the same context, usually recorded the lowest grade with the exception of the two productive use tests. However, in the productive use tests, which demand contextual knowledge of target vocabulary, the using the same context lead to better results since it helped learners focus on contextual information of the target words and was not strongly related to the task that was done.

5.2. Pedagogical Implications

Based on the major findings described in section 5.1, this study presents the following pedagogical implications on L2 vocabulary learning.

1) Implementing more productive tasks than receptive ones in the classroom context may be effective for vocabulary learning, especially regarding its retention. This is because it enables learners to gain and retain much more productive vocabulary knowledge as well as a little more or at least a similar level of receptive vocabulary knowledge.

2) Rather than just assuming vocabulary instruction through diverse contexts is always the most effective, the effectiveness of the context should be carefully judged based on other variables such as task types, students' English proficiency and so on. This study proposed the possibility of interaction effect between task types and sentence contexts on vocabulary retention. Teachers should consider the cognitive load and difficulty level of each sentence contexts of target vocabulary before designing, modifying, comparing, choosing, or implementing vocabulary tasks.

5.3. Limitations and Suggestions

First, this research was conducted with 117 Korean middle school students living in Gwanak-gu, Seoul, which makes it difficult to generalize the major findings for a larger population. The effect of vocabulary treatment may fluctuate according to students' age, their residence, their average language ability, or their motivations to learn English. Further research is suggested to employ a sufficient number of students from diverse backgrounds, randomly sampled for multiple variables so the finings can be more applicable to a larger population.

Second, this study did not take students' proficiency levels into account. Since the cognitive load of vocabulary task treatment may affect research results, different results could be revealed in the effects of task type and sentence contexts on vocabulary learning if students are classified into different

proficiency groups.

Third, the word items utilized in the vocabulary task treatment were limited in number, level and parts of speech. Only eight target words, including six nouns and two verbs, were chosen out of thirty-word items in the Lv 1000 and Lv 2000 word list (Academic Word List, Coxhead, 2000), but the limitation of word selection made it difficult to generalize the significant findings. Therefore, future studies should contain a larger and more diverse list of words from different levels using different parts of speech.

Fourth, the number of sentence contexts in the vocabulary task treatment was also limited. Compared with task effects, the effect of sentence contexts was relatively marginal, which can be derived from the limited number of sentence contexts. In addition, the multiple context groups failed to show its superiority when performing productive task. The result can differ with sufficient number of sentence contexts. In future studies, it would be beneficial to provide more context sentences during vocabulary instruction to determine if context will have a larger effect.

Fifth, there was a problem in assessing the productive use of word knowledge. As an alternative to the free composition evaluating the productive use of word knowledge in the last phase of the VKS (Paribakht & Wesche, 1997), this study implemented two productive use tests: gap-filling and word rearrangement. However, there remains a limitation because these assessments guarantee to evaluate students' productive use of word knowledge using the proper context. Moreover, in the assessment of the word rearrangement test, a

partial point was given when the target word was put in the right position, but other word cluster was not. However, if a student put a target word in the right place, it means that he/she had grammatical or contextual knowledge of the word, so it might not be reasonable to give a partial point rather than a full point. Therefore, a more sophisticated and segmented assessment will be necessary to accurately assess students' productive use of word knowledge.

In spite of these limitations, the findings from this study propose meaningful information about the effect of productive versus receptive task and sentence contexts on Korean middle school learners' English vocabulary learning.

REFERENCES

Amiryousefi, M., & Kassaian, Z. (2010). The effects of reading only vs. reading plus enhancement activities on vocabulary learning and production of Iranian pre-university students. *English Language Teaching, 3*(2), p94.

An, H.-S., & Min, C. K. (2011). The Effects of Receptive vs. Productive Vocabulary Instruction. *Journal of the Korea English Education Society, 10*(2), 1-22.

An, H.-S., & Min, C. K. (2014). The Effects of Sentence contexts on Vocabulary Learning. *Studies in English Education, 19*(2), 127-149

Anderson, J. R. (1990). *Cognitive psychology and its implications*: WH Freeman/Times Books/Henry Holt & Co.

Anderson, R., & Freebody, P. (1981). Vocabulary knowledge. In 1. T. Guthrie (Ed.), Comprehension and teaching: Research reviews (pp. 77-117). Newark. *DE: International Reading Association*.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests* (Vol. 1): Oxford University Press.

Bainbridge, J. V., Lewandowsky, S., & Kirsner, K. (1993). Context effects in repetition priming are sense effects. *Memory & Cognition, 21*(5), 619-626.

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human perception and*

performance, 10(3), 340.

Bao, G. (2015). Task type effects on English as a Foreign Language learners' acquisition of receptive and productive vocabulary knowledge. *System, 53*, 84-95.

Barcroft, J. (2004). Effects of sentence writing in second language lexical acquisition. *Second Language Research, 20*(4), 303-334.

Bensoussan, M., & Laufer, B. (1984). Lexical guessing in context in EFL reading comprehension. *Journal of Research in Reading, 7*(1), 15-32.

Berry, D. C., & Broadbent, D. E. (1987). The combination of explicit and implicit learning processes in task control. *Psychological research, 49*(1), 7-15.

Bolger, D. J., Balass, M., Landen, E., & Perfetti, C. A. (2008). Context variation and definitions in learning the meanings of words: An instance-based learning approach. *Discourse Processes, 45*(2), 122-159.

Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental Vocabulary Acquisition from Reading, Reading-While-Listening, and Listening to Stories. *Reading in a Foreign Language, 20*(2), 136-163.

Carroll, B., & Drum, P. (1982). *Effects of context in facilitating unknown word comprehension.* Paper presented at the New inquiries in reading research and instruction. Rochester, NY: National Reading Conference.

Chapelle, C. A. (1994). Are C-tests valid measures for L2 vocabulary research? *Second Language Research, 10*(2), 157-187.

Choi, J.-Y. (2007). *The Effect of Receptive and Productive Tasks on Lexical*

Knowledge Development. (Unpublished master's thesis), Michigan State University.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 213-238.

Day, R. R., Omura, C., & Hiramatsu, M. (1992). Incidental EFL vocabulary learning and reading. *Reading in a Foreign Language, 7*, 541-541.

De La Fuente, M. J. (2002). Negotiation and oral acquisition of L2 vocabulary. *Studies in second language acquisition, 24*(01), 81-112.

Drum, P. A., & Konopak, B. C. (1987). Learning word meanings from written context. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 73-87). Hillsdale, NJ: Lawrence Erlbaum Associates.

Eckerth, J., & Tavakoli, P. (2012). The effects of word exposure frequency and elaboration of word processing on incidental L2 vocabulary acquisition through reading. *Language Teaching Research, 16*(2), 227-252.

File, K. A., & Adams, R. (2010). Should Vocabulary Instruction Be Integrated or Isolated? *TESOL Quarterly, 44*(2), 222-249.

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

Gass, S., Mackey, A., Alvarez-Torres, M. J., & Fernandez-Garcia, M. (1999). The effects of task repetition on linguistic output. *Language Learning, 49*(4), 549-581.

Gass, S. M., Behney, J., & Plonsky, L. (2013). *Second language acquisition: An introductory course*. New York: Routledge.

Ghabanchi, Z., & Ayoubi, E. S. (2012). Incidental vocabulary learning and recall by intermediate foreign language students: The influence of marginal glosses, dictionary use, and summary writing. *Journal of International Education Research (JIER), 8*(2), 85-96.

Griffin, G., & Harley, T. A. (1996). List learning of second language vocabulary. *Applied Psycholinguistics, 17*(04), 443-460.

Hazrat, M. (2015). The Effects of Task Type and Task Involvement Load on Vocabulary Learning. *Waikato Journal of Education, 20*(2).

Heidari-Shahreza, M. A., & Tavakoli, M. (2012). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition by Iranian EFL Learners. *The Language Learning Journal*(ahead-of-print), 1-16.

Hemmati, P., & Asmawi, A. B. Incidental vocabulary Learning and Retention through Reading a Graded Reader among Iraninan EFL Learners. *The Online Journal of New Horizons in Education*, 114.

Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in second language acquisition, 21*(02), 303-317.

Hsueh-Chao, M. H., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language, 13*(1), 403-430.

Hu, H.-c. M. (2013). The Effects of Word Frequency and Contextual Types on Vocabulary Acquisition from Extensive Reading: A Case Study. *Journal of Language Teaching and Research, 4*(3), 487-495.

Hu, H.-c. M., & Nassaji, H. (2012). Ease of inferencing, learner inferential strategies, and their relationship with the retention of word meanings

inferred from context. *Canadian modern language review, 68*(1), 54-77.

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

Jenkins, J. R., Stein, M. L., & Wysocki, K. (1984). Learning vocabulary through reading. *American educational research journal, 21*(4), 767-787.

Jeon, E., & Shin, Y. (2011). Receptive and productive vocabulary learning using a word list in L2. *Elementary English Education, 17*(1), 395-416.

Joe, A. (1998). What effects do text-based tasks promoting generation have on incidental vocabulary acquisition? *Applied linguistics, 19*(3), 357-377.

Kim, S. S. (2013). *Analyses of Receptive and Productive Korean EFL Vocabulary: Computer-based Vocabulary Learning Program.* Arizona State University.

Kim, S. Y., & Lee, S. H. (2008). Learning strategies and instructional approaches for Korean EFL learners' productive vocabulary development. *English Language Teaching, 20*(4), 237-259.

Koda, K. (1989). The effects of transferred vocabulary knowledge on the development of L2 reading proficiency. *Foreign language annals, 22*(6), 529-540.

Laufer, B. (1998). The development of passive and active vocabulary in a second language: same or different? *Applied linguistics, 19*(2), 255-271.

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied linguistics,*

22(1), 1-26.

Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language testing, 16*(1), 33-51.

Laufer, B., & Paribakht, T. S. (1998). The relationship between passive and active vocabularies: Effects of languagelearning context. *Language Learning, 48*(3), 365-391.

Laufer, B., & Rozovski-Roitblat, B. (2011). Incidental vocabulary acquisition: The effects of task type, word occurance and their combination. *Language Teaching Research*, 1362168811412019.

Lee, H. (2003). The effects of production and comprehension for focus on form and second language acquisition. *Journal of the Applied Linguistics Association of Korea, 19*(2), 41-68.

Llach, M. d. P. A. (2009). The effect of reading only, reading and comprehension, and sentence writing in lexical learning in a foreign language:: some preliminary results. *Revista española de lingüística aplicada*(22), 9-34.

Martinez, M. E. (2010). *Learning and cognition: The design of the mind*. Boston, MA: Pearson College Division.

McKeown, M. G., Beck, I. L., Omanson, R. C., & Pople, M. T. (1985). Some effects of the nature and frequency of vocabulary instruction on the knowledge and use of words. *Reading research quarterly*, 522-535.

Meara, P. (2009). *Connected words: Word associations and second language vocabulary acquisition* (Vol. 24): John Benjamins Publishing.

Melka, F. J. (1997). Receptive vs. productive aspects of vocabulary. In N.

Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition, and pedagogy* (pp. 84-102). Cambridge: Cambridge University Press.

Milton, J. (2009). *Measuring second language vocabulary acquisition* (Vol. 45): Multilingual Matters.

Mondria, J.-A., & Wiersma, B. (2004a). Receptive, productive, and receptive+ productive L2 vocabulary learning: What difference does it make. In B. P & L. B (Eds.), *Vocabulary in a second language: Selection, acquisition, and testing* (pp. 79-102). Philadelphia, PA: John Benjamins.

Mondria, J.-A., & Wiersma, B. (2004b). Receptive, productive, and receptive+ productive L2 vocabulary learning: What difference does it make. *Vocabulary in a second language: Selection, acquisition, and testing*, 79-100.

Mondria, J.-A., & Wit-de Boer, M. (1991). The Effects of Contextual Richness on the Guessability and the Retention of Words in a Foreign Language1. *Applied linguistics, 12*(3), 249-267.

Nagy, W. E. (1995). On the role of context in first-and second-language vocabulary learning: Champaign, Ill.: University of Illinois at Urbana-Champaign, Center for the Study of Reading.

Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American educational research journal, 24*(2), 237-270.

Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading research quarterly*, 233-253.

Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Boston: Heinle & Heinle Publishers.

Nation, I. S. P. (2001). *Learning vocabulary in another language*: Ernst Klett Sprachen.

Nation, I. S. P. (2013). Teaching & learning vocabulary.

Nation, I. S. P., & Coady, J. (1988). Vocabulary and reading. *Vocabulary and language teaching, 97*, 110.

Nation, I. S. P., & Gu, P. Y. (2007). *Focus on vocabulary*. Sydney: NCELTR,Macquarie University.

Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. *Second language vocabulary acquisition: A rationale for pedagogy*, 174-200.

Paribakht, T. S., & Wesche, M. (1999). Reading and "incidental" L2 vocabulary acquisition. *Studies in second language acquisition, 21*(02), 195-224.

Paribakht, T. S., & Wesche, M. B. (1993). Reading comprehension and second language development in a comprehension-based ESL program. *TESL Canada journal, 11*(1), 09-29.

Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental Vocabulary Acquisition from an Authentic Novel: Do" Things Fall Apart"? *Reading in a Foreign Language, 22*(1), 31-55.

Pichette, F., De Serres, L., & Lafontaine, M. (2011). Sentence reading and writing for second language vocabulary acquisition. *Applied linguistics*, amr037.

Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning, 52*(3), 513-536.

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14*(3), 191-201.

Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language testing, 10*(3), 355-371.

Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge university press

Rott, S. (2007). The effect of frequency of input-enhancements on word learning and text comprehension. *Language Learning, 57*(2), 165-199.

Ryoo, Y.-s. (2009). Effects of Two Types of Vocabulary Practice: Receptive and Productive. *Foreign languages education, 16*(1), 79-99.

Schatz, E. K., & Baldwin, R. S. (1986). Context clues are unreliable predictors of word meanings. *Reading research quarterly*, 439-453.

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language testing, 18*(1), 55-88.

Schouten-van Parreren, C. (1989). Vocabulary learning through reading: Which conditions should be met when presenting words in texts. *AILA review, 6*(1), 75-85.

Son, J. (2007). The Effects of Vocabulary Exercises on EFL Vocabulary Learning and Retention. *English Language Education, 13*(4), 167-192.

Song, J., & Sardegna, V. G. (2014). EFL learners' incidental acquisition of English prepositions through enhanced extensive reading instruction. *RELC Journal, 45*(1), 67-84.

Srichamnong, N. (2008). Incidental EFL Vocabulary Learning: The Effects of Interactive Multiple-Choice Glosses: Design.

Stahl, S. A. (1986). Three principles of effective vocabulary instruction. *Journal of Reading*, 662-668.

Stallman, A. C. (1991). *Learning vocabulary from context: Effects of focusing attention on individual words during reading*. University of Illinois at Urbana-Champaign.

Sternberg, R. J. (1987). Most vocabulary is learned from context. In M. McKeown & M. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. pp. 89-105). Hillsdale, NJ: Lawrence Erlbaum Associates.

Waring, R. (1997). A study of receptive and productive learning from word cards. *Studies in Foreign Languages and Literature, 21*(1), 94-114.

Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader. *Reading in a Foreign Language, 15*(2), 130-163.

Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in second language acquisition, 27*(01), 33-52.

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied linguistics, 28*(1), 46-65.

Webb, S. (2008). The Effects of Context on Incidental Vocabulary Learning. *Reading in a Foreign Language, 20*(2), 232-245.

Webb, S. (2012). Repetition in Incidental Vocabulary Learning *The Encyclopedia of Applied Linguistics*: John Wiley & Sons, Inc.

Weinfurt, K. P. (2000). Repeated measures analysis: ANOVA, MANOVA, and HLM. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding more multivariate statistics* (pp. 317-361). Washington, DC: American Psychological Association.

Wesche, M., & Paribakht, T. S. (1996). Assessing Second Language Vocabulary Knowledge: Depth Versus Breadth. *Canadian modern language review, 53*(1), 13-40.

Wilkins, D. A. (1972). *Linguistics in language teaching*: E. Arnold, 1973.

Zimmerman, C. B. (1997). Historical trends in second language vocabulary instruction. *Second language vocabulary acquisition*, 5-19

APPENDIX 1. Consent Form

연구참여자용 설명서 및 동의서

연구 과업 명: 수용적/생산적 과업과 문장 문맥이 한국 중학교 학생들의 영어 어휘 보유와 지식에 미치는 영향 **(The Effects of Receptive/Productive Tasks and Sentence Contexts on English Vocabulary Retention and Knowledge of Korean Middle School Students)**

연구 책임자명: 김이경 (서울대학교 사범대학 영어교육과 석사과정, 학생)

본 연구는 과업 유형, 문맥의 다양성, 그리고 그 조합이 한국 중학생 영어 학습자의 어휘학습에 미치는 영향에 대해 알아보는 연구입니다. 귀하는 한국인 영어 학습자로서 본 연구의 대상에 적합하다고 판단되기 때문에 이 연구에 참여하도록 권유 받았습니다. 이 연구를 수행하는 서울대학교 소속의 연구원, <u>김이경</u>,은 연구에 대한 모든 설명을 해주고 실험에 대한 절차를 책임지고 진행할 것입니다. 귀하가 본 실험에 참여하기 위한 의사를 결정하기 전에 본 연구의 수행목적과 내용에 대해 이해하는 것이 중요합니다. 다음 제시사항을 자세히 읽어보신 후 참여 의사를 밝혀주시길 바라며, 필요에 따라 가족이나 선생님께 의논 드려 보시고 결정해주십시오. 내용을 모두 숙지하고 질문이 있다면 담당 연구원에게 질문해주십시오. 담당 연구원이 자세하게 설명해줄 것입니다.

1. 이 연구는 왜 실시합니까?

이 연구는 과업 유형, 문맥의 다양성이 한국 중학생 영어 학습자의 어휘학습에 미치는 영향을 알아보기 위해서 실시합니다.

2. 얼마나 많은 사람이 참여합니까?

영어를 외국어로 배우는 중 3 한국인 영어학습자 120 명 (미성중학교 총 4 개 반)이 연구에 참여할 것입니다.

3. 만일 연구에 참여하면 어떤 과정이 진행됩니까?

만일 귀하가 참여의사를 밝혀 주시면 본 연구에 들어가기에 앞서 4 개의 반의 동질성을 검사하기 위해서 중간고사 영어 내신 성적이 사용될 예정입니다. 성적 정보는 학생 개인정보 없이 숫자로만 제공될 것입니다. 본 연구는 본교 영어 수업의 일환으로 진행될 것입니다. 한 수업 당 45 분씩 총 3 회 차로 진행되며 연구 과정은 크게 총 3 단계 (어휘 과업, 사후 테스트, 지연 사후 테스트)로 진행될 것입니다.

1) 어휘 과업 단계는 총 2 회 차에 걸쳐 진행될 것입니다. 과업의 종류 (수용적 과업(R)/ 생산적 과업(P))와 문맥의 다양성(다양(D)/ 동질(S)) 따라 총 4 개의 반 (RD, RS, PD, PS)으로 구성되며 귀하가 속한 반에 따라 각기 다른 과업을 수행하게 될 것입니다. 수용적 과업(R)은 주어진 영어 문장을 한국어로 해석하는 것이며 생산적 과업(P)은 주어진 한국 문장을 영어로 작문 하는 것입니다. 또한 문맥이 다양(D)한 집단은 해당 어휘에 대한 다양한 예문으로 과업을 하고 문맥이 동질(S)한 집단은 해당 어휘에 대한 같은 예문으로 과업을 진행하게 될 것입니다.

2) 사후 테스트는 2 번째 수업이 끝난 직후에 진행될 것입니다. 두 차례의 수업 시수가 끝난 후 과업의 종류와 문맥의 다양성이 학생들의 어휘 학습에 미치는 영향을 보기 위해 사후 테스트가 진행될 것입니다.

3) 1 주일 후 영어 수업 시간에 어휘의 보유 정도를 알아보기 위하여 지연 사후 테스트를 실시할 것입니다.

4. 연구 참여 기간은 얼마나 됩니까?

