EFFECTS OF DIMENSIONS OF WORD KNOWLEDGE AND THEIR ACCESSIBILITY ON DIFFERENT LEVELS OF READING COMPREHENSION IN ADOLESCENT EFL LEARNERS

By

Yunjeong Choi

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Curriculum, Instruction, and Teacher Education—Doctor of Philosophy

2017

ABSTRACT

EFFECTS OF DIMENSIONS OF WORD KNOWLEDGE AND THEIR ACCESSIBILITY ON DIFFERENT LEVELS OF READING COMPREHENSION IN ADOLESCENT EFL LEARNERS

By

Yunjeong Choi

Researchers have investigated that different dimensions of word knowledge (breadth versus depth) may play distinct roles in second language (L2) reading comprehension. Yet, little research has addressed how learners' efficiency of accessing those dimensions of knowledge functions in their comprehension (i.e., the issue of knowledge availability versus accessibility in Cremer and Schoonen, 2013), not to mention comprehension at different levels. To fill gaps in previous research and enrich our knowledge about the lexical basis of L2 reading comprehension, the present study examined how different dimensions of L2 word knowledge and their respective accessibility function concurrently in explaining different levels of L2 reading comprehension in adolescent English as a Foreign Language (EFL) learners in Korea.

To achieve this research goal, the present study defined lexical competence by distinguishing between vocabulary breadth and depth knowledge as well as between knowledge availability and accessibility. A number of tasks were employed to measure participants' lexical competences: three paper-and-pencil tests for knowledge availability and three researcher-developed, computer-based tests for knowledge accessibility documenting response accuracy as well as latency. Reading comprehension was measured with the Gates-MacGinitie Reading Comprehension Test (MacGinitie & MacGinitie, 1992), which assessed both literal and inferential comprehension. In addition, the participants' working memory and grammatical knowledge were also measured.

The following key research findings were revealed. First, Confirmatory Factor Analysis analyses confirmed that knowledge of individual word meanings (breadth availability) and the ability to activate that knowledge efficiently (breadth accessibility), and knowledge of meaning relations among words (depth availability) and the ability to access that knowledge efficiently (depth accessibility) are distinctive constructs under the conceptualization of lexical competence. Second, availability and accessibility of depth knowledge were the only predictors that made significant contributions to reading comprehension when the impacts of those of breadth knowledge and working memory were controlled for. Within vocabulary depth, availability played a more important role than did accessibility. Availability and accessibility of breadth knowledge made no unique contribution to reading comprehension over and beyond those of depth knowledge. Third, the patterns of predictive roles of each lexical competence in reading comprehension did not seem substantially different between literal and inferential comprehension, confirming a stronger impact of depth knowledge, both availability and accessibility, than of breadth knowledge, and of availability of depth knowledge than of accessibility. However, the involvement of depth knowledge, particularly availability, seemed greater in inferential comprehension than in literal comprehension.

These findings enriched our understanding about the lexical basis of L2 reading comprehension, particularly, the importance of efficient access to semantic network knowledge for textual reading and potentially differential involvement of semantic network knowledge in different levels of comprehension. Pedagogically, they suggested that vocabulary instruction should provide a learning environment where new words are provided in a meaningful reading context so that learners can establish a semantic network in their lexicon and continue developing, expanding, and consolidating the lexicon as they learn new words.

ACKNOWLEDGEMENTS

I have received huge support and encouragement from a great number of individuals in this journey of pursuing a doctorate. I would like to put into words the gratitude and appreciation I have for the people who heartened me with love and support as I completed my study.

I would like to start by thanking my advisor, Dr. Dongbo Zhang, for his excellent guidance and persistent support. His great insights always gave me numerous chances of thinking and rethinking about the key essence of the study, and helped me think, read, and write critically and logically. He also continually and convincingly conveyed a spirit of excitements as well as challenges and adventures in research and scholarship, which helped me understand the field of my research and academia. I would also like to acknowledge with appreciation the great academic support I have received from my dissertation committee, Dr. Charlene Polio, Dr. Paula Winke, and Dr. Doug Hartman. Their insightful feedback inspired me to think deeper about the significance and implication of my work in the field of second language reading both theoretically and pedagogically.

I am also grateful to my best MSU colleagues, Yu-Han Hung, Abraham Ceballos, Iwan Syahril, Gerardo Aponte-Martinez, Hyojeong Kim, and Elise Dixon. Not only did they give me tons of scholarly support when they read my work and discuss with me, but they were also always there for me as I went through the ups and downs of my doctoral life. They have exemplified for me what it means to be real friends. Their attitudes to face life’s challenges with positive energy and firm faith always challenged me to follow the same in my life. It was truly my privilege to meet them and work with them together in such a critical period of my life.

I must also thank my dear friends, Dongsook Han, Jin Hee Lee, and Unhee Joo, whose presence at MSU has always been an amazing shelter for me in the struggle of doctoral life.

Their friendship, caring, encouragement, and support helped me stay on track and go forward when I felt I was losing my way. I am so fortunate to have them in my MSU life and call them friends forever.

Completing this degree would not have been possible without the incredible support from my family. I would like to express my sincere gratitude to all of them. My parents instilled in me values of goal-determination, hard work, and dedication to whatever I committed to, which I believe was a strong guidance to keep me going in my doctoral study as well. I really appreciate their constant support, trust, and caring from Korea. I am also truly grateful to my sisters and brothers who always encouraged me to pursue my goal. Their words of support and love meant a lot to me and made me feel blessed. My parents-in-law were very supportive of me as well and gave me strength to pursue a degree. I would like to thank their ongoing support and prayers for me. I also have to give special thanks to my soul-sisters, HyoungJee and Seyeon, who shared my doctoral journey in spirit. They always held me up and never failed to let me know I am not alone and they are always there for me. I cannot thank them enough for their unshakable friendship and support over the years of getting a degree.

My wholehearted gratitude goes in a special way to my husband, Young June Sah, and my son, Joonu Sah. Young June has been a great husband, great friend, great colleague, and great academic mentor to me. His words and thoughts always inspired me to seek true meaning in my work and life. I am deeply thankful for his love and care, and that I could walk with him in this journey from the beginning. Joonu is also a true blessing to me. He has made me happy and blessed since his first day coming to my family, and this gave me positive vibe and energy I could depend on to keep working with passion and vision. This dissertation would not have been possible without the two most precious people in my life. Lastly, I want to thank God for his true love that kept me in faith and gave me strength, courage, and wisdom to complete this journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

The purpose of reading is comprehension. One reads a text to construct meanings from the text being read, and how much and how well one comprehends the text depends on a complex interplay of many factors. In school settings, students' comprehension skills profoundly affect their school achievements (RAND Reading Study Group, 2002; Snow, Burns, & Griffin, 1998). Their academic progress heavily relies on how they understand, analyze, and interpret information from texts they encounter across content areas (i.e., reading to learn).

Of many factors influencing reading comprehension, word knowledge plays a predominant role. The instrumentalist hypothesis proposed by Anderson and Freebody (1981) postulated that word knowledge is a major causal factor directly impacting one's text comprehension. It suggests a commonsense explanation of the relationship between vocabulary knowledge and reading comprehension: the more words one knows in a text, the more comprehensible the text is to one. According to Perfetti and Hart's (2001, 2002) Lexical Quality Hypothesis, skilled reading comes from high-quality lexical knowledge, indicated by well-specified orthographic-, phonological-, and semantic representations of words. In other words, individual differences in the quality of word knowledge would also lead to variance in the degree of text comprehension.

Considerable empirical evidence supports the close relationships of word knowledge with reading comprehension in English-speaking students. For example, word knowledge has been found to be a strong predictor of reading comprehension (Cain & Oakhill, 2014; Ouellette, 2006; Tannenbaum, Torgesen, & Wagner, 2006), thereby profoundly impacting students' academic achievement in school (Carlo et al., 2004; Snow, 2010; Snow et al., 1998). In addition, positive effects of instruction on words have been found on students' reading comprehension development, beyond the development of word knowledge itself (e.g., Beck, Perfetti, & McKeown, 1982; see Wright & Cervetti, 2016 for a recent review). Similar findings also have been reported for reading in a second language (Henriksen, Albrechtsen, & Haastrup, 2004; Hirsh, Nation, & others, 1992; M. H.-C. Hu & Nation, 2000; Laufer, 1992b; Laufer & Ravenhorst-Kalovski, 2010; Schmitt, Jiang, & Grabe, 2011; Verhoeven, 2000).