약 3 주 동안 일주일에 1 번씩 총 3 회 한 회당 45 분씩 참여하도록 요청받을 것입니다.

5. 참여 도중 그만두어도 됩니까?

예, 귀하께서 실험 참여에 불편함을 느낀다면 언제든지 어떠한 불이익 없이 그만 둘 수 있습니다. 만일 귀하께서 연구에 참여하시는 것을 중단하고 싶다면 담당 연구원에게 즉시 말씀해주십시오.

6. 부작용이나 위험요소는 없습니까?

본 연구는 기존에 진행되는 본교 영어 수업의 일부로써 진행되고 동질성을 위해 쓰이는 중간고사 성적은 학생 정보 없이 오로지 점수로만 제공되기 때문에 안전에 대한 위협이 없을 것으로 예상됩니다. 연구에 대한 자료는 오직 연구만을 위하여 사용하고 결코 외부에 노출하거나 다른 용도로 사용하지 않을 것이며 이 연구결과를 작성할 때 참여자들의 이름을 모두 익명으로 표기하여 제시할 것입니다. 또한 연구 참여 도중 발생하는 문제나 불편함이 있다면 즉각적으로 담당 연구원(김이경)에게 말씀해주십시오. 바로 불편사항을 처리하겠습니다. 이외에도 연구 참여 도중 발생할 수 있는

부작용이나 위험 요소에 대한 질문이 있으면 담당 연구원(김이경)에게 즉시 문의해 주십시오.

7. 이 연구에 참여시 참여자에게 이득이 있습니까?

귀하가 이 연구에 참여하는데 있어서 직접적인 이득은 없습니다. 그러나 귀하가 제공하는 정보는 과업 유형, 문맥의 다양성, 그리고 그 조합이 한국 중학생 영어 학습자의 어휘학습에 미치는 영향을 이해하는 데 도움이 될 것입니다.

8. 만일 이 연구에 참여하지 않는다면 불이익이 있습니까?

귀하는 언제든지 본 연구에 참여하지 않을 자유가 있습니다. 또한, 귀하가 본 연구에 참여하지 않아도 귀하에게는 어떠한 불이익도 없습니다.

9. 연구에서 얻은 모든 개인 정보의 비밀은 보장됩니까?

개인정보관리책임자는 서울대학교 소속 연구원 김이경 (010-62565-3083)입니다. 저는 이 연구를 통해 얻은 모든 개인 정보의 비밀 보장을 위해 최선을 다할 것입니다. 이 연구에서 얻어진 개인 정보가 학회지나 학회에 공개 될 때 귀하의 개인 정보는 사용되지 않을 것입니다. 그러나 만일 법이 요구하면 귀하의 개인정보는 제공될 수도 있습니다. 또한 모니터 요원, 점검 요원, 생명윤리심의위원회는 연구 참여자의 개인 정보에 대한 비밀 보장을 침해하지 않고 관련규정이 정하는 범위 안에서 본 연구의 실시 절차와 자료의 신뢰성을 검증하기 위해 연구 결과를 직접 열람할 수 있습니다. 귀하가 본 동의서에 서명하는 것은, 이러한 사항에 대하여 사전에 알고 있었으며 이를 허용한다는 동의로 간주될 것입니다.

10. 이 연구에 참가하면 대가가 지급됩니까?

귀하의 연구 참여시 감사의 뜻으로 소정의 간식이 지급될 예정입니다.

11. 연구에 대한 문의는 어떻게 해야 됩니까?

본 연구에 대해 질문이 있거나 연구 중간에 문제가 생길 시 다음 연구 담당자에게 연락하십시오.

이름: 김 이경 전화번호: 010-6256-3083

만일 어느 때라도 연구참여자로서 귀하의 권리에 대한 질문이 있다면 다음의 서울대학교 생명윤리심의위원회에 연락하십시오.

서울대학교 생명윤리심의위원회 (SNUIRB) 전화번호: 02-880-5153

동 의 서

1. 나는 이 설명서를 읽었으며 담당 연구원과 이에 대하여 의논하였습니다.

2. 나는 위험과 이득에 관하여 들었으며 나의 질문에 만족할 만한 답변을 얻었습니다.

3. 나는 이 연구에 참여하는 것에 대하여 자발적으로 동의합니다.

4. 나는 이 연구에서 얻어진 나에 대한 정보를 현행 법률과 생명윤리심의위원회 규정이 허용하는 범위 내에서 연구자가 수집하고 처리하는데 동의합니다.

5. 나는 담당 연구자나 위임 받은 대리인이 연구를 진행하거나 결과 관리를 하는 경우와 보건 당국, 학교 당국 및 서울대학교 생명윤리심의위원회가 실태 조사를 하는 경우에는 비밀로 유지되는 나의 개인 신상 정보를 직접적으로 열람하는 것에 동의합니다.

6. 나는 언제라도 이 연구의 참여를 철회할 수 있고 이러한 결정이 나에게 어떠한 해도 되지 않을 것이라는 것을 압니다.

7. 나의 서명은 이 동의서의 사본을 받았다는 것을 뜻하며 연구 참여가 끝날 때까지 사본을 보관하겠습니다.

_____ _____ _____

연구참여자 성명 서 명 날짜 (년/월/일)

_____ _____ _____

동의서 받은 연구원 성명 서 명 날짜 (년/월/일)

_____ _____ _____

연구책임자 성명 서 명 날짜 (년/월/일)

_____ _____ _____

법정 대리인 성명(참여자와 관계) 서 명 날짜 (년/월/일)

APPENDIX 2. Receptive Task

반: _____ 이름: _____

★주어진 영어 단어의 뜻을 쓰고 주어진 문장을 우리말로 해석해보세요

sacrifice		wander	
He made a **sacrifice** of himself to save his town**.**		She **wander**ed aimlessly around the streets.	
firm		**charity**	
He works for an aircraft **firm.**		Many **charities** sent money to the victims.	
inquire		**dispose**	
I will **inquire** about how to get there		She **disposed** books in order**.**	
recognize		**proclaim**	
I could not **recognize** my old friend**.**		The president **proclaimed** a state of emergency.	

★주어진 영어 단어의 뜻을 쓰고 주어진 문장을 우리말로 해석해보세요

sacrifice		wander	
A war involves the **sacrifice** of many lives**.**		We **wander**ed back towards the car..	
firm		**charity**	
The accounting **firm** audited the company		She does a lot of work for **charity**.	
inquire		**dispose**	
I **inquire**d about the reason of his long absence.		The DVDs are **dispose**d in alphabetical order**.**	
recognize		**proclaim**	
I **recognize** the need for safety**..**		He **proclaim**ed her a traitor.	

★주어진 영어 단어의 뜻을 쓰고 주어진 문장을 우리말로 해석해보세요

sacrifice		wander	
He helped them at the **sacrifice** of himself		Those sheep **wander** all over the place.	
firm		charity	
I am not a member of the **firm.**		Any money that is left over will go to **charity**.	
inquire		dispose	
I will **inquire** into what happened.		He **dispose**d a fleet in a straight line**.**	
recognize		proclaim	
You can **recognize** this tune**.**		The citizens **proclaim**ed him as their king.	

APPENDIX 3. Productive Task

반: _____ 이름: _____

★주어진 뜻에 맞는 영단어를 쓰고 주어진 문장을 영작하여 쓰세요

	희생; 희생물		거닐다, 돌아다니다
그는 그의 마을을 지키기 위하여 스스로 희생을 했다.		그녀는 거리를 정처 없이(aimlessly) 돌아다녔다.	
	회사, 사무소		자선[구호] 단체
그는 항공(aircraft)회사에서 일한다		많은 구호 단체들이 그 희생자들에게 돈을 보냈다	
	(…에게) 묻다		배치하다, 배열하다
나는 그곳에 가는 방법을 물어보겠다		그녀는 책을 순서대로 배치했다.	
	~ 알아보다[알다]		선언[선포]하다
나는 옛 친구를 알아볼 수 없었다.		대통령이 국가(state) 비상사태 (emergency)를 선포했다.	

★주어진 뜻에 맞는 영단어를 쓰고 주어진 문장을 영작하여 쓰세요

	희생; 희생물		거닐다, 돌아다니다
전쟁은 많은 생명의 **희생**을 수반한다(involve).		우리는 천천히 **거닐며** 다시 차 있는 쪽으로 갔다.	
	회사, 사무소		자선[구호] 단체
그 회계(accounting) **사무소**는 그 회사(company)의 회계 감사를 했다(audit).		그녀는 많은 **자선 단체** 활동을 하였다	
	(…에게) 묻다		배치하다, 배열하다
나는 그의 오래 부재의 이유를 **물었다**.		DVD들은 알파벳 순으로 **배열되어 있다.**	
	~ 알아보다[알다]		선언[선포]하다
난 안전의 필요성을 **안다**.		그는 그녀를 반역자(traitor)라고 **선포했다.**	

★주어진 뜻에 맞는 영단어를 쓰고 주어진 문장을 영작하여 쓰세요

	희생; 희생물		거닐다, 돌아다니다
그는 자신의 **희생**으로 그들을 도왔다.		저 양들은 사방을 **돌아다닌다**.	
	회사, 사무소		자선[구호] 단체
저는 그 **회사** 사람이 아닙니다		남은(left over) 돈은 **자선 단체**에 보낼 것이다	
	(…에게) 묻다		배치하다, 배열하다
나는 무슨 일이 있었는지 **물어볼 거야**		그는 함대(fleet)를 일렬로 **배치했다.**	
	~ 알아보다[알다]		선언[선포]하다
너는 이 곡조(tune)를 **알 수 있다.**		국민은 그를 왕으로 **선포하였다.**	

APPENDIX 4. Active Word Learning Test; Immediate

반 _____이름: _____

★주어진 우리 말에 해당하는 영어 단어를 쓰세요. 일부만 써도 좋으니 최대한 기억나

는 대로 써주세요

1	거닐다, 돌아다니다	
2	희생; 희생물	
3	배치하다, 배열하다	
4	~ 알아보다[알다]	
5	선언[선포]하다	
6	(...에게) 묻다	
7	자선[구호] 단체	
8	회사, 사무소	

APPENDIX 5. Recognition Test and

Passive Word Learning Test; Immediate

반 _____이름: _____

★주어진 영어 단어를 아는 정도에 따라 0-3에 동그라미(o) 치세요.

☞ 각 숫자는 아래와 같은 단어 지식의 정도를 의미합니다.

0	전에 한번도 본적이 없고 모르는 단어다
1	전에 본적은 있지만 의미를 모른다
2	전에 본적이 있고 그 의미를 대충 짐작한다고 생각한다
3	전에 본적이 있고 그 의미를 안다 ★ 3을 택할 경우 맨 오른쪽에 그 영어 단어에 해당하는 우리말 뜻을 쓰세요.

1	recognize	0	1	2	3	
2	sacrifice	0	1	2	3	
3	dispose	0	1	2	3	
4	wander	0	1	2	3	
5	firm	0	1	2	3	
6	proclaim	0	1	2	3	
7	charity	0	1	2	3	
8	inquire	0	1	2	3	

이 수업을 통해 배운 단어가 아닌 알고 있던 단어가 있다면 체크해주세요!

□ recognize □ sacrifice □ dispose □ wander
□ proclaim □ charity □ inquire □ firm

APPENDIX 6. Gap-Filling Test; Immediate

반 _____이름: _____

★문맥에 맞도록 [보기]에서 알맞은 단어를 선택하여 빈칸에 쓰세요.

[보기]

sacrifice, adopt, firm, inquire, wander, raise,

dispose, proclaim, pour, recognize, charity

1. I will _____ for the shoes at the department store.

2. Don't _____ around alone after midnight.

3. She did not _____ me when she saw me.

4. He is working for an engineering _____.

5. The local _____ will raise money for the poor.

6. She _____ed her books in order.

7. My grandfather made a _____ of his life in the World War Ⅱ.

8. Sir Winston Churchill was _____ed honorary U.S. citizen.

APPENDIX 7. Word Reordering Test; Immediate

반 _____이름: _____

★주어진 단어를 재배열하여 문법적으로 의미 있고 뜻이 통하는 문장을 만드세요

1. disposed/ his soldiers /he/ for the war

    _____

2. him /recognize / I /did / not/ at once

    _____

3. a sacrifice/ of her life / for her family/ made / she

    _____

4. a law / they / for / firm/work

    _____

5. to the school/ how / I / about/ inquired/to get

    _____

6. wandering/ she/ the streets/ around / is

    _____

7. its independence/ the new government/ proclaimed/ in Venezuela

    _____

8. all his money/ to charity/ sent/he

    _____

APPENDIX 8. Active Word Learning Test; Delayed

반 _____이름: _____

★주어진 우리 말에 해당하는 영어 단어를 쓰세요. 일부만 써도 좋으니 최대한 기억나는 대로 써주세요

1	(...에게) 묻다	
2	자선[구호] 단체	
3	회사, 사무소	
4	거닐다, 돌아다니다	
5	희생; 희생물	
6	배치하다, 배열하다	
7	~ 알아보다[알다]	
8	선언[선포]하다	

APPENDIX 9. Recognition Test and

Passive Word Learning Test; Delayed

반 _____이름: _____

★주어진 영어 단어를 아는 정도에 따라 0-3에 동그라미(o) 치세요.

☞ 각 숫자는 아래와 같은 단어 지식의 정도를 의미합니다.

0	전에 한번도 본적이 없고 모르는 단어다
1	전에 본적은 있지만 의미를 모른다
2	전에 본적이 있고 그 의미를 대충 짐작한다고 생각한다
3	전에 본적이 있고 그 의미를 안다 ★ 3을 택할 경우 맨 오른쪽에 그 영어 단어에 해당하는 우리말 뜻을 쓰세요.

1	inquire	0	1	2	3	
2	charity	0	1	2	3	
3	recognize	0	1	2	3	
4	sacrifice	0	1	2	3	
5	dispose	0	1	2	3	
6	wander	0	1	2	3	
7	firm	0	1	2	3	
8	proclaim	0	1	2	3	

APPENDIX 10. Gap-Filling Test; Delayed

반 _____이름: _____

★문맥에 맞도록 [보기]에서 알맞은 단어를 선택하여 빈칸에 쓰세요.

[보기]

sacrifice, adopt, firm, inquire, wander, raise,

dispose, proclaim, pour, recognize, charity

1. I work at a law _____

2. I will made a _____ of my life to my country

3. Don't _____ around late at night.

4. He will _____ me at once.

5. He _____ed liberty throughout all the land.

6. The concert will raise money for local _____es.

7. She _____ed her clothes and shoes.

8. I will _____ for the book at the bookstore

.

APPENDIX 11. Word Reordering Test; Delayed

반 _____이름: _____

★주어진 단어를 재배열하여 문법적으로 의미 있고 뜻이 통하는 문장을 만드세요

1. a sacrifice/ for her children/ made / she/ of her happiness

    _____

2. I /an engineering / for / firm/work

    _____

3. the way / I / about / to the station/ inquired.

    _____

4. recognize / I /did / the animal /not

    _____

5. he/ to wander/ the streets/ around / likes

    _____

6. all his property/ donated/ to charity/ the old man/

    _____

7. his soldiers /he/ disposed/ for the battle

    _____

8. emperor/he/ proclaimed/ himself

    _____

국 문 초 록

본 연구는 수용적, 생산적 어휘 과업과 동일, 다양한 문장 문맥이 한국 중학교 학생들의 영어 어휘 학습에 미치는 영향에 대해 다음 두 가지 관점에서 검증하고자 한다. 우선, 본 연구는 이 두 가지 변인이 전반적인 어휘 학습에 미치는 영향을 단어 인지 시험, 소극적/적극적 단어 학습 시험, 두 개의 생산적 어휘 사용 시험 (빈칸 넣기, 단어 재배열 시험)의 총점을 분석하여 조사하고자 한다. 둘째로, 이 각각의 시험에서 측정된 세부적인 어휘 지식 학습과 보유를 이 두 개의 변인에 근거하여 살펴보고자 한다.

단어의 수용적, 생산적인 면모는 의사소통 과정의 두 가지 근간을 이루기 때문에 두 면모를 다 살펴보는 것은 학문적인 의미가 있다. 비록 많은 연구가 단어 학습에 있어 수용적 과업에 비해 생산적 과업의 우월성에 동의하지만 그에 반하는 연구들 역시 존재해 연구 결과가 동일하지 않다는 점에서 이 두 과업을 비교하는 연구가 더 필요한 실정이다. 한국에서 단어 과업이 대개 수용적 학습에만 의존하는 경향이 있다는 점 역시 생산적 학습에 대한 필요성을 주창한다.

어휘 학습에 있어 다른 중요한 요소인 문장 문맥의 영향에 대한 결과 역시 논란의 여지가 있다. 어휘 학습에 있어 문맥의 필요성에 대해서는 많이 연구되었지만 어떤 식으로 문장 문맥을 제공하는 것이 더 효과적인지에 대한 연구는 극히 미미한 실정이기 때문이다.

즉, 이 두 가지 변인은 어휘 학습에 중요한 영향력을 행사하지만 이들의 상호작용에 대한 연구는 거의 없다. 그러므로 본 연구는 동일/다양한 문장 문맥이 특정 단어 과업 안에서 주어졌을 때 그들이 상호적으로 어휘 지식 발달

에 미치는 영향력에 대해 살펴보고자 한다. 또한 어휘 지식의 다양한 면모를 살펴보기 위해서 다섯 개의 다양한 시험이 행해질 것이다.

본 연구의 참여자인 117 명의 한국 중학교 3 학년 학습자는 수용적/생산적 과업과 동일/다양한 문장 문맥이 조합된 4가지 과업 (RS, RD, PS, PD) 중 하나를 수행하고 즉시 사후 평가와 1 주일 후 이루어진 지연 사후 평가에 응하였다. 두 사후 평가는 앞서 말한 다섯 개의 시험으로 구성되어 있다.

전반적인 어휘 학습을 살펴본 결과, 과업의 종류는 어휘 학습과 보유에 상당하게 영향을 미치는 반면 문장 문맥은 그렇지 못했다. 그러나 어휘 보유에서 두 변인의 상호작용이 드러났다. 같은 문장 문맥을 제공하는 것이 생산적 과업에서는 긍정적인 효과를 내었지만 수용적 과업에서는 그렇지 못했다. 어휘 수용에 관한 한 다섯 개의 시험의 각각의 결과 역시 두 개의 생산적 어휘 사용 시험을 제외하고는 과업의 종류의 영향은 상당하지만 문장 문맥의 영향은 미미하다는 전반적 어휘 학습의 분석과 비슷한 결과를 냈다.

무엇보다 생산적 과업은 어휘의 생산적 사용 지식을 제외한 어휘 지식의 여러 단계에서 통계적으로 상당한 보유력을 지니고 있었다. 또한 과업과 상호 작용 시 문장 문맥은 특히 소극적 어휘 학습 시험과 적극적 어휘 학습 시험에서 강력한 영향력을 행사하였다. 반면 어휘의 생산적 사용 시험에서 어휘 지식의 보유를 측정할 때는 과업의 종류보다는 문장 문맥의 다양성 여부가 상당한 영향력을 보였다. 결과에 근거하여 본 연구는 과업의 종류와 문맥의 다양성에 대한 연구의 제언을 결론부에 제시한다.

주요어: 생산적 수용적 어휘 과업, 문장 문맥, 어휘 학습과 보존, 어휘 지식
학 번: 2011-23632

교육학석사학위논문

The Effects of Receptive/Productive Tasks and Sentence Contexts on English Vocabulary Retention and Knowledge of Korean Middle School Students

수용적/생산적 과업과 문장 문맥이 한국 중학교
학생들의 영어 어휘 보유와 지식에 미치는 영향

2016년 2월

서울대학교 대학원

외국어교육과 영어전공

김 이 경

The Effects of Receptive/Productive Tasks and Sentence Contexts on English Vocabulary Retention and Knowledge of Korean Middle School Students

by

EE KYOUNG KIM

A Thesis Submitted to

the Department of Foreign Language Education

in Partial Fulfillment of the Requirements

for the Degree of Master of Arts in Education

At the

Graduate School of Seoul National University

February 2016

The Effects of Receptive/Productive Tasks and Sentence Contexts on English Vocabulary Retention and Knowledge of Korean Middle School Students

수용적/생산적 과업과 문장 문맥이 한국 중학교
학생들의 영어 어휘 보유와 지식에 미치는 영향

지도교수 이 병 민

이 논문을 교육학 석사 학위논문으로 제출함
2016년 2월

서울대학교 대학원
외국어교육과 영어전공
김 이 경

김이경의 석사학위논문을 인준함
2016년 2월

<table>
<tr><td>위 원 장</td><td>_____</td></tr>
<tr><td>부위원장</td><td>_____</td></tr>
<tr><td>위   원</td><td>_____</td></tr>
</table>

The Effects of Receptive/Productive Tasks and Sentence Contexts on English Vocabulary Retention and Knowledge of Korean Middle School Students

APPROVED BY THESIS COMMITTEE:

_____

Jin-Wan Kim, COMMITTEE CHAIR

_____

Sun-Young Oh

_____

Byungmin Lee

ABSTRACT

The present thesis attempts to investigate the effects of task type (productive versus receptive) and sentence contexts (same versus diverse) on the vocabulary learning of Korean middle school English students in two areas: overall vocabulary learning, and the gain and retention of specific vocabulary knowledge. First, this study will look at the impact that the two variables have on overall vocabulary learning; measured using the sum score of five different tests (recognition, passive word learning, active word learning, and two productive vocabulary use tests: gap-filling and word reordering). Second, the gain and retention of specific vocabulary knowledge measured by the five tests will be compared to verify the impact of the two variables.

The receptive and productive aspects of vocabulary have been derived from the two fundamental communication processes, so both aspects are worth studying. Although many studies agree on the superiority of productive tasks over receptive tasks in vocabulary instruction, the results between these studies have been inconsistent; therefore, more research is needed on the impact of these two tasks.

Moreover, vocabulary tasks in Korea largely depend on receptive vocabulary instruction rather than productive instruction, which goes against the majority of findings from previous research that suggest productive vocabulary instruction is more effective. Context, the other important factor for vocabulary learning, has been a controversial issue in the vocabulary instruction research.

Many studies were conducted to determine whether context should be provided for vocabulary learning but few studies were conducted on how to effectively provide context for vocabulary instruction. In other words, these two factors are significant factors influencing vocabulary learning, but few studies have been conducted to investigate the relationship between these two variables. Therefore, this study attempts to integrate sentence contexts into the types of tasks so that the interactive effect of both variables on vocabulary knowledge development can be examined. Besides, this study attempts to scrutinize the multifaceted features of lexical knowledge, so five different sorts of assessment have been implemented.

In this study, 117 3rd grade middle school students in Korea completed one of four different treatment combinations, each having a different combination of the two task types and two sentence contexts (receptive task and same context - RS, receptive task and diverse context - RD, productive task and same context - PS, productive task and diverse context – PD), and took immediate and one-week delayed post-tests. Each of the two tests was composed of five different tests.

Regarding overall vocabulary learning, the results of this study revealed that task type was a factor that significantly affected vocabulary learning in both immediate word gain and its retention but sentence contexts were not. However, the interaction effect between the two variables was shown in word retention. The same context had a positive effect on the productive task but not on the receptive task. The findings from the individual analysis of the five vocabulary

tests showed similar results regarding word retention with the exception of the two productive use tests. The task effect was substantial, while that of context was not.

Above all, the productive task was statistically shown to have considerable power to help students retain several stages of vocabulary knowledge with the exception of the productive use tests. When combined with the task, sentence contexts had a strong effect on vocabulary learning in passive and active word learning tests. On the other hand, the retention of word knowledge, measured by the productive use of vocabulary tests, was influenced fundamentally by sentence contexts rather than task type. Results and the implications regarding task types and sentence contexts are discussed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1.
INTRODUCTION

This chapter introduces the research by presenting the purpose of the study. Section 1.1 discusses the purpose of the study. Section 1.2 presents the research questions, and Section 1.3 outlines the overall structure of the study.

1.1. The Purpose of the Study

Vocabulary is one of the significant factors in language learning since lexical knowledge is the most fundamental and essential for actual communication. Thus, Wilkins (1972) stated that "without grammar very little can be conveyed, without Lexis nothing can be conveyed" (p.11), representing that vocabulary mostly conveys its meaning in order to comprehend and produce messages.

Although many practitioners and learners agree on the importance of vocabulary instruction and often ascribe communication breakdown to the lack of vocabulary knowledge, vocabulary is one of the most neglected issues in the ESL research field (Zimmerman, 1997). Because of insufficient lexical input in EFL/ESL settings, it is a significant challenge for EFL/ESL learners to possess sufficient lexical knowledge.

Regarding the sufficient amount of vocabulary knowledge needed, Nation (2006) and Schmitt (2008) advocated that English learners have to know about 8,000-9,000 words for reading and 5,000-7,000 words for speaking and listening.

Not surprisingly, many ESL/EFL students fail to reach that vocabulary level without explicit vocabulary instruction (Nation, 2006). This creates a demand for more effective vocabulary instruction in ESL/EFL education settings.

Vocabulary instruction, especially in Korea, depends largely on students and their rote memorization of isolated single words. Most vocabulary tasks employed were mostly receptive-oriented (Kim, 2013). Receptive-centered vocabulary instruction may lead to discrepancies between English learners' comprehension and their production of words. Korean learners of English may have no difficulties retrieving some words for receptive uses such as reading and listening but it is difficult for them to retrieve them for productive purposes such as writing and speaking. Hence, effective vocabulary learning that can provoke both receptive and productive vocabulary knowledge is required.

The efficiency of vocabulary learning can be enhanced when words are provided with definitions and contextual clues and processed at a deeper level (Stahl & Fairbanks, 1986). It is necessary to select, sequence and present vocabulary appropriately and to choose the right tasks that integrate vocabulary knowledge development into communication when designing effective vocabulary instruction (Nation, 2001). Therefore, it is important to guide learners by providing them with appropriate task types and context for more effective and efficient vocabulary instruction.

As to the vocabulary task, its receptive and productive aspects have been explored a lot in the previous studies. The receptive and productive aspects of vocabulary have been derived from the two fundamental communication

processes: comprehension and production (Nation, 2001). Based on input and output process of communication, receptive task and productive task contribute considerably to vocabulary learning (An & Min, 2011; Shintani, 2011; Stahl & Fairbanks, 1986). Moreover, a word is presented in a relevant context in the process of communication which implies that context can be a more useful tool for language learning, especially vocabulary learning (Sternberg, 1987). For these reasons, the effects of task types and context on vocabulary learning are important factors to be explored

Many researchers have studied how differing task types, receptive and productive vocabulary instruction, involve lexical knowledge development. Although the majority of research agreed on the priority of productive tasks over receptive tasks in vocabulary instruction (Hulstijn & Laufer, 2001; Kim, 2013; Son, 2007; Pichette, De Serres, & Lafontaine, 2011; Webb, 2005), there were some studies opposed to this result (Barcroft, 2004). That is, the efficacy of receptive versus productive tasks on language learners' vocabulary learning has been open to debate (Webb, 2005). Moreover, vocabulary tasks in Korea depend largely on receptive vocabulary instruction rather than productive instruction, which required the productive vocabulary teaching. For that reason, exploring the effects of both task types, receptive and productive, on vocabulary knowledge gain and retention may provide valuable data that can enhance current vocabulary instruction in Korea.

Sentence contexts, the other important factor for vocabulary learning, have been a controversial issue in the field of vocabulary instruction as well. A lot of

researchers suggested the positive effect of sentence contexts on vocabulary learning by simulating schema and providing sufficient cognitive cues so as to help reinforce word retention (An & Min, 2014; J. R. Anderson, 1990; Bolger, Balass, Landen, & Perfetti, 2008; Schouten-van Parreren, 1989). Other researchers questioned the positive impact of sentence contexts since they tend to increase the cognitive load (File & Adams, 2010; Mondria & Wit-de Boer, 1991) and scatter learners' attention with too many cues (Hu & Nassaji, 2012; Nation & Coady, 1988).

Many studies were conducted to determine whether sentence contexts should be provided for vocabulary learning, but few studies were conducted how to provide the sentence contexts for vocabulary learning. Only An and Min (2014), Bolger and Balass et al. (2008) and Sternberg's (1987) studies dealt with the effect of sentence contexts. That is, they compared the differential effect between the diverse contexts and the same context on vocabulary learning. Although both studies proved the benefits of the diverse sentence contexts, the sentence contexts were only given through the receptive tasks and the number of experiments was small. Thus, further investigation of the role of the sentence contexts in vocabulary learning is required.

When a word was given with its sentence context without its definition, learners guess its meaning from the sentence. However, context guessing can be influenced by other variables (Nagy, 1995) and can lead to a false grasp of the word definition. Thus, this present study provides a definition of the target vocabulary which was regarded as an important factor for effective vocabulary

instruction (for example, An & Min, 2014; Bolger et al.; 2008; Stahl & Fairbanks, 1986). In addition, as suggested by Mondria & Wit-de Boer (1991), a single sentence context per a target word can diminish learners' cognitive load and prevent their attention from scattering due to excessive cues (Kim, 2013; Pichette et al., 2011), so this study employed a sentence context per target vocabulary for word learning tasks.

Overall, previous studies on receptive and productive vocabulary learning merely focused on comparing the effects of the two tasks. This study attempts to integrate sentence contexts into the types of tasks so that the interactive effect of both variables on vocabulary knowledge development can be examined.

It can be another significant issue to define lexical knowledge because of its multifaceted feature (Laufer & Nation, 1999), which needs to be reflected properly in studies on vocabulary instruction. In Korea, however, the vocabulary tests were usually limited to simple tests requiring 1:1 translations of context-excluded word items. This assessment measure is not sufficient enough to measure multifaceted vocabulary knowledge. The issue calls for a more comprehensive assessment.

Lexical processing needs to be investigated further to comprehend what it means to "know" a word and to further discover the constructs of vocabulary knowledge (Nation, 2001). In the endeavor to evaluate ESL learners' vocabulary knowledge, Paribakht and Wesche (1993) created a Vocabulary Knowledge Scale (VKS). The VKS requires learners to self-report their knowledge with five levels of word recognition, ranging from passive word knowledge to its

composition, showing that their vocabulary develops from partial to full knowledge (Nation, 2001).

This study extracted stage 1 to 3 of the VKS for the word recognition test and stage 4 for the passive word learning test. Because of the huge gap in difficulty level between stage 4 (retrieval of a target word) to stage 5 (free writing with a target word), learners may fail to prove their productive use knowledge of words (Bolger et al., 2008). The negative results of stage 5 can be triggered by the lack of their language proficiency, not by their vocabulary knowledge deficiency.

Therefore, assessing the learners' productive use knowledge of words by the VKS had a limitation, especially in ESL/EFL learners with low language skills. Specifically, most Korean middle school students are not familiar with composition, so it is difficult to assess their productive use of word knowledge through free-writing. This present study employed two productive use tests, gap-filling and word reordering, in order to compensate the limitation of the VKS. In addition, the VKS omitted the retrieval of word items from its meaning, called active word knowledge; this study also added the active word learning test.

According to An and Min (2011), the context-included tests led to a significant difference from the context-excluded tests, so both types of assessments are required for an in-depth understanding of lexical knowledge. To understand the overall depth of vocabulary knowledge and its development, five test items were employed to evaluate word recognition, passive and active word knowledge, productive word use in proper context and grammar.

To summarize, the impact of task types (productive versus receptive) and sentence contexts (diverse versus same) on vocabulary learning are controversial issues. Furthermore, little research that demonstrates the influence of the two variables on lexical knowledge development considering various factors has been conducted. As a result, the present study investigates those interventions related to vocabulary instruction in the Korean EFL classroom setting to observe how those factors contribute to Korean middle school English learners' lexical knowledge development.

1.2. Research Questions

The focus of the present study is to investigate the effects of receptive/productive task and sentence contexts on the vocabulary learning of Korean middle school English learners from the following two perspectives. First, this study looks into the impact of task types (productive versus receptive) and sentence contexts (same versus diverse) on the overall vocabulary learning of Korean middle school English learners. Here, the overall learning refers to the sum of five different test scores. Second, the gain and retention of specific vocabulary knowledge measured by five different tests are compared regarding task types and sentence contexts. The participants in this study completed one of four treatments with different task and context combinations and took the five sorts of the immediate and delayed post-tests. Every experimental process was

thoroughly developed and administrated to answer the following research questions.

1. How do the type of task (receptive versus productive) and sentence contexts (diverse versus same) influence Korean middle school English learners' overall immediate vocabulary learning and its retention?

2. How do these two factors influence Korean middle school English learners' vocabulary knowledge measured by five different vocabulary tests?

1.3. Organization of the Thesis

The present study consists of five chapters. Chapter 1 introduces the purpose of the study and presents the research questions. Chapter 2 provides an overview of the literature review on vocabulary knowledge and the effect of task and context on vocabulary learning. In Chapter 3, the methodology of this study is described regarding the research design, the participants, the procedure, the instruments, the treatment, the assessment, and the data analysis. Chapter 4 presents the results and discusses the research findings. Finally, Chapter 5 concludes the research with a summary of the significant findings and shows the implications of the present study and the suggestions for further research.

CHAPTER 2.
LITERATURE REVIEW

The current chapter presents the literature overviews about the effect of task type—receptive and productive—and context on vocabulary knowledge development. Section 2.1 discusses vocabulary knowledge in specifying its three components: receptive and productive aspects, breadth and depth, and context of vocabulary use. Section 2.2 details the main issues involved in this study—receptive versus productive task and context that calls for comprehending vocabulary knowledge development

2.1. Vocabulary Knowledge

The issue of "knowing" vocabulary had been demonstrated and debated among a large number of previous studies. Bachman and Palmer (1996) stated that vocabulary knowledge is the ability to use general and concrete words precisely with the appropriate contexts. However, word knowledge is a multifaceted construct (Laufer & Nation, 1999) that calls for proper reflection in vocabulary acquisition research. Therefore, many researchers tried to demonstrate the vocabulary knowledge construction (Chapelle, 1994; S. M. Gass, Behney, & Plonsky, 2013; Henriksen, 1999; Laufer & Nation, 1999; Read, 2000).

Chapelle (1994) divided vocabulary ability into three components: the

context of vocabulary, fundamental procedures of vocabulary knowledge, and metacognitive strategies of vocabulary use. On the other hand, some researchers (Henriksen, 1999; Read, 2000) defined vocabulary knowledge with three different aspects: "partial–precise knowledge," "depth of knowledge," and "receptive–productive control," during its gradual development stages (p. 304). Partial–precise knowledge refers to the progressive vocabulary development. As mentioned in Read's (2000) study, breadth and depth of knowledge are the quantity and quality of learners' vocabulary knowledge. Receptive and productive control of vocabulary knowledge was related to its comprehension and production. Previous studies, such as Gass et al. (2013), suggested similar components of lexical knowledge such as "production and reception," "knowledge and control," and "breadth and depth".

All things considered, three exemplary components of vocabulary knowledge were frequently cited: reception and production, vocabulary breadth and depth, and the context of vocabulary use, which are discussed in the following sections in more detail.

2.1.1. Reception and Production

The receptive and productive aspects of vocabulary have been doubtlessly regarded as to exist derived from the two fundamental communication processes: comprehension and production (Nation, 2001). However, there is no

clear-cut way of distinguishing between receptive and productive aspects in word knowledge. Rather, vocabulary knowledge was considered to be gradually developed from receptive to productive phases (Laufer & Paribakht, 1998; Melka, 1997). It is complicated to conceptualize which part of the continuum is occupied by the receptive aspect or that of the productive, and even more intricate to put the absolute threshold where vocabulary is developed from receptive to productive phases (Read, 2000). However, the segregation to put vocabulary knowledge on either stage can be practical (Melka, 1997).

Many researchers coined their definitions in a bid to delineate the term "receptive" and "productive" (Gass et al., 2013; Henriksen, 1999; Meara, 2009; Nation, 2001). Nation (2001), for example, described the "receptive" phase as the process to receive language input and to comprehend its meaning through listening or reading. The "productive" phase, on the other hand, is the procedure to generate language output and deliver a particular message through speaking or writing. The terms "receptive" and "productive" from previous studies were used to entail the receptive and productive facets of language processes and the use of receptive and productive language skills.

Also, the terms "receptive" and "productive" are often described as the corresponding terms, "active" and "passive," which are related to one another. Meara (1990), for instance, delineated "active vocabulary" can be activated through word association. "Passive vocabulary," in contrast, can only be triggered by a receptive stimulus such as reading and listening (Meara, 1990).

The division between the receptive and productive aspects of vocabulary

knowledge entails different facets of vocabulary knowledge such as lexical procedures, language skills, and word associations (Gass et al., 2013). In other words, the distinction between the two aspects is a complex mixture of several aspects of vocabulary knowledge (Henriksen, 1999). However, reception and production themselves, as the primary domains of vocabulary knowledge (Melka, 1997), will be observed in this study to explore the development of vocabulary knowledge.

2.1.2. Vocabulary Breadth and Depth

According to previous research, "breadth" and "depth" of vocabulary knowledge were key issues in language development (S. M. Gass et al., 2013). Milton (2009) defined the "breadth" as a learner's vocabulary size and "depth" as the quality of the learner's lexical knowledge. That is, breadth of word knowledge shows how many words someone knows and depth refers to what they known about those words (Milton, 2009).

Previous research addressed that degrees of knowledge (Melka, 1997; Paribakht & Wesche, 1993; Wesche & Paribakht, 1996), and word association (Meara, 2009; Read, 2000) were eloquently related to breadth and depth of lexical knowledge.

Regarding "breadth" of vocabulary knowledge, several studies were conducted. Goulden, Nation, and Read's (1990) study, for example, indicated

that English native speakers know about 20,000-word families on average. Nation (2006) concluded that English learners are required to know about 8,000-9,000 word families for reading and 6,000-7,000 for speaking. Schmitt (2008) reached a similar conclusion that language learners have to know about 8,000-9,000 word families for reading, and 5,000-7,000 for speaking and listening. Specifically, for written or oral communication, at least 98-99% of vocabulary should be possessed by English language learners (Hsueh-Chao & Nation, 2000).

Doubtlessly, not many students can reach this stage (Nation, 2006). Based on those findings, vocabulary should be strategically selected for vocabulary instruction, especially in EFL/ESL settings, to achieve vocabulary knowledge in a more effective and efficient way.

As part of an endeavor to gauge vocabulary breadth and depth and its receptive and productive aspects representing word knowledge development, Paribakht and Wesche (1993) created a Vocabulary Knowledge Scale (VKS) to evaluate EFL/ESL learners' vocabulary knowledge. The VKS asks learners to self-report their knowledge of each word by responding to the following statements

(1) I have never seen this word.

(2) I have seen this word before, but I don't know what it means.

(3) I have seen this word before, and I think it means _____.

   (Synonym or translation)

(4) I know this word. It means _____. (Synonym or translation)

(5) I can use this word in a sentence.

(Paribakht & Wesche, 1993, p. 15)

This assessment shows that a learner's vocabulary develops from partial to full knowledge through the "semantization" process (Nation, 2001) including the specific nature of this development from word recognition to its productive use in context.

Considering the lexical developmental stages mentioned above, vocabulary breadth and depth should be considered in the vocabulary learning processes (Milton, 2009; Nation, 2013; Nation & Gu, 2007).

2.1.3. Context of Vocabulary Use

A large number of studies detailed earlier asserted that the context in which a word is used makes up a significant part of the lexical ability. (Bachman & Palmer, 1996; Chapelle, 1994; Martinez, 2010).

Chapelle (1994), for instance, regarded the context in which vocabulary is used as one of the three major components of vocabulary knowledge. Bachman and Palmer (1996) considered lexical knowledge as the knowledge of words and their appropriate use in the appropriate context. They believed that the development of vocabulary knowledge calls for the ability to use vocabulary in the right context as well as its incremental gain (Bachman & Palmer, 1996). Martinez (2010) also pointed out that the acquisition of vocabulary knowledge is

the procedure during which a learner deliberates a target word, retrieves its lexical information and uses it in a proper context.