However, a notable limitation of those studies is a lack of attention to multidimensionality of word knowledge. For example, with reference to the perhaps best-known distinction between the size or breadth (i.e., how *many* words one knows focusing on form-meaning correspondence) and the depth (i.e., how *well* one knows those words) dimensions of word knowledge (Nagy & Anderson, 1984; Schmitt, 2014), previous research on the relationship of word knowledge with reading comprehension has focused primarily on the breadth dimension (Grabe, 2009; Schmitt, 2010b, 2014), such as the level of lexical coverage of a text for adequate unassisted comprehension and the correlational relationships between measured vocabulary knowledge (M. H.-C. Hu & Nation, 2000; Na & Nation, 1985; I. S. P. Nation, 2006) and reading comprehension (Mancilla-Martinez & Lesaux, 2010; Price, Meisinger, Louwerse, & D'Mello, 2016; Proctor, Carlo, August, & Snow, 2005; Shiotsu & Weir, 2007; Zhang, 2012). Little research attention, though, has been given to the depth dimension of word knowledge. This hinders us from comprehensive understanding of the function of word knowledge in reading comprehension.

Perhaps part of the challenge in addressing the role of vocabulary depth in reading comprehension is the lack of a clear conceptualization or definition of the construct of vocabulary depth. Usually, vocabulary breadth is definable relatively clearly because, by definition, it is presented with a quantifiable index, but this is not so for vocabulary depth (Chapelle, 1994; Henriksen, 1999; Meara, 1996; Read, 2000, 2004; Wesche & Paribakht, 1996). For example, if vocabulary depth refers to the quality of word knowledge or how well one knows the words one knows (i.e., form-meaning connections), how should ‘quality’ then be defined? What exactly does it mean to say one knows a word ‘well’? Read (2004), for example, distinguished three different lines of depth: precision of meaning, comprehensive word knowledge, and network knowledge. Haastrup and Henriksen (2000) defined depth of knowledge as knowing a word’s paradigmatic (e.g., antonymy, synonymy, hyponymy, gradation) and syntagmatic (e.g., collocations) relations to other words in the lexicon. And Qian (1999) and Li and Kirby (2015) in their studies on English L2 learners’ reading comprehension, also included morphological knowledge as one of their depth of vocabulary knowledge measures, in addition to learners’ knowledge of word meaning relations.

Despite the varied conceptualizations of vocabulary depth, the present study follows Read’s (2004) and focuses particularly on the network knowledge aspect of vocabulary depth. This aspect of vocabulary depth is particularly interesting because it focuses not just on knowledge of individual words as isolated entries in the mental lexicon, but also on semantic links or relations among those words. Developmentally, as one learns more words, those words will be accommodated in the existing network of already acquired words. As Read stated, “depth can be understood in terms of learners’ developing ability to distinguish semantically related words and, more generally, their knowledge of the various ways in which individual words are linked to each other” (p. 219). Based on such a conceptualization, psychometrically, vocabulary depth can be assessed with a focus on learners’ word association ability, with tasks such as the Word Associates Test (WAT) (Read, 1993, 1998) that often address three types of associations between a word and its associates: paradigmatic (superordinates, synonyms), syntagmatic (collocates), and analytic (words representing a key element of the meaning of the target word). For example, in a study comparing relative contributions of vocabulary size and depth to the reading comprehension of adult learners of English as a Second language (ESL) in Canada, Qian (1999) adopted Read's (1993) WAT to measure learners' paradigmatic and syntagmatic association ability to index their depth of word knowledge. Adopting the word association format, Horiba (2012) developed a vocabulary depth measure that covered paradigmatic, syntagmatic, and analytic associations for her Chinese- and Korean-speaking learners of Japanese.

In addition to the breadth *versus* depth distinction in the study of word knowledge and reading comprehension, another critical distinction that needs to be considered is of availability *versus* accessibility of different dimensions of knowledge one possesses (Cremer & Schoonen, 2013). Availability refers to knowledge one has in the mental lexicon, typically measured with offline vocabulary tests, such as the Vocabulary Levels Test (I. S. P. Nation, 1990; Schmitt, Schmitt, & Clapham, 2001) and WAT (Read, 1993, 1998). Critical to note is knowledge itself or availability of the knowledge, does not seem to be a sufficient condition for efficient textual comprehension. In other words, learners need not only various dimensions of word knowledge available for constructing propositional meanings and establishing a situation model (Kintsch, 1998), but also to access the knowledge with speed. Accessibility refers to such efficiency to activate the knowledge available. In the process of comprehending a text, the efficiency of recognizing words in the text themselves and accessing their meanings (i.e., efficient word recognition) as well as their relationships (e.g., collocation) is crucial, because effortless access to word meanings and relationships free up one's cognitive resources for participating in higher-level comprehension processes (Beck & McKeown, 2007; Perfetti & Hart, 2002). Schmitt (2010b) also pointed out that if one does not achieve appropriate lexical-recognition speed, one has to spend too much capacity in word-by-word decoding, thus “meaning construction is impaired and the overall flow of the text cannot be understood” (p. 106). Thus, accessibility of word knowledge is a key impact in reading comprehension and should be considered in relevant research discussions.

Regarding the place of accessibility in the conceptualization of word knowledge, particularly in the context of studying the relationship between word knowledge and reading comprehension, some researchers have argued that it is considered as a separate dimension of word knowledge parallel to other dimensions, such as vocabulary size and depth discussed earlier (Chapelle, 1994; Tannenbaum et al., 2006). Others have regarded it as a sub-component of depth of word knowledge (Schmitt, 2014). One problem with these conceptualizations is that they seem to regard accessibility as a type of word knowledge that one acquires such as knowledge about word meanings or meaning relations. However, accessibility is not so much knowledge as a skill that functions when one utilizes the knowledge. In other words, accessibility stands as another layer of language skill parallel to availability in conceptualizing lexical competence. It involves *both* breadth and depth of vocabulary knowledge in that it considers not only how efficiently one recognizes individual words one knows during the process of comprehending a printed text, but also how rapidly one accesses meaning connections among words stored in one’s mental lexicon to construct textual meanings for efficient comprehension. Drawing on this conceptualization, in the present study, accessibility of word knowledge, together with availability of word knowledge, works as a dimension of lexical competence.

Until now, reading comprehension literature, including that on L2 reading comprehension, seems to have focused primarily on vocabulary breadth (i.e., [availability of] knowledge of word form-meaning connections) (Cain, Oakhill, & Lemmon, 2004; Henriksen et al., 2004; Laufer, 1992b; Muter, Hulme, Snowling, & Stevenson, 2004) and word recognition efficiency (e.g., Kim, Wagner, & Foster, 2011; Kim, Wagner, & Lopez, 2012; Koda, 1996; Silverman, Speece, Harring, & Ritchey, 2013). Despite a small number of studies addressing the role of vocabulary depth (e.g., [availability of] knowledge of word relationships, such as synonymy and collocation), typically in comparison with vocabulary size, in reading comprehension (e.g., Qian, 1999, 2002; Li & Kirby, 2015; Zhang & Yang, 2016), little research specifically has probed how efficiency in accessing word relationships or network knowledge (i.e., accessibility of vocabulary depth knowledge) would contribute to reading comprehension, together with other aspects of word knowledge. In a notable study on Dutch-speaking monolingual and bilingual children, Cremer and Schoonen (2013) investigated to what extent semantic word knowledge supported reading comprehension. They not only measured availability of semantic word knowledge, but also paid attention to accessibility aspect, the speed with which that knowledge was activated to better explain individual differences in comprehension skills among monolingual and bilingual children. Speed of access to semantic knowledge was found to account for additional variance in reading comprehension beyond the impact of decoding and availability of semantic knowledge. This highlights the importance of considering accessibility in explaining the role of word knowledge in reading comprehension.

The above review of the literature suggests that little is known about how various dimensions of word knowledge, created by the distinctions of breadth *versu*s depth and of availability *versus* accessibility, specifically function in reading comprehension. Accordingly, an interest of the present study is to examine the relationships between reading comprehension and dimensions of word knowledge and their accessibility, with a focus on adolescent Korean-speaking learners of English as a Foreign Language (EFL) in South Korea. It also is an interest of this study to examine how the (relative) contributions of those dimensions of word knowledge to reading comprehension may be sensitive to the types of comprehension assessed in the learners.

Researchers have documented several levels of comprehension (Kintsch, 1988, 1998), and those different levels may involve as resources different types of knowledge or cognitive skills learners have when reading. For example, lower-level comprehension or literal comprehension is characterized by one being able to find information explicitly stated in a text. To achieve this level, learners "easily" could identify individual words in a sentence and establish the sentence's propositional meaning by applying knowledge of the grammatical structure(s). In contrast, higher-level comprehension such as inferential comprehension needs higher-order cognitive skills in addition to basic language skills, such as inferencing skills and background knowledge, so one can go beyond the explicit meaning of the text and infer or interpret what the text is about.