On the whole, it is beneficial to measure diverse aspects of lexical knowledge in order to deal with the complexity of vocabulary learning development. This compensates for a single component of vocabulary knowledge which would hardly capture the dynamic aspects of vocabulary (Gass et al., 2013).

2.2. Research Issues in Vocabulary Instruction

Nation (2001) proposed that effective vocabulary instruction demands decisions to select, sequence and present vocabulary while choosing appropriate tasks in order to integrate lexical progress into communication. As an effort to discover effective vocabulary instruction, research has been conducted on various issues such as vocabulary knowledge development (Anderson & Freebody, 1981; Koda, 1989; Read, 2000), the relationship between vocabulary knowledge and language proficiency (Koda, 1989; Qian, 2002), word frequency (Balota & Chumbley, 1984; Eckerth & Tavakoli, 2012; Hu, 2013; McKeown, Beck, Omanson, & Pople, 1985; Rayner & Duffy, 1986; Rott, 2007), explicit versus implicit learning (Berry & Broadbent, 1987), incidental versus intentional learning (Brown, Waring, & Donkaewbua, 2008; Day, Omura, & Hiramatsu, 1992; Ghabanchi & Ayoubi, 2012; Heidari-Shahreza & Tavakoli, 2012;

Hemmati & Asmawi; Joe, 1998; Laufer & Hulstijn, 2001; Laufer & Rozovski-Roitblat, 2011; Paribakht & Wesche, 1999; Pellicer-Sánchez & Schmitt, 2010; Song & Sardegna, 2014; Srichamnong, 2008; Webb, 2008, 2012), vocabulary assessment (Bachman & Palmer, 1996; Laufer & Nation, 1999; Schmitt, Schmitt, & Clapham, 2001), vocabulary learning strategies, task effect on vocabulary learning (An & Min, 2011; Bolger et al., 2008; Hulstijn & Laufer, 2001; Kim Ji, 2014; S. S. Kim, 2013; Pichette et al., 2011; Ryoo, 2009) and the effect of context on vocabulary learning (An & Min, 2014; Bainbridge, Lewandowsky, & Kirsner, 1993; Bensoussan & Laufer, 1984; Bolger et al., 2008; Carroll & Drum, 1982; Mondria & Wit-de Boer, 1991; Nagy, 1995; Nagy, Anderson, & Herman, 1987; Nagy, Herman, & Anderson, 1985; Schatz & Baldwin, 1986; Stallman, 1991; Sternberg, 1987; Webb, 2008). Among these issues, this study will mainly examine the effects of the task type and context on vocabulary learning.

2.2.1. Effects of Receptive and Productive Tasks on Vocabulary Learning

The two task types, receptive and productive, are commonly assumed to reflect input and output of communication in a number of previous studies (Amiryousefi & Kassaian, 2010; An & Min, 2011; Bao, 2015; De La Fuente, 2002; Folse, 2006; Hazrat, 2015; Jeon & Shin, 2011; S. Y. Kim & Lee, 2008; Laufer, 1998; Lee, 2003; Llach, 2009; Melka, 1997; Mondria & Wiersma,

2004a; Waring, 1997; Webb, 2005). There has been a consensus that both receptive and productive tasks, based on input and output process of communication, contribute considerably to vocabulary learning (An & Min, 2011; Shintani, 2011; Stahl & Fairbanks, 1986).

Among several definitions of receptive and productive vocabulary learning, this study adopted Mondria and Wiersma's (2004) terminology, as follows:

(1) Receptive vocabulary learning is to learn the meaning of an L2 word. Learning a word is going from L2 to L1.

(2) Productive vocabulary learning is to express a concept using an L2 word. Learning a word is going from L1 to L2. (p. 38)

A large number of studies were conducted to discover the efficacy of receptive and productive tasks on learners' vocabulary learning, but the results were rather contradictory (Barcroft, 2004; Choi, 2007; Griffin & Harley, 1996; Hulstijn & Laufer, 2001; S. S. Kim, 2013; Laufer & Hulstijn, 2001; Son, 2007; Waring, 1997; Webb, 2005).

Most research has proved the superiority of the productive task over the receptive task on either immediate vocabulary gain (Pichette et al., 2011), vocabulary retention (Hulstijn & Laufer, 2001) or both (Kim, 2013; Son, 2007; Webb, 2005).

Hulstijn and Laufer (2001) conducted research about EFL students' incidental short-term and long-term vocabulary retention after three different tasks: one productive task (free writing) and two receptive tasks (reading with fill-in and reading only) with various task involvement loads. As predicted,

retention was higher in the productive task compared to the two receptive tasks. It was highest in the composition, lower in the fill-in-the-blank task with reading, and lowest in reading only.

Webb (2005) discovered how Japanese EFL students learned target vocabulary using three glossed sentences and a sentence composition task. Five elements of vocabulary knowledge—"orthography, syntax, association, grammatical functions, and meaning and form" (p. 33)—were assessed. With the sufficient amount of time for task completion, the productive task was more effective for vocabulary gain as well and its retention.

Son (2007) examined Korean university students' immediate vocabulary gain and its retention by comparing one productive task and two receptive tasks with differential task loads and the combination of all three tasks. Corresponding to Hulstijn & Laufer's (2001) research, among a single task, the composition task resulted in the highest scores in immediate and delayed post-tests. However, unlike other results, there was no significant difference between two repetitive tasks with differing involvement loads. This study only proved the differential impact between different task types, productive and receptive, rather than those of involvement loads.

Pichette et al. (2011) investigated the relative effect of reading and writing sentences for ESL French learners' incidental vocabulary learning. Different from the results of the Son (2007) and Webb (2005), which implied the superiority of productive task over the receptive on immediate and delayed tests, immediate recall scores showed superior recall for writing tasks over reading

tasks while delayed recall scores demonstrated no differences between them over time.

Compared to the research results that confirmed the dominance of productive task effect over that of receptive on overall vocabulary knowledge gain and retention, the results of Griffin and Harley, (1996) and Waring's (1997) research proposed that the vocabulary task types are widely influenced by the types of vocabulary knowledge. In other words, the receptive task made learners gain more receptive vocabulary knowledge, whereas the productive task led students to learn more productive vocabulary knowledge.

Some studies even proposed the dominance of receptive tasks over the productive task in vocabulary learning and retention. The results of Webb's (2005) first experiment, within the same limited amount of time, showed that the receptive task was superior to the productive one. Although, as time passed, the superiority of the receptive task disappeared, receptive vocabulary tasks still make up an important part of vocabulary learning, which was shown in Choi's (2007) study. Choi's study partially replicated Webb's (2005) study. Choi (2007) showed that receptive tasks yielded better gains in both receptive and productive vocabulary.

The majority of previous research agreed on the positive effects of productive tasks, whether it is partial or full, on overall vocabulary learning, or at least on productive vocabulary learning. Barcroft's (2004) research, however, showed the opposite results. Barcroft (2004) compared the effects of writing new sentence including target words with those of word-picture repetition on L2

Spanish learners' vocabulary learning. The research findings showed a strong negative effect from the productive task, suggesting that this task can inhibit learning word forms during the initial phases of L2 vocabulary acquisition

Although a large number of studies were conducted, the effect of receptive versus productive tasks on language learners' vocabulary learning is not conclusive (Webb, 2005). Despite the inconsistency of the research results, vocabulary tasks tend to be conducted receptively rather than productively, especially in EFL settings (Kim & Lee, 2008; Webb, 2005). According to Kim and Lee (2008), Korean EFL vocabulary instruction has mainly been conducted using receptive tasks rather than productive ones. This receptive-centered vocabulary instruction would hamper students' output production.

Therefore, investigating the effects of two types of task–receptive and productive–may provide important implications to improve current vocabulary instruction in Korea. In that sense, this study aims to compare the main effects of the differing tasks in order to provide meaningful information to compensate for the inconsistency of the previous research.

2.2.2. Effects of Context on Vocabulary Learning

Vocabulary knowledge is the ability to use general and specific word items in its precise context (Bachman & Palmer, 1996). In authentic communication, a word is generally presented with relevant context, whether it

is written or said. That is, understanding context can be more useful for language learning, especially vocabulary learning (Sternberg, 1987). However, previous studies have revealed rather controversial results on the effect of context.

The effects of context on vocabulary instruction have been consistently investigated through reading the research. Many studies agreed on the positive impact of diverse contextual information on vocabulary learning (An & Min, 2014; J. R. Anderson, 1990; Bolger et al., 2008; Schouten-van Parreren, 1989; Sternberg, 1987; Webb, 2008). Nagy, Herman, and Anderson (1985)' study showed that context lead to small but statistically reliable gains in word knowledge. The incidental learning from context through learners' reading resulted in a substantial vocabulary development during the school years. Schouten-van Parreren's (1989) experiment about comprehension and retention of vocabulary in texts revealed that context with an appropriate level of difficulty is beneficial to vocabulary learning. This study supposed that reading the same words in various context sentences would provide plentiful references to retrieve word meaning.

In line with the previous studies, Webb (2008) also noted a positive effect of context on vocabulary learning after comparing the effects of different context types. He divided Japanese EFL learners into two groups: one with more contextual clues and the other with less contextual clues and made them learn target vocabulary through reading. His research proved the superiority of the more informed context group over the other on the retrieval of vocabulary meaning, but not on that of the retrieval of its form. Rather, the number of

encounters had a greater effect on retrieving the forms of words. Based on his results, Webb calculated that sentence contexts would affect different features of vocabulary knowledge in different ways, which calls for assessing various dimensions of vocabulary knowledge.

However, there have been few studies conducted that examine the sentence contexts (diverse versus same) except for Sternberg (1987), Bolger et al. (2008) and An and Min's (2014) study. Sternberg (1987) demonstrated the possible effect of context variation on vocabulary learning. According to his research, a proper level of context variable helps learners get an overall understanding of the meanings of given words. He found that repetition of the same context sentence alone could not lead to the same favorable result as repetition of multipe context sentences.

Bolger et al. (2008) explored the effect of sentence contexts and use of definitional context on vocabulary learning. A group who repeatedly encountered target words in the same sentences and another group who met the same target words but in different sentences without definition were compared. The result showed that the multiple-context group had higher scores on comprehension of word meanings and on judging whether a newly given word was proper in context. Thus, they argued that the degree of sentence contexts has a significant influence on the learner's vocabulary learning, primarily performed in a receptive way.

An and Min (2014) examined two EFL Korean learner groups with different sentence contexts which practiced target vocabulary through the receptive task.

However, each of these groups had a different sentence contexts ; diverse context or the same context. Participants repeatedly practiced the target words in the given sentences through reading and listening. This research demonstrated that sentence contexts has a statistically meaningful influence on developing vocabulary knowledge as the diverse context group's test results showed better mean scores than those of the single context group.

In contrast, some researchers (Herman et al., 1987; Jenkins, Pany, & Schreck, 1978; Lawson & Hogben, 1996) doubted the effect of context on vocabulary gain and suggested that its efficacy is rather negligible when compared to direct vocabulary instruction. Jenkins, Stein, & Wysocki (1984) argued the rather modest effects of the context variable on vocabulary acquisition are due to the redundancy of cues in the text. Each word presented plenty of contextual clues that did not need to be understood receptively, so learners did not pay selective attention to each word item (Jenkins, Stein, & Wysocki, 1984).

Mondria and Wit-de Boer (1991) reported that guessing the meanings of words through diverse contexts did not show statistically meaningful effects on word gain. Rather, they proposed a negative correlation between contextual guessing and retrieving the meaning of words. This research discovered that the easier it is for learners to guess the meaning of words from their context, the faster they tend to forget them.

Corresponding to the previous research result, Nation and Coady (1988) explained the negative correlation between context and word learning. According to Nation and Coady, language learners seldom focus on the

meanings of individual words when too many contextual clues were given since they can readily comprehend the general messages from context. This may hamper retention of the target words in the end.

File & Adams (2010) compared three ESL university learners groups that focused on taking isolated vocabulary instruction without context sentences, integrated teaching, and incidentally learning vocabulary through context. The group learning words without context sentences achieved better mean scores on vocabulary tests than those with context sentences. Although some words were incidentally learned through reading, the number of words learned was much fewer than expected. They believed that the cognitive load of comprehending context might have hinder vocabulary learning.

Although the role of context on vocabulary learning has been studied a lot, they were mostly conducted in receptive learning settings such as reading and listening, so in most cases, context richness had a significant role in understanding the meanings of target words in receptive settings (Bolger et al., 2008). This calls for research exploring the effect of context on vocabulary learning in productive learning settings.

Also, to compensate for the misleading contextual information of a target word, which was mentioned in previous research, that cast doubt on context effect, both proper definitions and contexts for new words (Stahl, 1986) should be provided for effective vocabulary instruction (Bolger et al., 2008).   .

CHAPTER 3.
METHODOLOGY

This current chapter presents the methods used in this study. Section 3.1 introduces the research design. Section 3.2 discusses the participants. The procedures of the study are described in Section 3.3. Section 3.4 provides details on the instruments regarding the target words and sample sentences used. The treatment of the receptive task group and productive task group are explained in Section 3.5. The word learning assessment methods and their scoring procedures are described in Section 3.6. Section 3.7 describes the data analysis.

3.1. Research Design

A multifactorial design with no control group was implemented in this study. When it comes to the vocabulary treatment, task groups had the significant superiority over control groups in previous studies (An & Min, 2011; Stahl, 1986). Moreover, this study aims to compare differences across four treatment groups, depending on their task types and contexts.

The independent variables (2) were task types (receptive versus productive) and sentence contexts (same versus diverse). The dependent variables (5) were five types of vocabulary tests: the recognition test, the receptive translation test, the passive/active word learning test, and the two productive vocabulary use tests: gap-filling, and word reordering.

3.2. Participants

This study was conducted from June to July 2015. All participants (N=128) were third-grade middle school students from one co-educational middle school (M), located in Sinrim-dong, Gwanak-gu district in Seoul. Most of the learners have had at least five and half years of English education: three years in elementary school, and two and half years in middle school. Only data from students who signed a consent form were used in this study. Four intact classes were chosen based on the mean scores of English mid-term and final exams. These exams had been administrated in the target school during the 1st semester of 2015. Table 3.1 shows the mean scores and the standard deviations of the English scores of the four participating classes. One-way analysis of variance (ANOVA) confirmed the homogeneity of the participating classes ($p = .525$) (F $= .749$, $p > .05$).

Table 3.1

Descriptive Statistics of 4 Participating Classes

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
C 1	29	72.655	24.0071	4.4580	63.523	81.787	18.0	97.5
C 2	30	72.717	22.5773	4.1220	64.286	81.147	23.5	98.0
C 3	28	72.661	22.8559	4.3193	63.798	81.523	18.0	100.0
C4	30	72.717	18.8001	3.4324	65.697	79.737	19.0	95.0
Total	117	72.688	21.8290	2.0181	68.691	76.685	18.0	100.0

Note. The maximum test score was 100; C = Class

One of four treatments with a different combination of task types and contexts (RD, RS, PD, PS) was randomly assigned to each of the four homogeneous classes; RD to Class 1 (32 students), RS to Class 2 (32 students), PD to Class 3 (32 students), and PS to Class 4 (32 students).

To ascertain the effects of each treatment, students who identified more than two items as pre-known words or as re-encountered words during the tests (see Appendix 5) were also excluded from the data analysis (N=11). As a result, 117 students (29 for RD, 30 for RS, 28 for PD, 30 for PS) were selected for data analysis.

3.3. Procedure

The whole process was composed of three sessions; the task, immediate test, and delayed test. The research was conducted in the classes at middle school M under the guidance of the researcher with the help of one English teacher of Korean nationality. Before the experiment, the students were informed of the purpose of this study and then read and signed the consent form (see Appendix 1). Then, an orientation session was held in which the details about the process of the experiment were explained to participants. After the orientation session, the researcher asked students to practice the eight target words in class, which involved one of the following: learning words through a receptive task with diverse context sentences (RD), a receptive task with the same context sentences (RS), a productive task with diverse context sentences (PD), and a productive task with the same context sentences (PS).

Right after the task was done, the students' learning of the target words was measured by five different types of vocabulary tests: recognition, passive word learning, active word learning, and two productive word uses: gap-filling and word reordering. A second test was administered one week after. The tests were conducted in the following order: the active word learning test, the recognition test, the passive word learning test, then the two productive use tests, gap-filling and word reordering. There were three different versions per test that presented items in a random order to avoid any fixed-order effects (Puff, 1982).

3.4. Instruments

This section details the instruments (the target words and sample sentences) implemented in the present study.

3.4.1. Target Words

All The target words in the present study were eight words that were unknown to participants. Students unrelated to the participants of this study selected a total of thirty candidate target words (6 verbs, 5 adjectives, 19 nouns) from a list of Lv 1000 and Lv 2000 words in the Academic Word List (Coxhead, 2000). The students who selected these words have a similar level of English, regarding the mean scores of English mid-term and final exams, as the target students but were excluded from the study in order to make four treatment groups. To ensure the target students' absence of knowledge on the selected target words, students that already knew the meaning of more than two words were excluded from data analysis.

The researcher consulted with the teacher, who had taught the target students for two and half years, and selected eight target words from the thirty candidate words. Words were selected whose meanings did not overlap. Six nouns and two verbs were selected as target words to balance out the students' use of the words in context. Table 3.2 presents the selected words for the experiment.

Table 3.2

List of the Target Words

sacrifice	희생; 희생물
inquire	(…에게) 묻다
wander	거닐다, 돌아다니다
dispose	배치하다, 배열하다
firm	회사, 사무소
recognize	~ 알아보다[알다]
charity	자선[구호] 단체
proclaim	선언[선포]하다

3.4.2. Sample Sentences

In the research, two groups were presented with the target words using one of two different context conditions: diverse context sentences or the same context sentences. Students in the same context group were given only one sample sentence per target word during the task and practiced it three times whereas those in the diverse context group received three different sentences each time they were given a target word (see Appendix 2 for receptive task groups and Appendix 3 for productive task groups).

This sort of grouping aimed to save time to evaluate the quality and quantity of context clues, and represent sentence contexts in a more practical way. This methodology followed previous studies with a similar purpose (Bolger et al., 2008; Gass et al., 1999). The sample sentences were taken from various resources, including the Academic Word List (Coxhead, 2000), Naver online

concordances, and Your Dictionary web resources. The sentences were modified to balance the difficulty of context cues and the length of sentences.

Because different types of context cues are inclined to influence students' word learning differently (Drum & Konopak, 1987), the sample sentences should be chosen with great care to counterbalance the differential cognitive load required to acquire target vocabulary. In this study, the sentence contexts take focus, rather than the contextual clues. Therefore, each sample sentence was reviewed cautiously in order to avoid unexpected learning effects from other variables besides sentence contexts and repeated task.

Sample English sentences were used for all groups. The sentences were translated into Korean as L1 samples for the productive task. As mentioned earlier, the diverse context group was given three sentences per target word while the same context group was assigned a sentence that was repeated three times.

3.5. Treatment

During the task, the participants were given eight words that were unknown to them in the form of handouts and on screen projections. They were asked to practice using them in sentences and check their answers on the screen. A set of eight words were shown three times across two sessions.

3.5.1. The Receptive Task Groups

In the receptive task groups, the students practiced the target words in receptive ways. Students were guided to listen to and read the given sentences and translate L2 sentences (English) into L1 sentences (Korean) (see Appendix 2). At first, the learners were shown the form and sound of a target word. Then, the meaning of the word was given to them. In the following stage, the students were given a sample sentence that included the target word. They were given time to read the sentence, and they were advised to listen and repeat the sentence together. Subsequently, they were asked to translate the given L2 sentence into their L1 equivalent and then check the suggested answer on the screen.

Following the procedure stated above, the receptive groups carried out the treatment task using two different context conditions, as follows:

(1) Receptive Task + Same Context (RS): The students practiced the target words, found in the same sentences, three times each by reading and translating them (L2   L1).

(2) Receptive Task + Diverse Context (RD): The students practiced the target words, found in three different sentences, by reading and translating them (L2   L1).

3.5.2. The Productive Task Groups

In the productive task groups, the students completed the productive tasks by repeating the target words in sentences while writing and speaking (see Appendix 3). The learners were given the L1 meaning of the target word. Then, the target word for the given meaning was shown to the students on screen. They were asked to speak aloud all together. The researcher gave them time to write it down. An L1 sentence was given, and the learners were asked to translate it into the L2 sentence using the target word. After that, students were advised to check the suggested answer on the screen. They are asked to speak the sentence aloud together.

Following the procedure stated above, the productive groups carried out the treatment task using two different context conditions, as follows:

(1) Productive Task + Same Context (PS): The students practiced each target word three times using the same context sentences by writing an L2 sentence with the target word corresponding to the given L1 sentence.

(2) Productive Task + Diverse Context (PD): The students practiced the target words through the three different sentences by translating a given L1 sentence into an L2 sentence.   .

3.6. Assessment

To assess the students' knowledge of the target words, five types of vocabulary test were employed in this study—active word learning test, recognition test, passive word learning test, and two productive use tests: gap-filling, word ordering. These tests aimed to assess different developmental stages of the students' vocabulary knowledge. The vocabulary tests were intended to efficiently evaluate the students' ability to recognize the target word, retrieve its form and meaning, and use them in proper context (Nation & Gu, 2007; Yamashita, 2003). The tests weres also intended to evaluate the students' ability reorder the target word with good syntactic knowledge (Zwarts & Dras, 2007) with regards to the five specific stages of vocabulary knowledge: recognizing new words, getting their form and meanings, and using them in proper context and with appropriate grammar. The assessments were based on and revised from the developmental stages suggested in the VKS (Paribakht & Wesche, 1997) were also designed to better indicate the development of vocabulary knowledge.

This study was conducted to gauge exactly what students learn through vocabulary treatments by evaluating various parts of vocabulary knowledge development. Each vocabulary test has three different versions, in which the order of the target words was arranged in a different way to minimize the effect of repeated task at each time point. Each vocabulary test is described in detail in the following sections.

3.6.1. Active Word Learning Test

The active word learning test assesses the students' knowledge of the form of a target word associated with its meaning. In the test, the meaning of the target word was given, and the students were asked to retrieve the equivalent form. The active word learning test follows the format used in previous studies (Mondria & Wiersma, 2004; Webb, 2005, 2007, 2008). The active word learning test was the first test conducted. Appendix 4 and Appendix 5 illustrate sample active word learning test items for immediate and for delayed test each.

3.6.2. Recognition Test

The recognition test was designed to observe the gradual development of learners' vocabulary knowledge. Unlike other performance-based knowledge tests to assess learners' word knowledge through a given task, the recognition test adopted a self-report format to report the students' level of understanding of the words based on a scale. This arrangement was intended to measure the initial development of students' understanding of word form and meaning. It is expected that this test will serve as an appropriate tool to capture even partial or small progress in learners' knowledge, as in previous studies (Dale, 1965; Read, 2000; Paribakht & Wesche, 1997). The test intends to measure a learner's overall understanding of the target word on a modified version of the VKS

(Paribakht & Wesche, 1997), having learners answer on a four-point Likert-type scale of vocabulary knowledge (see Appendix 2 for immediate and Appendix 5 for delayed). The recognition test consists of eight test items.

3.6.3. Passive Word Learning Test

The passive word learning test measures the students' capability to retrieve the meaning of the target word, which contains the initial stage of vocabulary knowledge when a learner encounters a word (Nation & Gu, 2007). In contrast to the self-reported scaled recognition test, which has a similar purpose, this test aims to measure performance in a more direct way by letting them write down the meaning of the given word. Like the active word learning test, this test format follows that of previous studies (Webb, 2005, 2007, 2008) in which students translated the given L2 word without context into its L1 equivalent.

The recognition test and the passive word learning test were done concurrently to follow the revised format of the VKS from previous studies (Weinfurt, 2000). Appendix 2 and Appendix 5 illustrate a sample of the recognition test and the passive word learning test item.

3.6.4. Two Productive Use Tests: Gap-Filling and Word Reordering

The two productive word use tests measure how to retrieve an appropriate word in the given context (see Appendix 6 for immediate and Appendix 10 for delayed) and how to rearrange the given word clusters with appropriate syntactic and semantic meanings (see Appendix 7 for immediate and Appendix 11 for delayed). In the first test, incomplete sentences were given to students who had to complete them by filling in the proper words in the right context. Previous studies used this type of test as a retrieval cue to measure a learner's productive vocabulary knowledge (Laufer, 1998; Laufer & Rozovski-Roitblat, 2011). They adopted this test so as to evaluate learners' vocabulary use in context and monitor their vocabulary knowledge development that simple translation tests could not offer.

In the second test, a cluster of words, including the target word, were provided to the students to rearrange for meaning using proper grammar. The present study adopted this test from Zwarts & Darts's (2007) research to assess students' grammatical knowledge and observe aspects of vocabulary knowledge development that the direct translation test could not provide.

Some researchers criticize this sort of tests for not reflecting authentic aspects of the production procedure (Milton, 2009). They argue that a more genuine and suitable approach would employ essays to directly measure the students' productive vocabulary knowledge. Regardless of this criticism,

considering the English level of most Korean middle school students, it is almost impossible to make them write essays fluently. The practicality and feasibility of this test method cannot be ignored, and the present study has adopted it as one of the tests used.

3.6.5. Scoring

Two independent raters conducted scoring; the researcher of the present study, with one year of English teaching experience in high school, and another teacher, who has three years of experience teaching English in middle school. In particular, for the productive tests, scores for items showing disagreement were confirmed with a native English teacher. Excluding the self-reported recognition test, Pearson's r was calculated to check inter-rater reliability. The attained values were and 0.987 for the active word learning test, 0.979 for the passive word learning test, 0.999 for the Gap-Filling test and 0.986 for the word reordering test. Due to the very high inter-rater reliability, one of the rater's scores was randomly selected and included in the data analysis. The recognition test is a type of self-reporting measurement. The students' answers to the test items were scored according to the criteria shown in Table 3.3. These criteria are from a modified version of the VKS (Paribakht & Wesche, 1993)

Table 3.3

The Scoring Criteria for the Recognition Test

Score Knowledge Scale Description

0	I have never seen this word before, and I don't know this word at all.
1	I have seen this word before, but I don't know what it means.
2	I have seen this word before, and I think I partially know the meaning of the word.
3	I have seen this word before, and I know the meaning of this word.

The rest of the vocabulary tests are performance-based tests in which the students show their word knowledge by completing given test items. When scoring these tests, one aspect of the students' learning was considered: knowledge of the target form or meaning scored by asking learners to provide the form or meaning of the target word. A maximum of three points were assigned based on the following criteria in Table 3.4.

Table 3.4

Scoring Criteria for Performance-based Tests

Points	Criteria	Points by Criteria
Form/Meaning	correct	3
(3 points)	partially correct	1.5
	wrong	0

For form/meaning correctness, the points given to each answer ranged from 1.5 to 3, graded as shown in Table 3.4. Three points were given for the right answer and one and a half points for a partial or near-right answer. Awarding partial points makes it possible to be more sensitive to incomplete but still meaningful knowledge (Waring & Takaki, 2003).

Recognition Test

In the recognition test, students reported their vocabulary knowledge themselves according to a given scale, shown in Table 3.4. The score for each test item ranges from 0 to 3, so the maximum score for the full test is 24 (8×3).

Passive Word Learning Test

Two Korean English teachers scored the passive word learning tests according to the rubric for performance-based tests (see Table 3.4). Partial points were given to answers including a meaning semantically close to the right

answer, a decision made by the two English experts. The score for each test item ranges from 0 to 3, so the maximum score for the full test is 24 (8×3).

Active Word Learning Test

The active word learning tests scored by two Korean EFL teachers with assistance from one native-English-speaking teacher consistent with the rubric for performance-based tests (see Table 3.4). Partial points were given for spelling errors that did not distort the sounds of words. In particular, they got partial scores with 1) more than half of correct syllables or letters, 2) the correct consonant clusters with wrong vowels or 3) the correct vowels with reverted consonants. For example, a student who misspelled wander as wonder was given 1.5 points for the answer. The score for each test item ranges from 0 to 3, so the maximum score for the full test is 24 (8×3).

Productive Use Tests

Two Korean EFL teachers mentioned above scored the productive use tests. No partial points were awarded for the gap-filling test. In the word reordering test, partial points were given only if students put the target word in the right position but had a mistake in placing the remaining parts. The maximum total score for all types of tests used in this study was 24 (8×3) respectively.

3.7. Data Analysis

To adequately explore the research questions, analysis was conducted on each of the four main study groups to investigate the effects of task types and sentence contexts on the sum of five vocabulary test scores and individual test scores. Statistical analysis was implemented using SPSS for Windows (v. 22.0) to verify the research questions; how task types and sentence contexts involve vocabulary gains and retention while vocabulary knowledge development.

First, a set of two-way analysis of variance (ANOVA) was employed with task types and sentence contexts as independent variables and the total vocabulary test scores as a dependent variable. Univariate between-group analysis was followed. Second, a set of two-way Multivariate Analysis of Variance (MANOVA) was employed with task types and sentence contexts as independent variables and the five types of vocabulary tests as dependent variables. Univariate between-group analysis was followed.

CHAPTER 4.
RESULTS AND DISCUSSION

This chapter describes the results of the statistical analysis of the test scores and discusses the findings. Section 4.1 reports the sum of the five test scores and a discussion of the immediate post-test and delayed post-test depending on the assigned task type and sentence contexts. The effect of these two factors on the specific outcomes of each test item and their discussion of the immediate post-test and the delayed post-test are described in Section 4.2.

4.1. The Effects of Task Type and Sentence Contexts on the Overall Immediate Vocabulary Learning and Retention

To investigate the impact of the task type and sentence contexts on general vocabulary gains and its retention, the participating students were divided into one of four treatment groups. Each group was assigned one of two task types and one of two contexts (RS, RD, PS, PD) and all groups took the immediate post-test and delayed test.

To analyze the effects of the task type and sentence contexts on the overall vocabulary learning, the sum of five vocabulary test scores were analyzed. Table 4.1 summarizes the overall means and the standard deviations of the test scores,

according to the four groupings with a combination of different task types and
contexts. The data is then represented in chart format in Figure 4.1. .

Table 4.1

Descriptive Statistics of Immediate Test

Task	Context	Mean	Std. Deviation	N
	Same	76.6500	27.04216	30
Receptive	Diverse	75.2759	30.73928	29
	Total	75.9746	28.67542	59
	Same	97.2333	26.29925	30
Productive	Diverse	85.7143	31.34794	28
	Total	91.6724	29.17349	58
	Same	86.9417	28.40985	60
Total	Diverse	80.4035	31.20846	57
	Total	83.7564	29.85768	117

Note. The scores are the sum of five vocabulary test scores; the maximum score is 120.

Estimated Marginal Means of Whole Test Score

Figure 4.1

Overall Test Scores by Task and Context on Immediate Test

As shown in Table 4.1 and Figure 4.1, the productive groups had significant superiority over the receptive groups for vocabulary learning in the immediate tests. In sentence contexts, however, although the same context groups showed slightly higher mean scores than the diverse context ones. No observable difference was noticed between the two different context groups. Regarding Figure 4.1, no interaction between the two variables was represented. Therefore, in the immediate test, the PS and PD groups were ahead of RD and RS groups but differences between the same task groups were marginal. In particular, the mean scores of the two receptive groups were practically the same.

In order to verify the statistically significant differences between each variable, a set of two-way analysis of variance (ANOVA) were conducted. As seen in Table 4.2, below, the task type, not the context, had significant main effects on the initial learning of the target vocabulary.

Table 4.2

Effects of Task and Context on the Immediate Test

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	7031.443	1	7031.443	8.430	.004	.069
Context	1214.597	1	1214.597	1.456	.230	.013
Task * Context	751.983	1	751.983	.902	.344	.008
Error	94254.949	113	834.115			

p<.05

In the immediate test score analysis, the primary impact of the task was shown, $F(1, 115) = 8.430$, $p = .004$, $\eta2 = .069$, but the statistically meaningful impact of sentence contexts was not, $F(1, 115) = 1.425$, $p = .235$, $\eta2 = .230$. There was no significant interaction effect between the task and the context in

the immediate test, F (1, 115) = .902, p = .344, $\eta2$ = .008.

According to the result of the immediate test, the productive groups had are remarkable superiority over the receptive groups for vocabulary learning in the immediate tests. In sentence contexts, however, no observable difference was noticed between the two different context groups. Thus, the task type, and not sentence contexts, influenced immediate word gain. To be specific, the productive task is more beneficial for immediate vocabulary learning than the receptive task but whether the task is provided in the same context or in diverse contexts does not meaningfully affect immediate vocabulary learning.

The productive groups outperformed the receptive groups in vocabulary retention as well, as shown below in Table 4.3 and Figure 4.2. Compared to the immediate word gain, the productive task proved more beneficial when it comes to maintaining words. No observable difference was shown between the two different context groups, but the interaction between the two variables is shown in Figure 4.2. This means that sentence contexts may affect vocabulary retention differently depending on the type of the task.

Table 4.3

Descriptive Statistics of Delayed Test

Task	Context	Mean	Std. Deviation	N
	Same	52.2667	23.91246	30
Receptive	Diverse	56.1724	33.13912	29
	Total	54.1864	28.63472	59
	Same	89.2500	28.73054	30
Productive	Diverse	72.3393	31.44780	28
	Total	81.0862	31.00129	58
	Same	70.7583	32.16397	60
Total	Diverse	64.1140	33.05164	57
	Total	67.5214	32.62934	117

Note. The scores are the sum of five vocabulary test scores; the maximum score is 120

Figure 4.2

Overall Test Scores by Task and Context on Delayed Test

Looking at the delayed test scores, the PS group was still ahead of others, followed by PD (see Figure 4.3 and Table 4.3). The same context affected vocabulary learning more positively than diverse contexts during the productive task session. The context effect, however, were reversed when students did the receptive task; the RD outperformed the RS for word retention. Overall, the PS task had more durability of overall vocabulary knowledge than the PD task. On the other hand, the RD task had stronger durability than the RS task in the delayed posttest. That is, the sentence contexts conjugating the task type may not sufficiently influence immediate word gain but may influence its retention

Estimated Marginal Means of Vocabulary_Test

*cf. TC= Task type * Context*

Figure 4.3

Overall Test Scores by TC Immediate and Delayed Test

Table 4.4

Effects of Task and Context on the Delayed Test

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	20640.565	1	20640.565	23.807	.000	.174
Context	1235.749	1	1235.749	1.425	.235	.012
Task * Context	3166.106	1	3166.106	3.652	.059	.031
Error	97971.906	113	867.008			

p<.05

The statistical significance of the test result differences was checked through a set of two-way analysis of variance(ANOVA).The task type had the main effects: F (2, 114) = 30.52, p = 0.000,η2 =.174, but context did not. The p-value of the interaction effect was .059, which did not meet the statistically meaningful level (p<.05), but showed a certain power of the sentence contexts on learners' word retention.

To sum up, task type was a factor that significantly affected vocabulary learning, and this ultimately shows support for previous studies (Hulstijn & Laufer, 2001; Kim, 2013; Webb, 2005) that argued the superiority of the productive task over the receptive task on vocabulary knowledge development.

The productive task group produced higher overall scores than the receptive task group in both the immediate and the one-week delayed test. In the delayed test, the overall test score of the productive task groups decreased much less than those of the receptive task groups, which indicates that productive tasks have greater potential to help retain word knowledge.

Sentence contexts, on the other hand, were not a statically significant factor affecting vocabulary learning throughout the experiment. Even though no statistically meaningful differences were found between the groups, the descriptive statistics show that, when compared to the diverse context groups, the same context groups had a slightly higher mean score overall. This result is in contradiction to those of Bolger et al.'s (2008) and An and Min's (2014) study, which proposed the superiority of diverse contexts over the same context in vocabulary learning.

Regarding word retention, the interaction effect between two variables was shown. In delayed test, like the immediate test, the PS group was ahead of other groups. However, the scores of receptive groups significantly decreased. The result of the RS group, in particular, saw a very large decrease, so this group demonstrated the lowest ability to retain word knowledge. That is, the same context had a positive effect on the productive task but not on the receptive task.

The findings from overall test scores revealed that, depending on the assigned task type, effective context might differ. This suggests that the productive task was more demanding to the learners so that it made them focus more on the vocabulary itself when they repeatedly wrote the same sentence

rather than writing down different sentences. The receptive task, on the other hand, was less challenging, so students could benefit from several context sentences, with their focus on the target vocabulary itself.

4.2. The Effects of Task Type and Sentence contexts on the Immediate Learning and Retention of Specific Vocabulary Knowledge

In this section, the scores of the recognition, passive word learning, active word learning, gap-filling and word reordering tests are treated as five dependent variables and analyzed to investigate whether the task type and sentence contexts affect them differently.

The results of each vocabulary test are explained in greater detail in Sections 4.2.1 to 4.2.5.

4.2.1. Recognition Test

Table 4.5, Table 4.6 and Figure 4.4 represent the descriptive statistics of the recognition test results in the immediate and delayed post-test. The productive task groups showed slightly higher mean scores than the receptive groups on the immediate test, but the difference was negligible (see Table 4.5). Concerning the

sentence contexts, the difference between the same context group and diverse context group was marginal, and the mean scores of the groups were practically the same. Regarding the combination of the two factors (see Table 4.5 and Figure 4.4), even though the PS group had slightly higher mean scores than the other groups, no observable difference was noticed among the four different treatment groups for immediate word gain. Overall, the effect of task type and sentence contexts on the immediate gain of word recognition was subsidiary.

Table 4.5

Descriptive Statistics of the Immediate Recognition Test

Task	Context	Mean	Std. Deviation	N
	Same	19.2000	5.47345	30
Receptive	Diverse	19.5172	5.77322	29
	Total	19.3559	5.57631	59
	Same	21.3333	3.57514	30
Productive	Diverse	20.0357	5.70563	28
	Total	20.7069	4.72770	58
	Same	20.2667	4.70797	60
Total	Diverse	19.7719	5.69466	57
	Total	20.0256	5.19526	117

Note. The maximum score for each test is 24.

Figure 4.4

Recognition Test Scores by TC on Immediate and Delayed Tests

As shown below in Table 4.6, the productive task groups were more capable of recognizing the target words than the receptive groups. Concerning sentence contexts, the difference between the delayed test scores of the same context groups and those of diverse context groups was trivial. As represented in Table 4.6 and Figure 4.4, the scores of the PS and PD groups outperformed RD and RS groups, but no observable difference was noticed among the same task groups. That is, the task type affects retaining word recognition knowledge, not the sentence contexts

Table 4.6

Descriptive Statistics of the Delayed Recognition Test

Task	Context	Mean	Std. Deviation	N
	Same	15.7000	5.71839	30
Receptive	Diverse	16.7586	6.68503	29
	Total	16.2203	6.18136	59
	Same	20.5667	3.77545	30
Productive	Diverse	18.5000	5.88469	28
	Total	19.5690	4.97401	58
	Same	18.1333	5.39449	60
Total	Diverse	17.6140	6.30973	57
	Total	17.8803	5.83858	117

Note. The maximum score for each test is 24.

The statistical significance of the test result differences was checked through a set of two-way MANOVA tests, shown below in Table 4.7 and Table 4.8. The results of the recognition test scores in Table 4.7 revealed that neither task nor context, served as a between-subjects variable, had statistically evocative influence over the immediate recognition test scores (Task, $F(5,109) = 1.903$, p = .171, $\eta2 = .017$; Context, $F(5,109) = .260$, p = .611, $\eta2 = .002$).

However, in the delayed test, shown in Table 4.8, the effect of task type was statistically meaningful; $F(5,109) = 10.161$, p = .002, $\eta2 = .083$, but still no

meaningful effect of context was observed; $F(5,109) = .236$, $p = .628$, $\eta 2 = .002$. As to interaction, there was no significant interaction effect in both tests (immediate, $F(5,109) = .706$, $p = .403$, $\eta 2 = .006$; delayed, $F(5,109) = 2.273$, $p = .134$, $\eta 2 = .020$).