The distinction between higher- and lower- level comprehensions adds another layer of complexity in addressing different dimensions of word knowledge and their accessibility in reading comprehension. In other words, how different dimensions of word knowledge and their accessibility influence reading comprehension might vary as a function of the level of comprehension. For example, learners' word knowledge of different dimensions may be engaged in different ways to meet the different demands of comprehension processes. Likewise, learners' accessibility of word knowledge may be involved differentially in different types of comprehension.

Until now, little research has addressed how different dimensions of word knowledge possibly would influence different levels of comprehension in different ways, and how accessibility functions in those relationships. In a recent study on Chinese-speaking EFL learners, Li and Kirby (2015) examined the relationship between breadth and depth of vocabulary, and their relative contribution to different aspects of L2 (English) reading comprehension measured with multiple choice questions and summary writing. The study showed that vocabulary breadth significantly predicted reading comprehension measured with multiple-choice questions whereas vocabulary depth was the stronger predictor of summary writing, which demanded higher levels of processing and memory skills. A similar finding was reported in a study by Zhang and Yang (2016) which examined the relative contributions of vocabulary breadth and depth to reading comprehension that touched on different levels of comprehension among adult learners of Chinese as a Second Language. However, neither study considered the role of accessibility of word knowledge in reading comprehension, such as in Cremer and Schoonen (2013) reviewed earlier.

To address these research gaps, the present study aims to examine how different dimensions of L2 word knowledge (i.e., availability of vocabulary breadth and depth knowledge) and their accessibility function concurrently in explaining different levels of the L2 reading comprehension of adolescent EFL learners in Korea. Theoretically, this study is expected to provide a clearer picture of what constitutes L2 word knowledge, and of how different dimensions of word knowledge relate to each other. More importantly, it is expected to shed light on the relationships of different dimensions of word knowledge and their accessibility with reading comprehension, and generate insights into a more comprehensive model for the lexical basis of L2 reading comprehension. Pedagogically, this study will suggest curricular and instructional implications of what teachers and other practitioners could pay attention to when teaching vocabulary, and how their instructions could be delivered in ways conducive to promoting L2 learners' engagement with words, and to developing deep knowledge of vocabulary and rapid knowledge access for efficient textual reading (and other literacy purposes).

CHAPTER TWO

LITERATURE REVIEW

This chapter reviews existing literature on conceptualizations of word knowledge and its multidimensionality, and the relationships between different dimensions of word knowledge and reading comprehension. In addition to reviewing the distinction between vocabulary breadth and depth (i.e., availability), accessibility of word knowledge is also discussed as an additional layer of lexical skill that influences reading comprehension. This review also addresses how different dimensions of word knowledge and their accessibility work together to contribute to comprehension of different levels, that is, how the (relative) contributions of different dimensions of word knowledge and their accessibility to reading comprehension may be a function of the level of comprehension in question.

2.1. Lexical basis of reading comprehension

Reading comprehension is a process of extracting and constructing meaning from a written text (RAND Reading Study Group, 2002). While reading, the reader recognizes individual words in the text and integrates both the meanings of those words and his/her personal knowledge, actively to construct the meaning of what is being read. All these processes are simultaneous and often automatic. Hence, for successful comprehension, one needs different types of knowledge, skills, and strategies, such as metalinguistic awareness, word recognition skills, linguistic knowledge (e.g., vocabulary and grammatical knowledge), inferencing skills, comprehension strategies, and world knowledge (Cain, 2005; Duke & Carlisle, 2011; Grabe, 2009; K. Nation & Norbury, 2005; Perfetti, 1999).

Among various factors impacting on reading comprehension, word knowledge is fundamental in distinguishing skilled readers from less-skilled readers. According to Perfetti's

(1985) Verbal Efficiency theory, which contends that skilled reading depends on efficiency of word-level processing, an automatic and effortless word identification - efficient process - can preserve more processing resources for higher-level comprehension. Efficiency here is not simply speed alone, but rather about how efficiently a reader identifies form and meaning components of words in a sentence, which are the basic components of comprehension. Readers who can retrieve meanings they need from each word in a given context are more skillful in reading than those who cannot.

More recently, in his Lexical Quality Hypothesis (LQH), Perfetti (Perfetti, 2007; Perfetti & Hart, 2002) regarded differences in reading skills as essentially the differences in readers' lexical quality, which refers to "the extent to which a mental representation of a word specifies its form and meaning components in a way that is both precise and flexible" (Perfetti, 2007, p. 359). In other words, a high-quality representation of a word means strong orthographic (i.e., spelling), phonological (i.e., sound), and semantic (i.e., meaning) representations (Perfetti & Hart, 2002). Precision and flexibility in form-meaning correspondence both matter in lexical quality, because one needs to know that "*knight and night*" (Perfetti, 2007, p. 359) are not the same (i.e., precision), and to understand that "*roaming charge*" (Perfetti, 2007, p. 359) is a type of fee charged by a mobile company, not a battle-maneuver (i.e., flexibility). Lexical representation is in a continuum from no knowledge to the full, coherent representation of a word. Each person's lexical representation varies in terms of what and how much one knows about a word, which consequently leads to individual differences in comprehension.

The Lexical Quality Hypothesis provides a basis for understanding the critical import of word knowledge and individual differences in reading comprehension. Skilled comprehenders tend to have stronger word representations and more detailed or nuanced knowledge about word forms and meanings. In contrast, poor comprehenders' poor lexical quality suggests that their comprehension processes for constructing a situation model of a given text tend to be interfered with by their semantically-incomplete lexical representations (Hamilton, Freed, & Long, 2013). Moreover, the effortful, inefficient lexical processing coming from their poor quality of lexical knowledge easily consumes available cognitive resources, which otherwise could be invested in higher-level comprehension. Central to textual comprehension is the word knowledge a reader brings to reading.

The following sections further discuss how word knowledge is represented, what its roles are in reading comprehension, what other factors need consideration, and how relationships among key factors have been evidenced in previous literature.

2.2. What does it mean to know a word?

Knowing a word does not mean merely that one can provide a definitional meaning of it, which is perhaps the most expected response when one is asked to justify that one knows the word. Beyond the knowledge of a simple form-meaning relationship, however, many aspects to know about a particular word are learned in an incremental process throughout one's life (Read, 2000). Researchers have attempted to specify those aspects, in diverse ways of conceptualization, better to explain what it means to know a word or what word knowledge is (Chapelle, 1994; Daller et al., 2007; Henriksen, 1999; Kieffer & Lesaux, 2012; Meara, 1996; Nagy & Anderson, 1984; I. S. P. Nation, 2001; Richards, 1976; Schmitt, 2010a)

For example, Richards (1976) provided the following list of 8 assumptions concerning different aspects of what knowing a word means from the perspective of second language program design and vocabulary instruction.

1. The native speaker of a language continues to expand his vocabulary in adulthood, whereas there is comparatively little development of syntax in adult life.
2. Knowing a word means knowing the degree of probability of encountering that word in speech or print. For many words we also know the sort of words most likely to be found associated with the word.
3. Knowing a word implies knowing the limitations imposed on the use of the word according to variations of function and situation.
4. Knowing a word means knowing the syntactic behavior associated with the word.
5. Knowing a word entails knowledge of the underlying form of a word and the derivations that can be made from it.
6. Knowing a word entails knowledge of the network of associations between that word and other words in the language.
7. Knowing a word means knowing the semantic value of a word.
8. Knowing a word means knowing many of the different meanings associated with a word. (p. 83)

Chapelle (1994) defined vocabulary ability as involving three components: context of language use, vocabulary knowledge and processes, and metacognitive strategies. Under the component of vocabulary knowledge and processes, she outlined four different dimensions, including a) vocabulary size, b) knowledge of word characteristics (e.g., phonemic, morphemic, and collocational features), c) lexicon organization, and d) fundamental vocabulary processes. Henriksen (1999) defined lexical competence by proposing three dimensions: a) partial to precise knowledge, b) depth of knowledge, and c) receptive to productive use ability.