Table 4.7

Effect of Task and Context on Immediate Recognition Test

Effect	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	51.380	1	51.380	1.903	.171	.017
Context	7.023	1	7.023	.260	.611	.002
Task * Context	19.054	1	19.054	.706	.403	.006

p<.o5

Table 4.8

Effect of Task and Context on Delayed Recognition Test

Effect	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	319.049	1	319.049	10.161	.002	.083
Context	7.425	1	7.425	.236	.628	.002
Task * Context	71.366	1	71.366	2.273	.134	.020

p<.o5

Because word recognition ability is the initial stage of vocabulary knowledge, according to the VKS (Paribakht & Wesche, 1993), no differences were observed depending on difference in task types and context. As for retention for this knowledge, however, the productive task had more durability than the receptive one (Hulstijin & Laufer, 2001).

4.2.2. Passive Word Learning Test

Table 4.9, Table 4.10 and Figure 4.5 show the descriptive statistics of the passive word learning test scores in the immediate test and the delayed test.

Table 4.9

Descriptive Statistics of the Immediate Passive Word Learning Test

Task	Context	Mean	Std. Deviation	N
Receptive	Same	17.5000	7.35199	30
	Diverse	17.4828	7.00444	29
	Total	17.4915	7.12118	59
Productive	Same	19.8000	5.37812	30
	Diverse	17.4643	7.07359	28
	Total	18.6724	6.30897	58
Total	Same	18.6500	6.49074	60
	Diverse	17.4737	6.97535	57
	Total	18.0769	6.72797	117

Note. The maximum score for each test is 24.

Regarding the task type, the productive groups received slightly higher passive word learning test scores than the receptive groups on the immediate test, but the difference was marginal. Sentence contexts also showed minimal difference between the same context groups and the diverse context groups. Table 4.9 and Figure 4.5 show that the PS group was a little ahead of others and no apparent mean differences across the remaining three groups were shown. That is, task type and sentence contexts do not affect immediate retrieval of word meaning.

Figure 4.5

Passive Word Learning Test Scores by TC on Immediate and Delayed Tests

Table 4.10

Descriptive Statistics of the Delayed Passive Word Learning Test

Task	Context	Mean	Std. Deviation	N
Receptive	Same	10.8167	6.12300	30
	Diverse	12.5690	8.35563	29
	Total	11.6780	7.29594	59
Productive	Same	18.4000	6.28956	30
	Diverse	15.0000	6.90411	28
	Total	16.7586	6.75592	58
Total	Same	14.6083	7.24516	60
	Diverse	13.7632	7.70674	57
	Total	14.1966	7.45329	117

Note. The maximum score for each test is 24.

In the delayed test, the outcomes turned out to be different. The productive task groups outperformed the receptive groups with considerable gaps. Meanwhile, sentence contexts did not represent meaningful differences among the two different context groups. In the delayed post-test shown in Table 4.10 and Figure 4.5, the PS group kept ahead of all of the other groups. Unlike the immediate posttest, there were apparent mean differences across the remaining groups. The mean score of the PS group decreased much less than other groups. Receptive task groups, however, especially the RS group, showed a sharper

decline amongst the four groups.

When univariate analyses were conducted to probe whether the differences between groups were statistically meaningful, the results showed that neither task type ($F(5,109) = .838$, $p = .362$, $\eta2 = .007$) nor context ($F(5,109) = .891$, $p = .347$, $\eta2 = .008$) had a significant main effect on the initial word learning. This data is shown below in Table 4.11.

Table 4.11

Effect of Task and Context on Immediate Passive Word Learning Test

Effect	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	38.033	1	38.033	.838	.362	.007
Context	40.452	1	40.452	.891	.347	.008
Task * Context	39.275	1	39.275	.865	.354	.008

p<.o5

In the delayed test score analysis, the main effect of the task was shown, $F(5,109) = 15.120$, $p = .000$, $\eta2 = .118$, but no statistically meaningful effect of sentence contexts was shown, $F(5,109) = 19.837$, $p = .524$, $\eta2 = .004$. These results are shown below in Table 4.12. As for interaction, there was an

interaction effect but only in the delayed test (F(5,109) = 4.002, p = .048, η2= .034). That is, the sentence contexts worked differently corresponding to the types of tasks. The same context functioned positively on the productive task while it worked negatively on the receptive task. Thus, the receptive task with the same context had a weak power to retain vocabulary knowledge, especially inferring the meaning of the target word.

Table 4.12

Effect of Task and Context on Delayed Passive Word Learning Test

Effect	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	732.755	1	732.755	15.120	.000	.118
Context	19.837	1	19.837	.409	.524	.004
Task * Context	193.961	1	193.961	4.002	.048	.034

p<.o5

Contrary to Waring's (1997) argument that the receptive task outperformed the productive one when assessing receptive vocabulary knowledge, there was no difference in passive word knowledge gain through two different types of the task. Furthermore, regarding retention of passive word knowledge, the

productive task showed more potential than the receptive task. The results can be supported by the level of vocabulary knowledge suggested by the VKS (Paribakht & Wesche, 1997). The productive task is dealing with deeper word process and had a better impact on retention of receptive word knowledge

4.2.3. Active Word Learning Test

Table 4.13, Table 4.14 and Figure 4.6 display the descriptive statistics of active word learning test scores in immediate and delayed post-tests. The productive groups showed a considerably higher mean score than the receptive groups for both the immediate word learning (Table 4.13) and its retention (Table 4.14) in the active word learning. This test aimed at retrieving word forms, which were included in productive vocabulary knowledge, so it is doubtless that the productive task was more beneficial than the receptive task in immediate word knowledge gain and its retention. Regarding sentence contexts, no observable difference was noticed between the two different context groups in the immediate post-test.

The results of the active word learning tests among four treatment groups are displayed in Table 4.13, Table 4.14 and Figure 4.6. The PS group consistently earned the highest mean score, and the PD, RD and RS groups followed in either the immediate or the delayed posttest. This proved that there were adverse effects of sentence contexts on task types over the mean score of the active word

learning test. Even though learners repeatedly practiced the given task, the PS significantly outperformed the other groups, while, the RS group received the lowest active word learning test score.

Table 4.13

Descriptive Statistics of the Immediate Active Word Learning Test

Task	Context	Mean	Std. Deviation	N
	Same	6.3000	4.44623	30
Receptive	Diverse	9.9310	7.97174	29
	Total	8.0847	6.62685	59
	Same	19.6000	6.06346	30
Productive	Diverse	16.9286	7.95673	28
	Total	18.3103	7.10685	58
	Same	12.9500	8.52996	60
Total	Diverse	13.3684	8.64608	57
	Total	13.1538	8.55219	117

Note. The maximum score for each test is 24.

Table 4.14

Descriptive Statistics of Delayed Active Word Learning Test

Task	Context	Mean	Std. Deviation	N
Receptive	Same	2.0500	2.34649	30
	Diverse	6.5690	6.97153	29
	Total	4.2712	5.60428	59
Productive	Same	17.2500	6.71687	30
	Diverse	13.7679	8.43656	28
	Total	15.5690	7.72978	58
Total	Same	9.6500	9.14446	60
	Diverse	10.1053	8.47356	57
	Total	9.8718	8.78899	117

Note. The maximum score for each test is 24.

Figure 4.6

Recognition Test Scores by TC on Immediate and Delayed Tests

The statistical significance of differences in the test results were checked through a set of two-way multivariate analysis of variance (MANOVA). The task was found to have statistically significant impacts on both the immediate test ($(5,109) = 66.330$, p = .000, $\eta2$= .370), as shown in Table 4.15 and the delayed test ($(5,109) = 87.195$, p = .000, $\eta2$= .436) as represented in Table 4.16. In terms of context, on the other hand, there was no observable impact on both tests. As to interaction, there was interaction effect in both tests (immediate, $F(5,109) = 6.395$, p = .013, $\eta2$= .054; delayed, $F(5,109) = 11.126$, p = .001, $\eta2$= .090).

Table 4.15

Effect of Task and Context on

the Immediate Active Word Learning Test

Effect	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	3010.221	1	3010.221	66.330	.000	.370
Context	6.728	1	6.728	.148	.701	.001
Task * Context	290.223	1	290.223	6.395	.013	.054

p<.o5

Table 4.16

Effect of Task and Context on

the Delayed Active Word Learning Test

Effect	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	3665.766	1	3665.766	87.195	.000	.436
Context	7.855	1	7.855	.187	.666	.002
Task * Context	467.748	1	467.748	11.126	.001	.090

p<.o5

Overall, regarding the form retrieval of target words, repeating several context sentences seems to be more efficient with the productive task, in comparison to repeating the same context sentences with the receptive task. As the RS group was the least effective, repeatedly reading new words in the same context sentences does not seem to be a practical way to learn the word forms. Consequently, the results suggest that the productive task was more challenging for the students so it made them focus more on vocabulary spelling itself when repeatedly writing the same sentence rather than writing down different sentences. The receptive task, on the other hand, was less demanding, so using several context sentences helped them retrieve vocabulary from its corresponding meaning.

4.2.4. Two Productive Use Tests

This section shows the results and discussion of the two productive use tests conducted in this study. The result and discussion of the gap-filling test and word reordering test are described.

4.2.4.1. Gap-Filling Test

The descriptive statistics of the independent variables for the gap-filling test were shown below in Table 4.17 and Table 4.18. When it comes to either the

task type or the sentence contexts, no observable differences were shown between the two groups per each variable regarding immediate word gain (see Table 4.17). The gap of mean scores between groups was not apparent on the immediate test.

As shown in Figure 4.7, the PS group achieved the highest average score, followed by the RS, PD and RD groups on the immediate test. The difference between the RS and PD groups was trivial and the mean scores of the two groups were the same. Neither the type of the task nor the sentence contexts meaningfully affected the immediate use of words in the proper context sentence.

Table 4.17

Descriptive Statistics of the Immediate Gap-Filling Test

Task	Context	Mean	Std. Deviation	N
Receptive	Same	14.4500	7.70395	30
	Diverse	12.5172	8.63804	29
	Total	13.5000	8.16373	59
Productive	Same	17.1000	7.31248	30
	Diverse	14.1429	8.75051	28
	Total	15.6724	8.10542	58
Total	Same	15.7750	7.56577	60
	Diverse	13.3158	8.65439	57
	Total	14.5769	8.17287	117

Note. The maximum score for each test is 24.

Table 4.18

Descriptive Statistics of the Delayed Gap-Filling Test

Task	Context	Mean	Std. Deviation	N
Receptive	Same	9.1500	6.34381	30
	Diverse	8.0690	8.14130	29
	Total	8.6186	7.23994	59
Productive	Same	15.8833	7.37152	30
	Diverse	10.7143	7.84978	28
	Total	13.3879	7.97641	58
Total	Same	12.5167	7.61687	60
	Diverse	9.3684	8.03926	57
	Total	10.9829	7.95026	117

Note. The maximum score for each test is 24.

Figure 4.7

Gap-Filling Test Scores by TC on Immediate and Delayed Tests

The In the delayed test, in contrast, there was a considerable effect of both the type of the task and the variable of the context on word retention (see Table 4.18). The test scores of the productive groups significantly exceeded those of the receptive groups. This means the productive group showed more statistically meaningful durability than the receptive group for word knowledge regarding its productive use in the proper context. It implies that the receptive task made the context information of vocabulary harder to recall from memory after one week. In terms of sentence contexts, the same context group outperformed the diverse context group on both tests. The gap between groups was apparent in delayed

post-test.

To be brief, the productive task using the same context sentences was more efficient at helping students retrieve word forms and use them in context in comparison to other treatment. In mastering target words for use in relevant contexts, it seems that practicing the words productively through writing activities repeated in the same sentence is useful in the overall learning process.

Table 4.19 and Table 4.20 show the univariate results of a two-way MANOVA in the immediate test and delayed test. In regard to task type and sentence contexts, the differences were not statistically significant (Task, $F(5,109) = 2.032$, $p = .157$, $\eta2 = .018$; Context, $F(5,109) = 2.658$, $p = .106$, $\eta2 = .023$) in the immediate test. In the delayed test, however, both task type and sentence contexts had statistically meaningful effects on the test scores (Task, $F(5,109) = 11.596$, $p = .001$, $\eta2 = .093$; Context, $F(5,109) = 5.150$, $p = .025$, $\eta2 = .044$).

Table 4.19

Effect of Task and Context on Immediate Gap-Filling Test

Effect	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	133.570	1	133.570	2.032	.157	.018
Context	174.708	1	174.708	2.658	.106	.023
Task * Context	7.667	1	7.667	.117	.733	.001

p<.o5

Table 4.20

Effect of Task and Context on Delayed Gap-Filling Test

Effect	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	642.677	1	642.677	11.596	.001	.093
Context	285.419	1	285.419	5.150	.025	.044
Task * Context	122.106	1	122.106	2.203	.141	.019

p<.o5

Regarding the VKS (Paribakht & Wesche, 1997), the productive use of vocabulary was the most difficult part of learning vocabulary knowledge, so it is more helpful to conduct a deeper level task, which refers to the productive task. The results about context effect in this study yielded conflicting results from An and Min's (2014) previous study. The same context groups, and not the diverse context groups, had statistically meaningful effect on vocabulary knowledge regarding its contextual use. It is because students may focus on its contextual use better when the same context sentences were used repeatedly rather than when the diverse context sentences, using too many cues, were given (Hu & Nation, 2012; Nation & Coady, 1988).

4.2.4.2. Word Reordering Test

Table 4.21 and Table 4.22 show the descriptive statistics of the task type and context for the word reordering tests for immediate word gain and its retention. Regarding the task type, the productive group showed slightly higher test scores than the receptive group but the gap between the two task groups was minimal on the immediate test (see Table 4.21). When it comes to sentence contexts, however, the same context group considerably outperformed the diverse context group for immediate gain of the grammatical use of vocabulary.

Table 4.21

Descriptive Statistics of Immediate Word Reordering Test

Task	Context	Mean	Std. Deviation	N
Receptive	Same	19.2000	5.54853	30
	Diverse	15.8276	7.05372	29
	Total	17.5424	6.50417	59
Productive	Same	19.4000	6.24003	30
	Diverse	17.1429	6.22399	28
	Total	18.3103	6.28129	58
Total	Same	19.3000	5.85503	60
	Diverse	16.4737	6.63286	57
	Total	17.9231	6.37871	117

Note. The maximum score for each test is 24.

Table 4.22

Descriptive Statistics of Delayed Word Reordering Test

Task	Context	Mean	Std. Deviation	N
Receptive	Same	14.5500	6.58152	30
	Diverse	12.2069	7.60347	29
	Total	13.3983	7.13888	59
Productive	Same	17.1500	6.84200	30
	Diverse	14.3571	6.90928	28
	Total	15.8017	6.95786	58
Total	Same	15.8500	6.78377	60
	Diverse	13.2632	7.28692	57
	Total	14.5897	7.12227	117

Note. The maximum score for each test is 24.

In the immediate test, the mean scores of the PS and RS groups and those of the PD and RD groups were practically the same (see Figure 4.8). The outcome shows that the context, not the task type, has a significant effect on the test scores on the immediate test. Both task groups with the same context performed better than those with diverse context.

As for retention of lexical knowledge measured by the word reordering test, the productive group showed relatively higher test scores than the receptive group (see Table 4.22). Regarding the sentence contexts, the same context

groups significantly surpassed the diverse context groups for retention of the productive use of vocabulary

Figure 4.8

Word Reordering Test Scores by TC on Immediate and Delayed Tests

Figure 4.8 shows that the test scores of the receptive groups declined sharply on the delayed post-test. The PS group was still ahead of the other three groups, but the gap between the PS and RS groups significantly expanded. Although the mean score of the RS group was much higher than that of the PD group in the immediate test, there were no visible differences in the mean scores between the two groups on the delayed test.

The univariate results of the word reordering post-tests shown in Table 4.23 prove that the differences between groups were statistically significant in regard to sentence contexts (F(5,109) = 5.865, p = .017, η2= .049), but not task type (F(5,109) = .425, p = .516, η2= .004), as a between-subjects variable on the immediate test. In addition, there was no significant interaction effect between the task and the context in the immediate test (F(5,109) = .230, p = .632, η2= .002).

Table 4.23

Effect of Task and Context on Immediate Word Reordering Test

Effect	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	16.776	1	16.776	.425	.516	.004
Context	231.558	1	231.558	5.865	.017	.049
Task * Context	9.088	1	9.088	.230	.632	.002

*p<.o5

Although, in the delayed post-test, task type (F(5,109) = 3.374, p = .069, η2= .029) had a statistically meaningful effect on test scores, but it had a very limited effect (see Table 4.24). Sentence contexts (F(5,109) = 3.944, p = .049,

η2= .034), on the other hand, had a statistically meaningful effect on test scores. In addition, there was no significant interaction effect between the task and the context in the delayed test (F(5,109) = .030, p = .862, η2= .000).

Table 4.24

Effect of Task and Context on Delayed Word Reordering Test

Effect	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Task	164.871	1	164.871	3.374	.069	.029
Context	192.733	1	192.733	3.944	.049	.034
Task * Context	1.478	1	1.478	.030	.862	.000

*p<.o5

According to the VKS (Paribakht & Wesche, 1997), the productive use of vocabulary knowledge was the most difficult stage, so it is more helpful to conduct the productive task rather than the receptive task. This study, however, failed to prove the positive effect of the productive task in either word gain or its retention. However, comparing the effects of task on word gain, those on its retention was much more dominant. In terms of context effect, the same context groups had a statistically meaningful effect on vocabulary knowledge regarding

its use in both immediate and delayed tests, which is opposed to An and Min's (2014) findings. It is because students may focus on its productive use better when they repeatedly read the same context rather than read the different sentences all the time.

CHAPTER 5.
CONCLUSION

This chapter is composed of three sections. Section 5.1 summarizes the findings of the present study. The pedagogical implications of this study regarding English vocabulary education are discussed in Section 5.2. Finally, Section 5.3 describes the limitations of the present study and makes suggestions for the further research

5.1. Major Findings

This study investigated how task type (receptive versus productive) and sentence contexts (the same context versus diverse contexts) contribute to lexical knowledge development of Korean middle school students.

The first research question looked into the effect of receptive versus productive task and sentence contexts on overall vocabulary learning and retention. The impact of the each task and sentence contexts on the five specific components of vocabulary knowledge development were investigated in the second research question.

Task type was a factor that significantly affected vocabulary learning, and this ultimately show support for the previous studies (Hulstijn & Laufer, 2001; Kim, 2013; Webb, 2005) that argued the superiority of the productive task over the receptive task for developing vocabulary knowledge. The productive task

group produced higher overall scores compared to the receptive task group in both the immediate and the one-week delayed test. In the delayed test, the overall test scores of the productive task group decreased much less than those of the receptive task group, which indicates that productive task had more durability to retain word knowledge (Webb, 2005).

Sentence contexts themselves were not a statistically significant factor affecting vocabulary learning throughout this experiment. However, they played a crucial role when interacting with the type of task, especially in word retention. In other words, the more effective context might differ in relation to the assigned task types in this study. The same context groups were positively affected on the productive task but not on the receptive task. Thus, the PS group showed predominance in word gain and a much more statistically significant power in word retention among the four treatment groups. The RS group, on the other hand, revealed the lowest ability to retain word knowledge.

Since the productive task was more demanding for the students, it made them focus more on vocabulary itself when repeatedly writing the same sentence rather than writing down different sentences. The receptive task, on the other hand, was less challenging, so students could benefit from several context sentences and focus on the target vocabulary itself. However, this study only investigated the gain and retention of target words. Regarding additional word gain, the multiple context groups may have superiority over the same context groups.

The findings from the individual analysis of the five vocabulary tests

demonstrated that the productive task had significant superiority over the receptive task in vocabulary learning. In particular, the productive task had statistically considerable power to retain several phases of vocabulary knowledge, with the exception of the word reordering test. The context itself did not have much influence on the lexical knowledge development from word recognition to passive and active word knowledge. When combined with the task, the sentence contexts had a strong effect on vocabulary learning in passive and active word learning tests, especially in word form and meaning extraction. In line with the overall findings, the same context groups, and not diverse contexts, had a statistically meaningful effect on vocabulary knowledge. Thus, the PS group scored the highest among the four treatment groups in the five types of tests respectively. On the other hand, with the exception of the two productive use tests, the RS group scored the lowest. The results explain the cross effect between task type and context for word recognition and retrieval of word meaning and form.