Among many conceptualizations of word knowledge, Nation (2001) seems to provide the most comprehensive coverage of word knowledge aspects. According to Nation (2001), knowing a word includes at least nine different aspects, under each of which a further distinction is made between receptive and productive knowledge of the word. More specifically, Nation (2001) describes knowing a word as engaging form (i.e., spoken, written, and word parts), meaning (i.e., form-meaning, concept and referents, and associations), and use (i.e., grammatical functions, collocations, and constraints on use). To elaborate, here is an example given by Nation (2001):

Receptive knowledge of *underdeveloped* involves:

- Being able to recognise the word when it is heard [spoken]
- Being familiar with its written form so that it is recognized when it is met in reading [written]
- Recognizing that it is made up of the parts *under*-, -*develop*- and -*ed* and being able to relate these parts to its meaning [word parts]
- Knowing that *underdeveloped* signals a particular meaning [form and meaning]
- Knowing what the word means in the particular context in which it has just occurred [form and meaning]
- Knowing the concept behind the word which will allow understanding in a variety of contexts [concept and referents]
- Knowing that there are related words like *overdeveloped*, *backward*, and *challenged* [associations]
- Being able to recognize that *underdeveloped* has been used correctly in the sentence in which it occurs [grammatical functions]

- Being able to recognize that words such as territories and areas are typical collocations [collations]
- Knowing that underdeveloped is not an uncommon word and is not a pejorative word [constraints on use]

(from I. S. P. Nation, 2001, p. 26, with square brackets added)

2.3. Vocabulary breadth *versus* depth

Perhaps the best-known conceptualization of vocabulary knowledge is the distinction between size (or breadth) and depth (Anderson & Freebody, 1981; Read, 2004; Schmitt, 2014). While vocabulary size, also known as vocabulary breadth, is straightforward to define as how many words one knows (i.e., the *quantity* of word knowledge), vocabulary depth (i.e., how well one knows about the words) is not a simple construct easily defined. Much of what depth is, or what “how well” means precisely, or what vocabulary depth specifically entails still remain unsettled.

2.3.1. Vocabulary breadth

To discuss the number of words one knows presumes to define what a ‘word’ is, but that is not an easy concept. Some basic concepts offer understanding of what a word is. Type/token distinction is one of the useful terms when counting words. *Tokens* are the number of running words in a text regardless of repeated occurrences, while *types* are the number of the different words. A *lemma* is a collective term, which indicates the base and inflected forms of a word (e.g., communicate, communicates, communicated, communicating). While a lemma is a group of related word forms within a word class, *word family* encompasses all of the word forms which are related semantically (e.g., communication, communicative, communicator).

To inform language curriculum and teaching (e.g., how many words L2 learners need to know at which stage of language learning), there have been different estimates of the number of words in English. Goulden, Nation, and Read (1990) estimated about 114,000 base words (non-derived entries) in English based on *Webster's Third New International* Dictionary. They also found that university students of native speakers in New Zealand had a vocabulary size of about 17,000 word families.

Nagy and Anderson (1984) identified 88,500 distinct word families, with a word family defined as a head word excluding compounds, abbreviations, non-main entries in dictionaries, foreign words, and derived or variant words. In general, it is estimated that a well-educated L1 or native English speaker knows around 20,000 word-families (excluding proper names and transparently derived forms) (I. S. P. Nation, 2006). These are of course not the number of words L2 learners need to know to be able to read L2 texts. It is not possible even for native speakers to know all these words, nor do they need to. However, the figures highlight the importance of size of vocabulary by giving a sense that L2 learners need to have considerable size of vocabulary or at least a part of their vocabulary learning goal needs to increase size of vocabulary.

2.3.2. Vocabulary depth

While the conceptualization of vocabulary size seems straightforward in a quantitative sense with respect to form-meaning relationships, vocabulary depth or quality is a rather complex construct to be clearly laid out. Schmitt (2010) suggested that there had been two ways to conceptualize the depth or quality of knowledge of individual words. One is the ‘developmental’ approach (Read, 2000), which considers vocabulary acquisition as a continuum of mastery ranging from ‘do not know at all’ at the beginning, all the way to ‘full mastery of a lexical item in all contexts of use’ at the advanced end. A typical way of operationalizing vocabulary depth in this approach is to set up a set of scales describing different stages of vocabulary mastery. For example, the *Vocabulary Knowledge Scale* developed by Paribakht and Wesche (1997) specifies five stages of knowledge development from no knowledge (i.e., *I don't remember having seen this word before*) to full knowledge with production (i.e., *I can use this word in a sentence*) with the last three requiring demonstration of knowledge by producing synonyms/translations and a sentence. It is a self-report scoring system with possible scores 1-5 based on familiarity of the word and ability to use it in a semantically and grammatically appropriate way. Specifically, one's knowledge of a word can be described as 'The word is not familiar at all' (Score 1), 'The word is familiar but its meaning is not known' (Score 2), 'A correct synonym or translation is given.' (Score 3), 'The word is used with semantic appropriateness in a sentence.' (Score 4), and 'The word is used with semantic appropriateness and grammatical accuracy in a sentence' (Score 5). Although the developmental approach captures the incrementality of vocabulary acquisition (Nagy & Scott, 2000), it has been criticized for the fact that it is not appropriate to represent (the diverse aspects of) lexical knowledge in general (Read, 2000), nor does it provide clear judgement of each category of degree of vocabulary mastery (Schmitt, 2010b).

The other way to conceptualize vocabulary depth is a dimensional or components approach, which posits that word knowledge can be divided into multiple dimensions or components (Read, 2000). For example, Read (2004) outlined three distinct (but overlapped) aspects of word knowledge in L2 vocabulary acquisition:

1) The difference between having a limited, vague idea of what a word means, and having much more elaborated and specific knowledge of its meaning, which I will refer to as *precision of meaning*.

2) Knowledge of a word which includes not only its semantic features but also its orthographic, phonological, morphological, syntactic, collocational and pragmatic characteristics: *comprehensive word knowledge.*
3) The incorporation of the word into a lexical network in the mental lexicon, together with the ability to link it to - and distinguish it from - related words, which we can call *network knowledge*. (Read, 2004, p. 211)

What draws attention in relation to the present study is network knowledge, similar to what Schmitt (2014) called *lexical organization* and highlighted as the "most promising approach" (p. 31) to capture the characteristics of vocabulary depth. This approach considers vocabulary depth as learners' ability to link a newly-acquired word to already-known words stored in the existing mental lexicon, resulting in an expanded or denser network.

Lexical network is based on semantic relations among words, as mapped out through word associations. Typically, word associations are classifiable via paradigmatic-syntagmatic distinction. Words are paradigmatically associated when they form semantic relationships of coordinates, superordinates, subordinates, synonyms, or antonyms; and belong to the same word class. Syntagmatic association, on the other hand, refers to the collocational relationship between two words. Drawing on this paradigmatic-syntagmatic categorization, Read (1993, 1998) developed the Word Associates Test (WAT) to measure English L2 learners' vocabulary depth. Many WAT variants also have been developed for diverse assessment and research purposes (e.g., Qian, 1999, 2002; Qian & Schedl, 2004; Schoonen & Verhallen, 2008).

WAT is essentially a controlled, recognition-based association test that measures a test taker's ability to identify paradigmatic (i.e., synonyms) and syntagmatic associates (i.e., collocates) from distractors for a target word. In the revised version of Read's original WAT (Read, 1998), for example, a target adjective (e.g., *sudden*) and two groups of four words are given; the four adjectives in the left group include words in paradigmatic relationships with the target word together with adjectival distractors (e.g., *beautiful, quick, surprising,* and *thirsty*) and the other four nouns in the right group include words that have syntagmatic relationships with the target (e.g., *change, doctor, noise,* and *school*). For the present example, a test taker is supposed to select *quick* and *surprising* as the paradigmatic associates and *change* and *noise* as the syntagmatic associates of *sudden*. While all target words in the revised WAT have four associates, their distribution varies. In addition to a 2-2 distribution in the foregoing example, that is, two paradigmatic and two syntagmatic associates, there are also 3-1 and 1-3 distributions.

One of the critical issues regarding WAT, despite its wide use in research and its known validity and reliability in some validation studies (Read, 1993, 1998; Schmitt et al., 2011), is whether paradigmatic and syntagmatic associations are unidimensional so that a composite score of the two can be used to indicate one's network knowledge, which has been a common practice in the literature (e.g., Read, 1993, 1998; Qian, 1999; Qian, 2002; Zhang & Yang, 2016). Although significant correlations between the scores for the two types of association have been reported in some studies (Horiba, 2012; Qian & Schedl, 2004), others (e.g., Batty, 2012; Shin, 2015) who evaluated the factor structure of network knowledge beyond simple correlations suggested that paradigmatic and syntagmatic associations might represent two different dimensions of vocabulary depth. In Shin's (2015) study with Korean elementary school EFL learners, two different sets of items were designed to measure paradigmatic and syntagmatic association separately, assuming they were two different types of knowledge contributing to two underlying factors. Confirmatory factor analysis (CFA) indicated that paradigmatic and syntagmatic associations, despite their significant correlation (*r*=.83), seemed to tap rather different dimensions of vocabulary depth. Batty (2012) tested three CFA models with Japanese university students. Among the three models tested, the bifactor model demonstrated the best fit, where the primary explanatory factor was a single vocabulary g(eneral)-factor, with additional, uncorrelated second-order subskill factors for paradigmatic and syntagmatic items. That is, paradigmatic-syntagmatic associations may be distinct subskills of vocabulary depth, suggesting that it may be meaningful to measure the two dimensions separately or provide separate subscores for each if tested together as in WAT.