The retention of word knowledge measured by the productive use of vocabulary tests, on the other hand, was influenced fundamentally by sentence contexts rather than task type. Conflicting with results from a previous study (An & Min, 2014), the same context groups had a statistically meaningful effect on retention of the vocabulary knowledge regarding its contextual use. Using the same context sentence may help students focus more on the words contextual use compared with the diverse context sentences that demands higher cognitive loads.

In general, considering the effect of task type and context on overall

vocabulary learning, the task effect was substantial, while that of context was not. Examining the two variables together, however, shows that the effect of context was different from the task types, especially for retrieving word and meaning connection. The productive task, when completed within the same context, was always ahead of other treatment groups. The receptive task, when completed within the same context, usually recorded the lowest grade with the exception of the two productive use tests. However, in the productive use tests, which demand contextual knowledge of target vocabulary, the using the same context lead to better results since it helped learners focus on contextual information of the target words and was not strongly related to the task that was done.

5.2. Pedagogical Implications

Based on the major findings described in section 5.1, this study presents the following pedagogical implications on L2 vocabulary learning.

1) Implementing more productive tasks than receptive ones in the classroom context may be effective for vocabulary learning, especially regarding its retention. This is because it enables learners to gain and retain much more productive vocabulary knowledge as well as a little more or at least a similar level of receptive vocabulary knowledge.

2) Rather than just assuming vocabulary instruction through diverse contexts is always the most effective, the effectiveness of the context should be carefully judged based on other variables such as task types, students' English proficiency and so on. This study proposed the possibility of interaction effect between task types and sentence contexts on vocabulary retention. Teachers should consider the cognitive load and difficulty level of each sentence contexts of target vocabulary before designing, modifying, comparing, choosing, or implementing vocabulary tasks.

5.3. Limitations and Suggestions

First, this research was conducted with 117 Korean middle school students living in Gwanak-gu, Seoul, which makes it difficult to generalize the major findings for a larger population. The effect of vocabulary treatment may fluctuate according to students' age, their residence, their average language ability, or their motivations to learn English. Further research is suggested to employ a sufficient number of students from diverse backgrounds, randomly sampled for multiple variables so the finings can be more applicable to a larger population.

Second, this study did not take students' proficiency levels into account. Since the cognitive load of vocabulary task treatment may affect research results, different results could be revealed in the effects of task type and sentence contexts on vocabulary learning if students are classified into different

proficiency groups.

Third, the word items utilized in the vocabulary task treatment were limited in number, level and parts of speech. Only eight target words, including six nouns and two verbs, were chosen out of thirty-word items in the Lv 1000 and Lv 2000 word list (Academic Word List, Coxhead, 2000), but the limitation of word selection made it difficult to generalize the significant findings. Therefore, future studies should contain a larger and more diverse list of words from different levels using different parts of speech.

Fourth, the number of sentence contexts in the vocabulary task treatment was also limited. Compared with task effects, the effect of sentence contexts was relatively marginal, which can be derived from the limited number of sentence contexts. In addition, the multiple context groups failed to show its superiority when performing productive task. The result can differ with sufficient number of sentence contexts. In future studies, it would be beneficial to provide more context sentences during vocabulary instruction to determine if context will have a larger effect.

Fifth, there was a problem in assessing the productive use of word knowledge. As an alternative to the free composition evaluating the productive use of word knowledge in the last phase of the VKS (Paribakht & Wesche, 1997), this study implemented two productive use tests: gap-filling and word rearrangement. However, there remains a limitation because these assessments guarantee to evaluate students' productive use of word knowledge using the proper context. Moreover, in the assessment of the word rearrangement test, a

partial point was given when the target word was put in the right position, but other word cluster was not. However, if a student put a target word in the right place, it means that he/she had grammatical or contextual knowledge of the word, so it might not be reasonable to give a partial point rather than a full point. Therefore, a more sophisticated and segmented assessment will be necessary to accurately assess students' productive use of word knowledge.

In spite of these limitations, the findings from this study propose meaningful information about the effect of productive versus receptive task and sentence contexts on Korean middle school learners' English vocabulary learning.

REFERENCES

Amiryousefi, M., & Kassaian, Z. (2010). The effects of reading only vs. reading plus enhancement activities on vocabulary learning and production of Iranian pre-university students. *English Language Teaching, 3*(2), p94.

An, H.-S., & Min, C. K. (2011). The Effects of Receptive vs. Productive Vocabulary Instruction. *Journal of the Korea English Education Society, 10*(2), 1-22.

An, H.-S., & Min, C. K. (2014). The Effects of Sentence contexts on Vocabulary Learning. *Studies in English Education, 19*(2), 127-149

Anderson, J. R. (1990). *Cognitive psychology and its implications*: WH Freeman/Times Books/Henry Holt & Co.

Anderson, R., & Freebody, P. (1981). Vocabulary knowledge. In 1. T. Guthrie (Ed.), Comprehension and teaching: Research reviews (pp. 77-117). Newark. *DE: International Reading Association*.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests* (Vol. 1): Oxford University Press.

Bainbridge, J. V., Lewandowsky, S., & Kirsner, K. (1993). Context effects in repetition priming are sense effects. *Memory & Cognition, 21*(5), 619-626.

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human perception and*

performance, 10(3), 340.

Bao, G. (2015). Task type effects on English as a Foreign Language learners' acquisition of receptive and productive vocabulary knowledge. *System, 53*, 84-95.

Barcroft, J. (2004). Effects of sentence writing in second language lexical acquisition. *Second Language Research, 20*(4), 303-334.

Bensoussan, M., & Laufer, B. (1984). Lexical guessing in context in EFL reading comprehension. *Journal of Research in Reading, 7*(1), 15-32.

Berry, D. C., & Broadbent, D. E. (1987). The combination of explicit and implicit learning processes in task control. *Psychological research, 49*(1), 7-15.

Bolger, D. J., Balass, M., Landen, E., & Perfetti, C. A. (2008). Context variation and definitions in learning the meanings of words: An instance-based learning approach. *Discourse Processes, 45*(2), 122-159.

Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental Vocabulary Acquisition from Reading, Reading-While-Listening, and Listening to Stories. *Reading in a Foreign Language, 20*(2), 136-163.

Carroll, B., & Drum, P. (1982). *Effects of context in facilitating unknown word comprehension.* Paper presented at the New inquiries in reading research and instruction. Rochester, NY: National Reading Conference.

Chapelle, C. A. (1994). Are C-tests valid measures for L2 vocabulary research? *Second Language Research, 10*(2), 157-187.

Choi, J.-Y. (2007). *The Effect of Receptive and Productive Tasks on Lexical*

Knowledge Development. (Unpublished master's thesis), Michigan State University.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 213-238.

Day, R. R., Omura, C., & Hiramatsu, M. (1992). Incidental EFL vocabulary learning and reading. *Reading in a Foreign Language, 7*, 541-541.

De La Fuente, M. J. (2002). Negotiation and oral acquisition of L2 vocabulary. *Studies in second language acquisition, 24*(01), 81-112.

Drum, P. A., & Konopak, B. C. (1987). Learning word meanings from written context. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 73-87). Hillsdale, NJ: Lawrence Erlbaum Associates.

Eckerth, J., & Tavakoli, P. (2012). The effects of word exposure frequency and elaboration of word processing on incidental L2 vocabulary acquisition through reading. *Language Teaching Research, 16*(2), 227-252.

File, K. A., & Adams, R. (2010). Should Vocabulary Instruction Be Integrated or Isolated? *TESOL Quarterly, 44*(2), 222-249.

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

Gass, S., Mackey, A., Alvarez-Torres, M. J., & Fernandez-Garcia, M. (1999). The effects of task repetition on linguistic output. *Language Learning, 49*(4), 549-581.

Gass, S. M., Behney, J., & Plonsky, L. (2013). *Second language acquisition: An introductory course*. New York: Routledge.

Ghabanchi, Z., & Ayoubi, E. S. (2012). Incidental vocabulary learning and recall by intermediate foreign language students: The influence of marginal glosses, dictionary use, and summary writing. *Journal of International Education Research (JIER), 8*(2), 85-96.

Griffin, G., & Harley, T. A. (1996). List learning of second language vocabulary. *Applied Psycholinguistics, 17*(04), 443-460.

Hazrat, M. (2015). The Effects of Task Type and Task Involvement Load on Vocabulary Learning. *Waikato Journal of Education, 20*(2).

Heidari-Shahreza, M. A., & Tavakoli, M. (2012). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition by Iranian EFL Learners. *The Language Learning Journal*(ahead-of-print), 1-16.

Hemmati, P., & Asmawi, A. B. Incidental vocabulary Learning and Retention through Reading a Graded Reader among Irannan EFL Learners. *The Online Journal of New Horizons in Education*, 114.

Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in second language acquisition, 21*(02), 303-317.

Hsueh-Chao, M. H., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language, 13*(1), 403-430.

Hu, H.-c. M. (2013). The Effects of Word Frequency and Contextual Types on Vocabulary Acquisition from Extensive Reading: A Case Study. *Journal of Language Teaching and Research, 4*(3), 487-495.

Hu, H.-c. M., & Nassaji, H. (2012). Ease of inferencing, learner inferential strategies, and their relationship with the retention of word meanings

inferred from context. *Canadian modern language review, 68*(1), 54-77.

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

Jenkins, J. R., Stein, M. L., & Wysocki, K. (1984). Learning vocabulary through reading. *American educational research journal, 21*(4), 767-787.

Jeon, E., & Shin, Y. (2011). Receptive and productive vocabulary learning using a word list in L2. *Elementary English Education, 17*(1), 395-416.

Joe, A. (1998). What effects do text-based tasks promoting generation have on incidental vocabulary acquisition? *Applied linguistics, 19*(3), 357-377.

Kim, S. S. (2013). *Analyses of Receptive and Productive Korean EFL Vocabulary: Computer-based Vocabulary Learning Program.* Arizona State University.

Kim, S. Y., & Lee, S. H. (2008). Learning strategies and instructional approaches for Korean EFL learners' productive vocabulary development. *English Language Teaching, 20*(4), 237-259.

Koda, K. (1989). The effects of transferred vocabulary knowledge on the development of L2 reading proficiency. *Foreign language annals, 22*(6), 529-540.

Laufer, B. (1998). The development of passive and active vocabulary in a second language: same or different? *Applied linguistics, 19*(2), 255-271.

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied linguistics,*

22(1), 1-26.

Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language testing, 16*(1), 33-51.

Laufer, B., & Paribakht, T. S. (1998). The relationship between passive and active vocabularies: Effects of languagelearning context. *Language Learning, 48*(3), 365-391.

Laufer, B., & Rozovski-Roitblat, B. (2011). Incidental vocabulary acquisition: The effects of task type, word occurance and their combination. *Language Teaching Research*, 1362168811412019.

Lee, H. (2003). The effects of production and comprehension for focus on form and second language acquisition. *Journal of the Applied Linguistics Association of Korea, 19*(2), 41-68.

Llach, M. d. P. A. (2009). The effect of reading only, reading and comprehension, and sentence writing in lexical learning in a foreign language:: some preliminary results. *Revista española de lingüística aplicada*(22), 9-34.

Martinez, M. E. (2010). *Learning and cognition: The design of the mind*. Boston, MA: Pearson College Division.

McKeown, M. G., Beck, I. L., Omanson, R. C., & Pople, M. T. (1985). Some effects of the nature and frequency of vocabulary instruction on the knowledge and use of words. *Reading research quarterly*, 522-535.

Meara, P. (2009). *Connected words: Word associations and second language vocabulary acquisition* (Vol. 24): John Benjamins Publishing.

Melka, F. J. (1997). Receptive vs. productive aspects of vocabulary. In N.

Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition, and pedagogy* (pp. 84-102). Cambridge: Cambridge University Press.

Milton, J. (2009). *Measuring second language vocabulary acquisition* (Vol. 45): Multilingual Matters.

Mondria, J.-A., & Wiersma, B. (2004a). Receptive, productive, and receptive+ productive L2 vocabulary learning: What difference does it make. In B. P & L. B (Eds.), *Vocabulary in a second language: Selection, acquisition, and testing* (pp. 79-102). Philadelphia, PA: John Benjamins.

Mondria, J.-A., & Wiersma, B. (2004b). Receptive, productive, and receptive+ productive L2 vocabulary learning: What difference does it make. *Vocabulary in a second language: Selection, acquisition, and testing*, 79-100.

Mondria, J.-A., & Wit-de Boer, M. (1991). The Effects of Contextual Richness on the Guessability and the Retention of Words in a Foreign Language1. *Applied linguistics, 12*(3), 249-267.

Nagy, W. E. (1995). On the role of context in first-and second-language vocabulary learning: Champaign, Ill.: University of Illinois at Urbana-Champaign, Center for the Study of Reading.

Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American educational research journal, 24*(2), 237-270.

Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading research quarterly*, 233-253.

Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Boston: Heinle & Heinle Publishers.

Nation, I. S. P. (2001). *Learning vocabulary in another language*: Ernst Klett Sprachen.

Nation, I. S. P. (2013). Teaching & learning vocabulary.

Nation, I. S. P., & Coady, J. (1988). Vocabulary and reading. *Vocabulary and language teaching, 97*, 110.

Nation, I. S. P., & Gu, P. Y. (2007). *Focus on vocabulary*. Sydney: NCELTR,Macquarie University.

Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. *Second language vocabulary acquisition: A rationale for pedagogy*, 174-200.

Paribakht, T. S., & Wesche, M. (1999). Reading and "incidental" L2 vocabulary acquisition. *Studies in second language acquisition, 21*(02), 195-224.

Paribakht, T. S., & Wesche, M. B. (1993). Reading comprehension and second language development in a comprehension-based ESL program. *TESL Canada journal, 11*(1), 09-29.

Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental Vocabulary Acquisition from an Authentic Novel: Do" Things Fall Apart"? *Reading in a Foreign Language, 22*(1), 31-55.

Pichette, F., De Serres, L., & Lafontaine, M. (2011). Sentence reading and writing for second language vocabulary acquisition. *Applied linguistics*, amr037.

Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning, 52*(3), 513-536.

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14*(3), 191-201.

Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language testing, 10*(3), 355-371.

Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge university press

Rott, S. (2007). The effect of frequency of input-enhancements on word learning and text comprehension. *Language Learning, 57*(2), 165-199.

Ryoo, Y.-s. (2009). Effects of Two Types of Vocabulary Practice: Receptive and Productive. *Foreign languages education, 16*(1), 79-99.

Schatz, E. K., & Baldwin, R. S. (1986). Context clues are unreliable predictors of word meanings. *Reading research quarterly*, 439-453.

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language testing, 18*(1), 55-88.

Schouten-van Parreren, C. (1989). Vocabulary learning through reading: Which conditions should be met when presenting words in texts. *AILA review, 6*(1), 75-85.

Son, J. (2007). The Effects of Vocabulary Exercises on EFL Vocabulary Learning and Retention. *English Language Education, 13*(4), 167-192.

Song, J., & Sardegna, V. G. (2014). EFL learners' incidental acquisition of English prepositions through enhanced extensive reading instruction. *RELC Journal, 45*(1), 67-84.

Srichamnong, N. (2008). Incidental EFL Vocabulary Learning: The Effects of Interactive Multiple-Choice Glosses: Design.

Stahl, S. A. (1986). Three principles of effective vocabulary instruction. *Journal of Reading*, 662-668.

Stallman, A. C. (1991). *Learning vocabulary from context: Effects of focusing attention on individual words during reading.* University of Illinois at Urbana-Champaign.

Sternberg, R. J. (1987). Most vocabulary is learned from context. In M. McKeown & M. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. pp. 89-105). Hillsdale, NJ: Lawrence Erlbaum Associates.

Waring, R. (1997). A study of receptive and productive learning from word cards. *Studies in Foreign Languages and Literature, 21*(1), 94-114.

Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader. *Reading in a Foreign Language, 15*(2), 130-163.

Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in second language acquisition, 27*(01), 33-52.

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied linguistics, 28*(1), 46-65.

Webb, S. (2008). The Effects of Context on Incidental Vocabulary Learning. *Reading in a Foreign Language, 20*(2), 232-245.

Webb, S. (2012). Repetition in Incidental Vocabulary Learning *The Encyclopedia of Applied Linguistics*: John Wiley & Sons, Inc.

Weinfurt, K. P. (2000). Repeated measures analysis: ANOVA, MANOVA, and HLM. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding more multivariate statistics* (pp. 317-361). Washington, DC: American Psychological Association.

Wesche, M., & Paribakht, T. S. (1996). Assessing Second Language Vocabulary Knowledge: Depth Versus Breadth. *Canadian modern language review, 53*(1), 13-40.

Wilkins, D. A. (1972). *Linguistics in language teaching*: E. Arnold, 1973.

Zimmerman, C. B. (1997). Historical trends in second language vocabulary instruction. *Second language vocabulary acquisition*, 5-19

APPENDIX 1. Consent Form

연구참여자용 설명서 및 동의서

연구 과업 명: 수용적/생산적 과업과 문장 문맥이 한국 중학교 학생들의 영어 어휘 보유와 지식에 미치는 영향 **(The Effects of Receptive/Productive Tasks and Sentence Contexts on English Vocabulary Retention and Knowledge of Korean Middle School Students)**

연구 책임자명: 김이경 (서울대학교 사범대학 영어교육과 석사과정, 학생)

본 연구는 과업 유형, 문맥의 다양성, 그리고 그 조합이 한국 중학생 영어 학습자의 어휘학습에 미치는 영향에 대해 알아보는 연구입니다. 귀하는 한국인 영어 학습자로서 본 연구의 대상에 적합하다고 판단되기 때문에 이 연구에 참여하도록 권유 받았습니다. 이 연구를 수행하는 서울대학교 소속의 연구원, <u>김이경</u>,은 연구에 대한 모든 설명을 해주고 실험에 대한 절차를 책임지고 진행할 것입니다. 귀하가 본 실험에 참여하기 위한 의사를 결정하기 전에 본 연구의 수행목적과 내용에 대해 이해하는 것이 중요합니다. 다음 제시사항을 자세히 읽어보신 후 참여 의사를 밝혀주시길 바라며, 필요에 따라 가족이나 선생님께 의논 드려 보시고 결정해주십시오. 내용을 모두 숙지하고 질문이 있다면 담당 연구원에게 질문해주십시오. 담당 연구원이 자세하게 설명해줄 것입니다.

1. 이 연구는 왜 실시합니까?

이 연구는 과업 유형, 문맥의 다양성이 한국 중학생 영어 학습자의 어휘학습에 미치는 영향을 알아보기 위해서 실시합니다.

2. 얼마나 많은 사람이 참여합니까?

영어를 외국어로 배우는 중 3 한국인 영어학습자 120 명 (미성중학교 총 4 개 반)이 연구에 참여할 것입니다.

3. 만일 연구에 참여하면 어떤 과정이 진행됩니까?

만일 귀하가 참여의사를 밝혀 주시면 본 연구에 들어가기에 앞서 4 개의 반의 동질성을 검사하기 위해서 중간고사 영어 내신 성적이 사용될 예정입니다. 성적 정보는 학생 개인정보 없이 숫자로만 제공될 것입니다. 본 연구는 본교 영어 수업의 일환으로 진행될 것입니다. 한 수업 당 45 분씩 총 3 회 차로 진행되며 연구 과정은 크게 총 3 단계 (어휘 과업, 사후 테스트, 지연 사후 테스트)로 진행될 것입니다.

1) 어휘 과업 단계는 총 2 회 차에 걸쳐 진행될 것입니다. 과업의 종류 (수용적 과업(R)/ 생산적 과업(P))와 문맥의 다양성(다양(D)/ 동질(S)) 따라 총 4 개의 반 (RD, RS, PD, PS)으로 구성되며 귀하가 속한 반에 따라 각기 다른 과업을 수행하게 될 것입니다. 수용적 과업(R)은 주어진 영어 문장을 한국어로 해석하는 것이며 생산적 과업(P)은 주어진 한국 문장을 영어로 작문 하는 것입니다. 또한 문맥이 다양(D)한 집단은 해당 어휘에 대한 다양한 예문으로 과업을 하고 문맥이 동질(S)한 집단은 해당 어휘에 대한 같은 예문으로 과업을 진행하게 될 것입니다.