Another important, if not more important, but nonetheless unresolved issue is whether vocabulary depth so conceptualized (i.e., network knowledge) and measured (i.e., through WAT) is indeed distinct from vocabulary size. As one's vocabulary size increases, one's lexicon presumably would have more words linked to each other and the vocabulary network would become denser and more organized. This relationship is well captured in Meara's (2009) conceptualization of vocabulary size and organization using a diagram with nodes and links among them (Meara, 2009). In this model, vocabulary size corresponds to the number of nodes (little squares) in one's lexical network. Between the nodes are numerous connections, which represent how words in the lexicon connect with each other to form a network. Some words are connected to merely one or two other words, while some are associated with multiple words simultaneously, making the network dense and complex. To put it differently, as one acquires a new word, that is, adds a new node in the network, it subsequently adds one or more links to existing` words in the network. Consequently, it logically is expected that one's vocabulary size is closely tied to one's depth of vocabulary or network knowledge, or in Meara's term, "size and organization interact." (Meara, 2009, p.77)

The close relationship between size and depth of vocabulary has actually been evidenced by a strong correlation between them in some studies (Akbarian, 2010; Greidanus, Bogaards, van der Linden, Nienhuis, & de Wolf, 2004; Henriksen, 2008; Horiba, 2012; Mehrpour, Razmjoo, & Kian, 2011; Qian, 1999, 2002; Vermeer, 2001; Zhang, 2012). The question is whether the strong correlation suggests that vocabulary size and depth essentially belong to the same construct of vocabulary knowledge, or are distinct aspects despite their close relationship. Vermeer (2001), in two studies examining in L1 and L2 Dutch kindergartners the relationship between vocabulary size and depth, argued there essentially is no difference between size and depth based on the high correlations between a depth measure and two size measures. In Dutch primary classrooms she also found strong correlation between vocabulary size and the frequency with which the words occur in the oral and written language input. She concluded that a size test with a good sample of words could measure children's vocabulary knowledge as well as would a depth test.

In a study on the relationship between word knowledge (size and depth) and reading comprehension in L2 learners of Japanese, Horiba (2012) found that the correlation between the size and depth measures was .51 for Chinese-speaking learners, but for Korean-speaking learners was notably higher ($r = .90$). The extremely high correlation for Korean speakers suggests essentially no distinction between size and depth, which seems to be corroborated in the follow-up regression analysis: for both groups, there was no unique and significant contribution of vocabulary depth on reading comprehension when the effect of vocabulary size was considered.

On the other hand, many other studies showed that vocabulary size and depth tap distinct aspects of word knowledge. For example, in a study on Korean and Chinese ESL learners in Canada, Qian (1999) found that vocabulary depth (measured with Read's (1993)) significantly explained additional variance in the learners' reading comprehension, after controlling for the effect of vocabulary size (measured with the VLT (Nation, 1990)). In addition, over and above the contribution of depth, size also uniquely contributed to reading comprehension. Another study of Dutch primary-school children by Schoonen & Verhallen (1998) also compared the (relative) contributions of vocabulary size and depth to reading comprehension measured by two cloze passages. Similar to Qian's (1999) findings, although a strong correlation between size and depth was found, further regression analysis revealed that each vocabulary test uniquely contributed about 5–10% to the reading comprehension, over and above the impact of the other test. Taken together, these results suggested that vocabulary size and depth might be distinct aspects of vocabulary knowledge.

2.4. Vocabulary size, depth, and reading comprehension

2.4.1. Vocabulary size and reading comprehension

Meara (1996) highlighted the critical role of vocabulary size and stated, "all other things being equal, learners with big vocabularies are more proficient in a wide range of language skills than learners with smaller vocabularies" (p. 37). As indicated in Anderson and Freebody's (1981) instrumentalist hypothesis for a close relationship between vocabulary knowledge and reading comprehension, words are the building blocks of texts; without knowing the meanings of the words in a text, comprehension of it will not occur. Thus, a lens for understanding vocabulary-size importance for reading comprehension is via examination of the lexical coverage of a text (i.e., the ratio of known words in a given text) and readers' adequate, unassisted comprehension of the text (M. H.-C. Hu & Nation, 2000; Na & Nation, 1985). For example, in a study with adult English learners using a narrative text, Hu and Nation (2000) reported that around 98% of lexical coverage may be needed for L2 learners to gain "adequate comprehension" (p.419). They found that adequate comprehension also was possible with coverage of 95%, but only a small proportion of the participants achieved it. In a study with university students whose L1 was Hebrew or Arabic, Laufer (1992a) indicated that a 95% lexical threshold may be needed for adequate academic reading comprehension, which corresponds to knowing about 3000 word families or 5000 lexical items to be translated into Nation's (1993) criteria.

A more direct way of understanding how vocabulary size and reading comprehension are related is to measure them among learners and examine their correlational relationships. Henrikson (2004), for example, found in a cross-sectional study of Danish (L1) learners of English (L2) that the scores on English vocabulary size (measured by VLT) and the English reading test were highly correlated both for 10th grade (r=.85) and university (r=.79) student groups. Laufer (1992b) found that the higher L2 learners' vocabulary size, the higher they scored in a reading comprehension test; in addition, after controlling for their general academic ability, the participants' lexical level (measured by VLT) accounted for 26% of the unique variance in reading comprehension, indicating the critical role of vocabulary size in reading comprehension. Nassaji (2003) showed that receptive vocabulary knowledge was the strongest among the component skills associated with L2 reading for 60 university-level ESL learners. Similar close relationships between vocabulary size and reading comprehension have been reported also among young learners. Verhoeven's (2000) longitudinal study of native Dutch-speaking children (L1) and minority children (L2) in primary school, for example, reported a strong predictive effect of receptive vocabulary knowledge on reading comprehension for both groups in a series of Structural Equation Modeling (SEM) analyses. But the impact of vocabulary knowledge was much higher for the minority group than for the native-Dutch-speaking group. A similar finding was reported also in Droop and Verhoeven (2003) on third- and fourth-grade language-minority children in the Netherlands.

2.4.2. Vocabulary depth and reading comprehension

How vocabulary depth functions, or explains individual differences, in reading comprehension has been examined much less in the literature than research on vocabulary size, probably because of the inherent complexity in the concept of depth as reviewed earlier. In the literature, vocabulary depth has been operationalized and measured in diverse ways, and consequently, findings on the contribution of vocabulary depth to reading comprehension on one hand, and relative contributions of vocabulary size and depth on the other, were not always converged.

Ouellette (2006), for example, distinguished between vocabulary breadth and depth to examine the role of oral vocabulary in different reading skills, and conducted a study with 4th grade English-speaking students. Vocabulary breadth was measured with both receptive and expressive vocabulary tests, while vocabulary depth was assessed by word definition and synonym tests. Analysis showed that only vocabulary depth significantly and uniquely predicted reading comprehension. When controlling for depth, neither receptive nor expressive vocabulary breadth significantly accounted for any additional variance of reading comprehension. In their longitudinal study, Nation and Snowling (2004) explored the relationship between phonological, oral language (including expressive vocabulary and semantic skills), and reading skills, with English-speaking children from the age of 8.5 to 13 years. Although not specifically indicated in the study, the researchers' measures of expressive vocabulary knowledge and semantic skills (i.e., semantic fluency and synonym task) seemed to tap vocabulary size and depth, respectively. Part of the results highlighted the role of expressive vocabulary knowledge and semantic skills (measured by a semantic fluency and synonym judgement task) in explaining reading comprehension. That is, both expressive vocabulary and semantic skills uniquely predicted concurrent reading comprehension skills, and this relationship was found to be also the case 4.5 years later.

In the field of L2 reading comprehension, increasing attention also has been given to the role of vocabulary depth in reading comprehension and to whether vocabulary depth makes a unique contribution to reading comprehension independent of vocabulary size (e.g., Qian, 1999, 2002; Li & Kirby, 2005; Zhang & Yang, 2016). While findings of some studies supported a unique contribution of vocabulary depth to reading comprehension, and the contribution of vocabulary depth appeared stronger than that of vocabulary size (e.g., Qian, 1999, 2002), those of other studies painted a complex picture about the relationships.

Qian (1999), for example, explored the relationship between size and depth of vocabulary knowledge and reading comprehension with Korean and Chinese young adults who studied English as a second language. The participants' L2 English vocabulary size was measured with the VLT (Nation, 1990) and depth was measured with Read's (1993) WAT. Hierarchical regression analysis showed that both vocabulary size and depth significantly contributed to reading comprehension. Vocabulary size alone explained 60% of the variance in reading comprehension, while vocabulary depth alone also explained 68% of the variance in reading comprehension. This suggests that vocabulary size and depth were both significant contributors to reading comprehension. Further, vocabulary depth significantly added 11% of the unique portion of explained variance in reading comprehension after controlling for the effect of vocabulary size, whereas the unique variance explained by vocabulary size over and above the effect of depth was, albeit significant, only 3.7%. This finding highlighted the stronger predictive power of vocabulary depth than vocabulary size. Interestingly, when WAT, which measured vocabulary depth, was divided into its two sub-tests of paradigmatic (synonym) and syntagmatic (collocation), the two sub-tests yielded similar relationships with reading comprehension. Specifically, paradigmatic association produced a correlation of .79 with reading comprehension, and the correlation between syntagmatic association and reading comprehension was .78.

Qian (2002) conducted a similar study with students from mixed L1 backgrounds enrolled in an ESL program in a university in Canada. Results supported findings from the earlier study, confirming that vocabulary depth played a more important role in explaining reading comprehension (measured by an unpublished version of TOEFL reading comprehension test). Specifically, vocabulary depth alone explained 59% of the variance in reading comprehension, while vocabulary size alone did 54%. The unique contribution of vocabulary depth, over and above vocabulary size, to reading comprehension was 13%, which was higher than the 8% of unique variance in reading comprehension explained by vocabulary size. The predictive power of vocabulary depth was even higher than that of the scores on TOEFL vocabulary item measures, which alone accounted for 54% of the variance in reading comprehension. A stronger impact of vocabulary depth (WAT) than vocabulary size (VLT) on reading comprehension was found also in Mehrpour, Razmjoo, and Kian's (2010) study with Iranian EFL university students. Multiple regression analysis in this study revealed that the regression coefficient (Beta) for vocabulary depth (i.e., .46; $p<.01$) was larger than that of vocabulary size (i.e., .32, $p<.05$).

In contrast, in Huang's (2006) study with university Chinese ESL learners, regression analysis yielded that vocabulary size alone accounted for 50% of the variance of reading comprehension, which was bigger than the variance explained by vocabulary depth (44%). The unique variances explained by vocabulary size and depth after controlling for the effect of the other was 8.1% (F(l, 21) = 3.59, p= .072) and 2.4% (F(1, 21)=1.08, p= .311), respectively, although the contribution of each was not statistically significant. Similar results were revealed also in Farvardin and Koosha's (2011) study conducted with Iranian EFL college students. Their analysis revealed that vocabulary size alone accounted significantly for 61.5% of the variance in reading comprehension, while only 48.2% of the variance in reading comprehension was explained by vocabulary depth. The unique variances explained in reading comprehension by size and depth over and above the other was 17% and 3.7%, respectively, neither of which were statistically significant. In Horiba (2012)'s study with Chinese and Korean students learning Japanese (L2) (as well as native Japanese speakers), multiple and stepwise regression analysis revealed that for both groups of L2 learners, only vocabulary breadth was found to be a significant predictor of reading comprehension, while no significant role of vocabulary depth was identified.

2.5. Word knowledge, its accessibility, and reading comprehension

2.5.1. Accessibility of word knowledge

Until this point, what has been discussed regarding word knowledge is what word knowledge means and what it means that one has the knowledge. In other words, primarily it is concerned with the *availability* of word knowledge – the extent to which one has the knowledge of form-meaning connections (e.g., vocabulary size) and relationships between words in one's mental lexicon (e.g., vocabulary depth) (Cremer & Schoonen, 2013). In this sense, to measure one's vocabulary size and depth is to assess explicit, declarative knowledge, or knowledge available to the learner as a conscious representation; that is, one knows what it is one knows and what one knows can be accessed and reported typically in a paper-and-pencil vocabulary test

(Bialystok, 1981; Read, 2004). However, knowledge itself as demonstrated through an offline vocabulary test (e.g., VLT and WAT) does not necessarily suggest it readily can be accessed in real-time language use situations, including textual reading, without the capacity to activate it in a rapid, unconscious manner. Called by Cremer and Schoonen (2013) as the *accessibility* of word knowledge, it is an implicit, procedural knowledge that controls fluent or automatic processing in word recognition, language comprehension and production (Read, 2004). According to Nagy and Scott (2000), word knowledge is primarily procedural (i.e., a matter of knowing how) rather than declarative (i.e., knowing "that"; e.g., that Washington D.C. is the capital of the US). Knowing a word goes beyond having the knowledge (about word forms, meanings, and relationships) (i.e., availability); it is more about being able to do things with the knowledge that one has, including recognition of words in connected discourse, meaning access, and pronunciation of them, and more importantly, "do[ing] these things within a fraction of a second" (p. 273), which is an issue of how accessible the knowledge is for efficient processing of language.

To further unpack the complexity of the dimension of vocabulary knowledge, Daller et al. (2007) proposed the concept of *lexical space* where a learner's word knowledge is described with a metaphor of a three-dimensional space. The horizontal axis represents lexical breadth, defined as the number of words one knows, whereas the vertical axis indicates lexical depth, referring to how much one knows about the words one knows. The final axis, which exists in another dimension passing through the dimension of breadth and depth, is fluency, defined as "how readily and automatically a learner is able to use the words they know and the information they have on the use of these words" (p. 8).

Chapelle (1994) likewise also considered fluency a dimension of word knowledge parallel to other dimensions such as vocabulary size, word characteristics, and lexical organization. She called this concept *fundamental vocabulary processes* associated with lexical access. Fluency is critical in her definition of word knowledge, because it is the knowledge that encompasses a wide-range of essential word processing skills such as attending to relevant vocabulary features, encoding phonological and orthographic information, accessing semantic properties from the lexicon, integrating semantic representation of the text, parsing words into their morphological components; and composing words morphologically (p.166). Such a conceptualization is represented also in Qian's (2002) framework of vocabulary knowledge, which included automaticity of receptive-productive knowledge as the fourth dimension on top of vocabulary size, depth, and lexical organization.

Schmitt (2014), on the other hand, highlighted fluency as a sub-component of vocabulary depth. He argued that fluency is important "because it moves the conceptualization of lexical proficiency onward from simple knowledge to the ability to use that knowledge in both comprehension and production." (p. 8) In other words, one's knowledge of words does not necessarily lead to one's appropriate engagement of those words in real communication. Given that the ultimate goal of vocabulary learning is the (efficient) use of the knowledge for various language-use situations, fluency is arguably a pivotal capacity one needs to develop for mastery of word knowledge.

Despite diverse conceptualizations delineated above about accessibility of word knowledge, this study follows Cremer and Schoonen (2013) and argues that accessibility, which concerns automaticity of knowledge activation or access, is a separate layer of language ability closely tied to different aspects of word knowledge discussed earlier (i.e., availability), rather than a distinct aspect of vocabulary knowledge (Chapelle, 1994; Daller et al, 2007; Qian 2002) or a sub-component of vocabulary depth (Schmitt, 2014). As shown in Figure 1, accessibility pertains not only to the access of form-meaning connections, the concern of vocabulary breadth, but also to the access of word relationships or network knowledge or vocabulary depth.