2) 사후 테스트는 2 번째 수업이 끝난 직후에 진행될 것입니다. 두 차례의 수업 시수가 끝난 후 과업의 종류와 문맥의 다양성이 학생들의 어휘 학습에 미치는 영향을 보기 위해 사후 테스트가 진행될 것입니다.

3) 1 주일 후 영어 수업 시간에 어휘의 보유 정도를 알아보기 위하여 지연 사후 테스트를 실시할 것입니다.

4. 연구 참여 기간은 얼마나 됩니까?

약 3 주 동안 일주일에 1 번씩 총 3 회 한 회당 45 분씩 참여하도록 요청받을 것입니다.

5. 참여 도중 그만두어도 됩니까?

예, 귀하께서 실험 참여에 불편함을 느낀다면 언제든지 어떠한 불이익 없이 그만 둘 수 있습니다. 만일 귀하께서 연구에 참여하시는 것을 중단하고 싶다면 담당 연구원에게 즉시 말씀해주십시오.

6. 부작용이나 위험요소는 없습니까?

본 연구는 기존에 진행되는 본교 영어 수업의 일부로써 진행되고 동질성을 위해 쓰이는 중간고사 성적은 학생 정보 없이 오로지 점수로만 제공되기 때문에 안전에 대한 위협이 없을 것으로 예상됩니다. 연구에 대한 자료는 오직 연구만을 위하여 사용하고 결코 외부에 노출하거나 다른 용도로 사용하지 않을 것이며 이 연구결과를 작성할 때 참여자들의 이름을 모두 익명으로 표기하여 제시할 것입니다. 또한 연구 참여 도중 발생하는 문제나 불편함이 있다면 즉각적으로 담당 연구원(김이경)에게 말씀해주십시오. 바로 불편사항을 처리하겠습니다. 이외에도 연구 참여 도중 발생할 수 있는

부작용이나 위험 요소에 대한 질문이 있으면 담당 연구원(김이경)에게 즉시 문의해 주십시오.

7. 이 연구에 참여시 참여자에게 이득이 있습니까?

귀하가 이 연구에 참여하는데 있어서 직접적인 이득은 없습니다. 그러나 귀하가 제공하는 정보는 과업 유형, 문맥의 다양성, 그리고 그 조합이 한국 중학생 영어 학습자의 어휘학습에 미치는 영향을 이해하는 데 도움이 될 것입니다.

8. 만일 이 연구에 참여하지 않는다면 불이익이 있습니까?

귀하는 언제든지 본 연구에 참여하지 않을 자유가 있습니다. 또한, 귀하가 본 연구에 참여하지 않아도 귀하에게는 어떠한 불이익도 없습니다.

9. 연구에서 얻은 모든 개인 정보의 비밀은 보장됩니까?

개인정보관리책임자는 서울대학교 소속 연구원 김이경 (010-62565-3083)입니다. 저는 이 연구를 통해 얻은 모든 개인 정보의 비밀 보장을 위해 최선을 다할 것입니다. 이 연구에서 얻어진 개인 정보가 학회지나 학회에 공개 될 때 귀하의 개인 정보는 사용되지 않을 것입니다. 그러나 만일 법이 요구하면 귀하의 개인정보는 제공될 수도 있습니다. 또한 모니터 요원, 점검 요원, 생명윤리심의위원회는 연구 참여자의 개인 정보에 대한 비밀 보장을 침해하지 않고 관련규정이 정하는 범위 안에서 본 연구의 실시 절차와 자료의 신뢰성을 검증하기 위해 연구 결과를 직접 열람할 수 있습니다. 귀하가 본 동의서에 서명하는 것은, 이러한 사항에 대하여 사전에 알고 있었으며 이를 허용한다는 동의로 간주될 것입니다.

10. 이 연구에 참가하면 대가가 지급됩니까?

귀하의 연구 참여시 감사의 뜻으로 소정의 간식이 지급될 예정입니다.

11. 연구에 대한 문의는 어떻게 해야 됩니까?

본 연구에 대해 질문이 있거나 연구 중간에 문제가 생길 시 다음 연구 담당자에게 연락하십시오.

이름: 김 이경 전화번호: 010-6256-3083

만일 어느 때라도 연구참여자로서 귀하의 권리에 대한 질문이 있다면 다음의 서울대학교 생명윤리심의위원회에 연락하십시오.

서울대학교 생명윤리심의위원회 (SNUIRB) 전화번호: 02-880-5153

동 의 서

1. 나는 이 설명서를 읽었으며 담당 연구원과 이에 대하여 의논하였습니다.

2. 나는 위험과 이득에 관하여 들었으며 나의 질문에 만족할 만한 답변을 얻었습니다.

3. 나는 이 연구에 참여하는 것에 대하여 자발적으로 동의합니다.

4. 나는 이 연구에서 얻어진 나에 대한 정보를 현행 법률과 생명윤리심의위원회 규정이 허용하는 범위 내에서 연구자가 수집하고 처리하는데 동의합니다.

5. 나는 담당 연구자나 위임 받은 대리인이 연구를 진행하거나 결과 관리를 하는 경우와 보건 당국, 학교 당국 및 서울대학교 생명윤리심의위원회가 실태 조사를 하는 경우에는 비밀로 유지되는 나의 개인 신상 정보를 직접적으로 열람하는 것에 동의합니다.

6. 나는 언제라도 이 연구의 참여를 철회할 수 있고 이러한 결정이 나에게 어떠한 해도 되지 않을 것이라는 것을 압니다.

7. 나의 서명은 이 동의서의 사본을 받았다는 것을 뜻하며 연구 참여가 끝날 때까지 사본을 보관하겠습니다.

_____ _____ _____

연구참여자 성명 서 명 날짜 (년/월/일)

_____ _____ _____

동의서 받은 연구원 성명 서 명 날짜 (년/월/일)

_____ _____ _____

연구책임자 성명 서 명 날짜 (년/월/일)

_____ _____ _____

법정 대리인 성명(참여자와 관계) 서 명 날짜 (년/월/일)

APPENDIX 2. Receptive Task

반: _____ 이름: _____

★주어진 영어 단어의 뜻을 쓰고 주어진 문장을 우리말로 해석해보세요

sacrifice		wander	
He made a **sacrifice** of himself to save his town**.**		She **wander**ed aimlessly around the streets.	
firm		charity	
He works for an aircraft **firm.**		Many **charities** sent money to the victims.	
inquire		dispose	
I will **inquire** about how to get there		She **disposed** books in order**.**	
recognize		proclaim	
I could not **recognize** my old friend**.**		The president **proclaimed** a state of emergency.	

★주어진 영어 단어의 뜻을 쓰고 주어진 문장을 우리말로 해석해보세요

sacrifice		wander	
A war involves the **sacrifice** of many lives**.**		We **wander**ed back towards the car..	

firm		charity	
The accounting **firm** audited the company		She does a lot of work for **charity**.	

inquire		dispose	
I **inquire**d about the reason of his long absence.		The DVDs are **dispose**d in alphabetical order**.**	

recognize		proclaim	
I **recognize** the need for safety.**.**		He **proclaim**ed her a traitor.	

★주어진 영어 단어의 뜻을 쓰고 주어진 문장을 우리말로 해석해보세요

sacrifice		wander	
He helped them at the **sacrifice** of himself		Those sheep **wander** all over the place.	
firm		**charity**	
I am not a member of the **firm.**		Any money that is left over will go to **charity**.	
inquire		**dispose**	
I will **inquire** into what happened.		He **dispose**d a fleet in a straight line**.**	
recognize		**proclaim**	
You can **recognize** this tune**.**		The citizens **proclaim**ed him as their king.	

APPENDIX 3. Productive Task

반: _____ 이름: _____

★주어진 뜻에 맞는 영단어를 쓰고 주어진 문장을 영작하여 쓰세요

	희생; 희생물		거닐다, 돌아다니다
그는 그의 마을을 지키기 위하여 스스로 **희생**을 했다.		그녀는 거리를 정처 없이(aimlessly) **돌아다녔다.**	
	회사, 사무소		자선[구호] 단체
그는 항공(aircraft)**회사**에서 일한다		많은 **구호** 단체들이 그 희생자들에게 돈을 보냈다	
	(…에게) 묻다		배치하다, 배열하다
나는 그곳에 가는 방법을 **물어보겠다**		그녀는 책을 순서대로 **배치했다.**	
	~ 알아보다[알다]		선언[선포]하다
나는 옛 친구를 **알아볼** 수 없었다.		대통령이 국가(state) 비상사태 (emergency)를 **선포했다.**	

★주어진 뜻에 맞는 영단어를 쓰고 주어진 문장을 영작하여 쓰세요

	희생; 희생물		거닐다, 돌아다니다
전쟁은 많은 생명의 **희생**을 수반한다(involve).		우리는 천천히 **거닐며** 다시 차 있는 쪽으로 갔다.	
	회사, 사무소		자선[구호] 단체
그 회계(accounting) **사무소**는 그 회사(company)의 회계 감사를 했다(audit).		그녀는 많은 **자선 단체** 활동을 하였다	
	(…에게) 묻다		배치하다, 배열하다
나는 그의 오래 부재의 이유를 **물었다**.		DVD들은 알파벳 순으로 **배열되어 있다.**	
	~ 알아보다[알다]		선언[선포]하다
난 안전의 필요성을 **안다**.		그는 그녀를 반역자(traitor)라고 **선포했다.**	

★주어진 뜻에 맞는 영단어를 쓰고 주어진 문장을 영작하여 쓰세요

	희생; 희생물		거닐다, 돌아다니다
그는 자신의 **희생**으로 그들을 도왔다.		저 양들은 사방을 **돌아다닌다**.	
	회사, 사무소		자선[구호] 단체
저는 그 **회사** 사람이 아닙니다		남은(left over) 돈은 **자선 단체**에 보낼 것이다	
	(…에게) 묻다		배치하다, 배열하다
나는 무슨 일이 있었는지 **물어볼 거야**		그는 함대(fleet)를 일렬로 **배치했다.**	
	~ 알아보다[알다]		선언[선포]하다
너는 이 곡조(tune)를 **알 수 있다.**		국민은 그를 왕으로 **선포하였다**.	

APPENDIX 4. Active Word Learning Test; Immediate

반 _____이름: _____

★주어진 우리 말에 해당하는 영어 단어를 쓰세요. 일부만 써도 좋으니 최대한 기억나는 대로 써주세요

1	거닐다, 돌아다니다	
2	희생; 희생물	
3	배치하다, 배열하다	
4	~ 알아보다[알다]	
5	선언[선포]하다	
6	(...에게) 묻다	
7	자선[구호] 단체	
8	회사, 사무소	

APPENDIX 5. Recognition Test and

Passive Word Learning Test; Immediate

반 _____이름: _____

★주어진 영어 단어를 아는 정도에 따라 0-3에 동그라미(o) 치세요.

☞ 각 숫자는 아래와 같은 단어 지식의 정도를 의미합니다.

0	전에 한번도 본적이 없고 모르는 단어다
1	전에 본적은 있지만 의미를 모른다
2	전에 본적이 있고 그 의미를 대충 짐작한다고 생각한다
3	전에 본적이 있고 그 의미를 안다 ★ 3을 택할 경우 맨 오른쪽에 그 영어 단어에 해당하는 우리말 뜻을 쓰세요.

1	recognize	0	1	2	3	
2	sacrifice	0	1	2	3	
3	dispose	0	1	2	3	
4	wander	0	1	2	3	
5	firm	0	1	2	3	
6	proclaim	0	1	2	3	
7	charity	0	1	2	3	
8	inquire	0	1	2	3	

이 수업을 통해 배운 단어가 아닌 알고 있던 단어가 있다면 체크해주세요!

□ recognize □ sacrifice □ dispose □ wander □ proclaim □ charity □ inquire □ firm

APPENDIX 6. Gap-Filling Test; Immediate

반 _____이름: _____

★문맥에 맞도록 [보기]에서 알맞은 단어를 선택하여 빈칸에 쓰세요.

[보기]

---

sacrifice, adopt, firm, inquire, wander, raise,

dispose, proclaim, pour, recognize, charity

---

1. I will _____ for the shoes at the department store.

2. Don't _____ around alone after midnight.

3. She did not _____ me when she saw me.

4. He is working for an engineering _____.

5. The local _____ will raise money for the poor.

6. She _____ed her books in order.

7. My grandfather made a _____ of his life in the World War Ⅱ.

8. Sir Winston Churchill was _____ed honorary U.S. citizen.

APPENDIX 7. Word Reordering Test; Immediate

반 _____이름: _____

★주어진 단어를 재배열하여 문법적으로 의미 있고 뜻이 통하는 문장을 만드세요

1. disposed/ his soldiers /he/ for the war

   _____

2. him /recognize / I /did / not/ at once

   _____

3. a sacrifice/ of her life / for her family/ made / she

   _____

4. a law / they / for / firm/work

   _____

5. to the school/ how / I / about/ inquired/to get

   _____

6. wandering/ she/ the streets/ around / is

   _____

7. its independence/ the new government/ proclaimed/ in Venezuela

   _____

8. all his money/ to charity/ sent/he

   _____

APPENDIX 8. Active Word Learning Test; Delayed

반 _____이름: _____

★주어진 우리 말에 해당하는 영어 단어를 쓰세요. 일부만 써도 좋으니 최대한 기억나는 대로 써주세요

1	(...에게) 묻다	
2	자선[구호] 단체	
3	회사, 사무소	
4	거닐다, 돌아다니다	
5	희생; 희생물	
6	배치하다, 배열하다	
7	~ 알아보다[알다]	
8	선언[선포]하다	

APPENDIX 9. Recognition Test and

Passive Word Learning Test; Delayed

반 _____이름: _____

★주어진 영어 단어를 아는 정도에 따라 0-3에 동그라미(o) 치세요.

☞ 각 숫자는 아래와 같은 단어 지식의 정도를 의미합니다.

0	전에 한번도 본적이 없고 모르는 단어다
1	전에 본적은 있지만 의미를 모른다
2	전에 본적이 있고 그 의미를 대충 짐작한다고 생각한다
3	전에 본적이 있고 그 의미를 안다 ★ 3을 택할 경우 맨 오른쪽에 그 영어 단어에 해당하는 우리말 뜻을 쓰세요.

1	inquire	0	1	2	3	
2	charity	0	1	2	3	
3	recognize	0	1	2	3	
4	sacrifice	0	1	2	3	
5	dispose	0	1	2	3	
6	wander	0	1	2	3	
7	firm	0	1	2	3	
8	proclaim	0	1	2	3	

APPENDIX 10. Gap-Filling Test; Delayed

<div align="right">반 _____이름: _____</div>

★문맥에 맞도록 [보기]에서 알맞은 단어를 선택하여 빈칸에 쓰세요.

[보기]

---

sacrifice, adopt, firm, inquire, wander, raise,

dispose, proclaim, pour, recognize, charity

---

1. I work at a law _____

2. I will made a _____ of my life to my country

3. Don't _____ around late at night.

4. He will _____ me at once.

5. He _____ed liberty throughout all the land.

6. The concert will raise money for local _____es.

7. She _____ed her clothes and shoes.

8. I will _____ for the book at the bookstore

   .

APPENDIX 11. Word Reordering Test; Delayed

반 _____이름: _____

★주어진 단어를 재배열하여 문법적으로 의미 있고 뜻이 통하는 문장을 만드세요

1. a sacrifice/ for her children/ made / she/ of her happiness

   _____

2. I /an engineering / for / firm/work

   _____

3. the way / I / about / to the station/ inquired.

   _____

4. recognize / I /did / the animal /not

   _____

5. he/ to wander/ the streets/ around / likes

   _____

6. all his property/ donated/ to charity/ the old man/

   _____

7. his soldiers /he/ disposed/ for the battle

   _____

8. emperor/he/ proclaimed/ himself

   _____

국 문 초 록

본 연구는 수용적, 생산적 어휘 과업과 동일, 다양한 문장 문맥이 한국 중학교 학생들의 영어 어휘 학습에 미치는 영향에 대해 다음 두 가지 관점에서 검증하고자 한다. 우선, 본 연구는 이 두 가지 변인이 전반적인 어휘 학습에 미치는 영향을 단어 인지 시험, 소극적/적극적 단어 학습 시험, 두 개의 생산적 어휘 사용 시험 (빈칸 넣기, 단어 재배열 시험)의 총점을 분석하여 조사하고자 한다. 둘째로, 이 각각의 시험에서 측정된 세부적인 어휘 지식 학습과 보유를 이 두 개의 변인에 근거하여 살펴보고자 한다.

단어의 수용적, 생산적인 면모는 의사소통 과정의 두 가지 근간을 이루기 때문에 두 면모를 다 살펴보는 것은 학문적인 의미가 있다. 비록 많은 연구가 단어 학습에 있어 수용적 과업에 비해 생산적 과업의 우월성에 동의하지만 그에 반하는 연구들 역시 존재해 연구 결과가 동일하지 않다는 점에서 이 두 과업을 비교하는 연구가 더 필요한 실정이다. 한국에서 단어 과업이 대개 수용적 학습에만 의존하는 경향이 있다는 점 역시 생산적 학습에 대한 필요성을 주창한다.

어휘 학습에 있어 다른 중요한 요소인 문장 문맥의 영향에 대한 결과 역시 논란의 여지가 있다. 어휘 학습에 있어 문맥의 필요성에 대해서는 많이 연구되었지만 어떤 식으로 문장 문맥을 제공하는 것이 더 효과적인지에 대한 연구는 극히 미미한 실정이기 때문이다.

즉, 이 두 가지 변인은 어휘 학습에 중요한 영향력을 행사하지만 이들의 상호작용에 대한 연구는 거의 없다. 그러므로 본 연구는 동일/다양한 문장 문맥이 특정 단어 과업 안에서 주어졌을 때 그들이 상호적으로 어휘 지식 발달

에 미치는 영향력에 대해 살펴보고자 한다. 또한 어휘 지식의 다양한 면모를 살펴보기 위해서 다섯 개의 다양한 시험이 행해질 것이다.

본 연구의 참여자인 117 명의 한국 중학교 3 학년 학습자는 수용적/생산적 과업과 동일/다양한 문장 문맥이 조합된 4가지 과업 (RS, RD, PS, PD) 중 하나를 수행하고 즉시 사후 평가와 1 주일 후 이루어진 지연 사후 평가에 응하였다. 두 사후 평가는 앞서 말한 다섯 개의 시험으로 구성되어 있다.

전반적인 어휘 학습을 살펴본 결과, 과업의 종류는 어휘 학습과 보유에 상당하게 영향을 미치는 반면 문장 문맥은 그렇지 못했다. 그러나 어휘 보유에서 두 변인의 상호작용이 드러났다. 같은 문장 문맥을 제공하는 것이 생산적 과업에서는 긍정적인 효과를 내었지만 수용적 과업에서는 그렇지 못했다. 어휘 수용에 관한 한 다섯 개의 시험의 각각의 결과 역시 두 개의 생산적 어휘 사용 시험을 제외하고는 과업의 종류의 영향은 상당하지만 문장 문맥의 영향은 미미하다는 전반적 어휘 학습의 분석과 비슷한 결과를 냈다.

무엇보다 생산적 과업은 어휘의 생산적 사용 지식을 제외한 어휘 지식의 여러 단계에서 통계적으로 상당한 보유력을 지니고 있었다. 또한 과업과 상호 작용 시 문장 문맥은 특히 소극적 어휘 학습 시험과 적극적 어휘 학습 시험에서 강력한 영향력을 행사하였다. 반면 어휘의 생산적 사용 시험에서 어휘 지식의 보유를 측정할 때는 과업의 종류보다는 문장 문맥의 다양성 여부가 상당한 영향력을 보였다. 결과에 근거하여 본 연구는 과업의 종류와 문맥의 다양성에 대한 연구의 제언을 결론부에 제시한다.