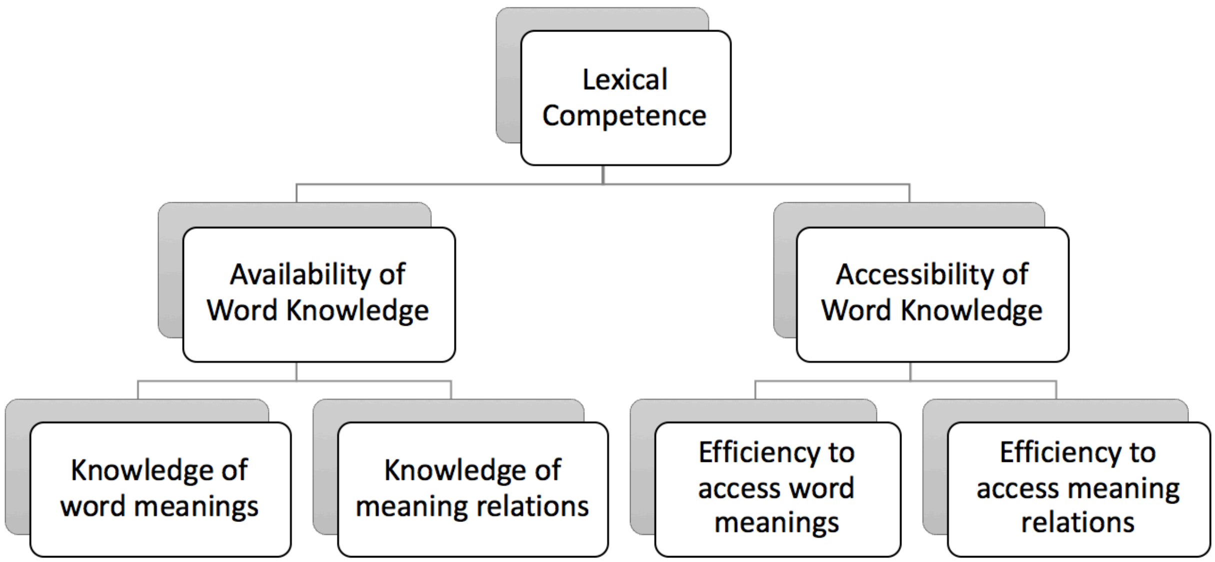

Figure 1: Hypothesized Conceptualization of Availability and Accessibility of Word Knowledge.

Given the importance of both availability and accessibility, it is reasonable to expect that both be considered in the vocabulary knowledge literature, including vocabulary assessment and, more importantly, relative to the focus of the present study, the examination of vocabulary knowledge in reading comprehension (see 2.5 Accessibility of word knowledge and reading comprehension). Unfortunately, most existing research has focused on the availability dimension with only limited attention to knowledge accessibility.

Accessibility has often been addressed by administering time-sensitive assessment tasks (Tannenbaum et al., 2006; van Gelderen et al., 2004). Typically, accessibility to form-meaning connections is measured by the speed or efficiency of word recognition. A lexical decision task is often used for this purpose, where participants are asked to indicate as fast as they can whether a given word is an existing word. Both response time and accuracy of the choices are collected.

For example, a study by van Gelderen et al. (2004) of Dutch EFL readers adopted a computer-based lexical decision task as a measure of word recognition speed. Also in a study by van Gelderen and his colleagues (2007), the speed of word recognition in Dutch (L1) and English (L2) was measured by a lexical decision task. The task contained letter strings (3–8 letters), and students were asked to decide as fast as they could whether a letter string was an existing word. Half of the letter strings consisted of existing monomorphemic words; the other half of phonologically- and orthographically-legal pseudowords. Letter strings were presented on laptop computer screens. Accuracy and reaction times (RTs) of responses were registered automatically. For analysis, only correct responses on existing words were taken into account.

Regarding the accessibility of network knowledge, Cremer and Schoonen (2013) developed a computerized semantic decision task (C-WAT), based on the WAT Schoonen and Verhallen (2008) developed to measure Dutch monolingual and bilingual children's availability of semantic knowledge. For the C-WAT, children were presented with a stimulus word followed by two other words, and were to identify the one that was semantically related to the stimulus word as quickly and accurately as possible. Their reaction time, which reflected the speed of accessing the semantic relationships between words, as well as the accuracy of the responses were collected. The finding showed that C-WAT performance had unique and significant predictive power in explaining reading comprehension over and above availability (measured with the WAT developed by Schoonen and Verhallen, 2008) and word decoding (this study will be reviewed in greater detail in the next section). This suggested that C-WAT tapped a skill distinct from WAT or availability; and accessibility of network knowledge constituted a different dimension of vocabulary competence.

2.5.2. Accessibility of word knowledge and reading comprehension

Successful textual comprehension requires word knowledge, including both vocabulary size and depth. However, availability of knowledge alone does not guarantee efficient reading comprehension. In addition to word knowledge - word meanings and meaning relationships – how efficiently readers can access that knowledge also is critical in the comprehension process (Perfetti, 1985; Perfetti & Hart, 2002). Theoretically, Perfetti's (1985) Verbal Efficient Theory and, more recently, the Lexical Quality Hypothesis, contend that fast access to word knowledge allows available processing resources to be devoted to comprehension (Verbal Efficiency Theory) and high-quality representations of words drive such rapid processing (Lexical Quality Hypothesis), contributing to better comprehension. Mezynski (1983) also stressed an important role of access or automaticity of word knowledge in comprehension, further arguing that learners need to be provided with chances for efficient use of words for vocabulary instruction to be successful.

In contrast to ample research on the efficiency of accessing knowledge of individual words, literature is limited on how the accessibility of lexical network knowledge may be (uniquely) important for reading comprehension. Nonetheless, a small number of studies did show that readers' efficiency in accessing knowledge of the relationship between words (e.g., synonymic relationship) could contribute uniquely to reading comprehension (Cremer & Schoonen, 2013; Ikeno, 2006; K. Nation & Snowling, 2004; Oakhill, Cain, & McCarthy, 2015). Cremer and Schoonen (2013) investigated the influences of word decoding, semantic word knowledge, and its accessibility on reading comprehension in both monolingual and bilingual children. Participants' availability of semantic word knowledge was measured with WAT (Schoonen & Verhallen, 2008). Accessibility of semantic word knowledge was assessed by a computerized semantic decision task (C-WAT), in which the participant was required to identify the semantically-related word as quickly as possible. Hierarchical regression analysis showed that the accessibility measure explained 2.5% of the unique variance in reading comprehension when controlling for the effects of (the availability of) semantic word knowledge and word decoding.

Oakhill et al. (2015) investigated the extent to which the relationship between vocabulary depth and reading comprehension was mediated by speed of access to word meanings in English-speaking children aged 9-10. They devised two ways to assess participants' word knowledge of synonym and hypernym. In the judgment tasks, the children saw, on a screen, words pairs (i.e., synonym and hypernym) and were asked to indicate if they were related (e.g. if "loud" and "noisy" mean the same in the synonym test, and if "lemonade" is a sort of "drink" in the hypernym test). In the production tasks, the children were required to provide a word that meant the same thing to a given word in the synonym test. They also were told to answer "type of" questions (e.g., "What is a pineapple a type of?") about a set of words given in the hypernym test. In both the judgment and the production tasks, children's response accuracy and speed were collected. Regression analysis revealed that children's judgment time in the synonym and hypernym tasks (i.e., accessibility of knowledge about semantic relationships between words) explained about 6.4% and 6.6%, respectively, in reading comprehension after controlling for their performance on the respective production task and other related variables. This suggests that accessibility of network knowledge is a legitimate construct to be examined in addition to its availability in the context of reading comprehension.

In a longitudinal study, Nation and Snowling (2004) investigated the reading development of 72 English-speaking children at two time points. Hierarchical regression analysis at Times 1 and 2 revealed that semantic skills tapped by semantic fluency and synonym judgement explained about 15.1% of the unique variance in reading comprehension at Time 1, when controlling for non-verbal ability, nonword reading, and phonological skills. Semantic skills at Time 1 also were predictive of reading comprehension at Time 2, over and above nonverbal intelligence, nonword reading, phonological skills, and reading comprehension at Time 1. Note that the unique variance explained by semantic skills was comparable to that by vocabulary size (4.5% *vs*. 4.9%, respectively), indicating vocabulary breadth and semantic depth (i.e., accessibility of knowledge of word relationships) may be equally important in predicting later reading comprehension. Considering individual differences in semantic processing influence differences in reading comprehension, difficulties in comprehension also might arise from inefficient processing skills of words as well as from lack of word knowledge.

The above findings about the unique contribution of semantic access efficiency, however, were not always produced in the literature. Some studies have shown that fluency of access to word meanings might not play a role in reading comprehension (Tannenbaum et al., 2006; van Gelderen et al., 2004). For example, in van Gelderen et al. study, researchers investigated different contributions of processing speed, vocabulary knowledge, grammar knowledge, and metacognitive knowledge of Dutch (L1) and English (L2) adolescent students. The speed of word recognition was tested with lexical decision tasks both in L1 and L2. Standardized regression weights from the analysis indicated that the speed of word recognition was not significant for either the explanation of L1 reading comprehension or for L2 reading comprehension. Vocabulary knowledge and metacognitive knowledge did make contributions to reading comprehension in L2, with relatively more significance of metacognitive knowledge. However, given that the correlation between word recognition speed and reading comprehension in L2 was substantial, despite no unique contribution to comprehension, the role of word recognition efficiency might be offset. Further studies definitely should work on this to verify the exact role of word recognition fluency in reading comprehension.

2.6. Word knowledge, its accessibility, and levels of reading comprehension

When reading, readers build multiple levels of mental representation for a text as a result of comprehension (Kintsch, 1988; Van Dijk & Kintsch, 1983): the *surface code*, *the textbase*, and *the situation model*. Surface code is a verbatim representation of a text where the meanings of individual words and phrases, and syntactic forms are represented as their exact wording. This representation is very short-lived and forgetful but is important because it is the basis for the textbase, where readers analyze the surface codes of the text into literal propositions (the smallest unit of idea/meaning) delivered by the text. Readers build textbase representations by drawing on the meaning of each proposition and integrating meanings across propositions. With this textbase, readers establish a situation of what is described in the text, or a situation model. The situation model goes beyond the explicit meaning of the text, representing the global meaning of the text that is constructed from integrating propositions (i.e., the textbase) and readers' world knowledge.

To address those different levels of representations with respect to readers' comprehension of a text, usually two types or levels of comprehension are distinguished: literal *versus* inferential. *Literal* comprehension is concerned with the reader's knowledge of the information stated explicitly in the text, which involves a textbase level of understanding. *Inferential* comprehension, on the other hand, requires the reader to go beyond the text level to integrate the information referenced in the text as well as combine his/her world knowledge with the information from the text (Eason, Goldberg, Young, Geist, & Cutting, 2012).

Research has found that different aspects of word knowledge and their accessibility may have differential levels of functions for different levels of comprehension (Cain & Oakhill, 2014; Eason et al., 2012; Oakhill et al., 2015; Zhang & Yang, 2016). Cain and Oakhill (2014), for example, explored how vocabulary breadth and depth (measured with definition and synonym tasks) on different types of inference making and literal comprehension with English-speaking children aged 10-11 in England. Different types of questions were developed to tap different aspects of comprehension. Literal questions were created to assess children's memories for locating literal information explicitly referenced in the text. Local cohesion questions were to tap the ability to generate local cohesion inferences by integrating two sentences in the text. Lastly, global coherence questions tapped the ability to produce global coherence inferences by integrating information in the text with general knowledge. Multiple regression analysis revealed that vocabulary knowledge in general predicted a greater proportion of variance in global coherence inferences than in local cohesion inferences. More importantly, results showed that different aspects of vocabulary knowledge differentially influenced inference making. Specifically, depth of vocabulary, rather than breadth, was found to be the more critical factor in explaining performance on global coherence inferences, over and above literal memory skills and word reading. Of note is that assessment of global coherence inferencing skills was based on relationships between different words and world knowledge such as setting, theme, or character identity. This suggested that this study's participants might have needed to depend on semantic network knowledge to form connections to their prior knowledge for better comprehension. Oakhill et al. (2015) later highlighted the importance of vocabulary depth and stated that "the rich and well-connected semantic representations of words will permit the rapid activation not only of the word meanings but also of related concepts, which can provide the basis for many of the inferences that are crucial for the construction of a coherent representation of a text" (p. 142).

To my knowledge, no study has considered dimensions of word knowledge and their accessibility together in relation to the levels of comprehension. Although, Tannenbaum et al. (2006) examined the relationships between three different dimensions of word knowledge (breadth, depth, and fluency) and reading comprehension with 3rd grade English-speaking students in the United States. Breadth was tapped by a multiple-choice, picture recognition task and an open-ended productive definition task, while depth was assessed by a measure of the ability to recognize multiple meanings and identify the diverse attributes (e.g., function, components, category, location) of words. For fluency, the WUF subtest of DIBELS (Good, Gruba, & Kaminski, 2002) was administered, which basically is an expressive vocabulary test to measure an ability to use given words quickly and correctly in sentences. Fluency measures also included a semantic categorization task asking children to name items in a particular semantic category (e.g., farm animals, fruits). Interestingly, Confirmatory Factor Analysis (CFA) revealed that the two-factor model (breadth and depth/fluency) provided a better fit to the data than the three-factor model where fluency was specified as separable from both breadth and depth. The result further showed that in this two-factor model, the unique contribution of depth/fluency to reading comprehension was minimal (2% of explained variance), and not significant, while the unique impact of breadth to reading comprehension was substantial (19% of explained variance).

At first sight, this finding seems to suggest it might be difficult to determine whether fluency, including efficient access of semantic relationships between words, is a construct distinguishable from either vocabulary breadth or depth, and whether it could have a unique role in reading comprehension. However, deeper analysis of the authors' measures suggests that such a conclusion may not hold. Specifically, the measures for vocabulary breadth and depth involved participants' word definition skills in one way or another, although the author only indicated it in the measure of breadth. For example, in the attribution task, a certain noun's specific attributes a child provides can easily be combined to define the word. So it is possible there was some shared variance between the breadth and depth measures, which perhaps resulted in some impact of depth offset by breadth. More importantly, while the researchers adopted time-sensitive measures for assessing the children's fluency, including naming words in the semantic categorization task (i.e., the accessibility of semantic network knowledge in the present context of discussion), those measures did not take into account the children's actual response times (such as in the case of the C-WAT in Cremer and Schoonen, 2013), which seems a more sensitive index of readers' actual efficiency of accessing their knowledge of semantic relationships between words.

That the level of comprehension may moderate the contribution of its correlates, including (different dimensions of) word knowledge, has received little attention in the L2 reading comprehension literature (E. H. Jeon & Yamashita, 2014). And studies with accessibility measures included together with availability measures are even fewer, if not nonexistent. Two recent L2 studies suggested that, like the findings of some L1 studies, such as Cain and Oakhill (2014), relative contributions of vocabulary depth and size to L2 reading comprehension may depend on the reading tasks in question or questions that tap different types or levels of comprehension. For example, Li and Kirby (2015) explored the relationship between vocabulary breadth and depth, and their effects on different aspects of English reading comprehension measured with multiple choice questions (Gates-MacGinitie Reading Comprehension Test, Level C; MacGinitie & MacGinitie, 1992) and summary writing. Chinese high school students in

English-immersion classes participated in this study. Their vocabulary breadth was assessed with Gates-MacGinitie Vocabulary Subtest (Level C, Grade 3), while vocabulary depth was gauged with multiple instruments measuring precision of knowledge, polysemy, and morphological awareness. A series of hierarchical regression analyses revealed that only vocabulary breadth predicted Gates-MacGinitie Reading Comprehension significantly over and above vocabulary depth, whereas only vocabulary depth predicted summary writing after controlling breadth. This finding makes good sense because in summary writing, one needs to employ one's deep vocabulary knowledge such as selecting appropriate words to meet the purpose of using the words in the context and providing correct definitions and forms of the words. All of these seem to require one to have higher level of processing and good memory skills.

It is noted that vocabulary depth measure used in Li and Kirby (2015) did not focus on network knowledge although they extensively conceptualized depth of vocabulary. Such an issue was further addressed in Zhang and Yang (2016) with a focus on adult L2 learners of Chinese. To measure participants' reading comprehension, they used two tasks with passages of different lengths, long and short, followed by multiple-choice questions. Questions for the long-passage reading task were mostly about literal comprehension, whereas those for the short-passage reading task assessed inferential comprehension. Among other findings, vocabulary depth, measured with a Chinese WAT developed by researchers themselves following Read (1998), was found a significant and unique predictor of reading comprehension over and above vocabulary size, which was measured with a picture selection task. More importantly, it was found that the (relative) contributions of the two dimensions of vocabulary knowledge to reading comprehension differed depending on what types of texts were read or what levels of comprehension were assessed. Specifically, vocabulary size was more predictive of long-passage comprehension with a focus on literal comprehension, whereas vocabulary depth was more important for short-passage comprehension with a focus on inferential comprehension.

This finding is interesting because although it is in agreement with previous findings that vocabulary size is sufficient for literal level of processing, the text itself (i.e., long passage) was "lexically much richer and grammatically more complex" (p. 9). Reading such a passage would have required readers to depend on their lexical-semantic knowledge (i.e., vocabulary depth) at the time of reading, to better comprehend the story. In other words, findings about the relative contributions in this study might result from the "convergence" of literal comprehension with lexically and grammatically more complex texts and inferential comprehension with lexically and grammatically simple texts. Thereby, to obtain a clearer picture of the relative contribution of vocabulary size and depth, there seems a need for texts of similar lexical and grammatical complexities, but focusing on different types/levels of comprehension. Specifically, it would be interesting to include inferential questions in long-passage comprehension tasks and literal questions in short-passage comprehension tasks in their study, to see how different dimensions of word knowledge work on different levels of comprehension, or how text types and question types interact.

In addition, Zhang and Yang (2016) did not differentiate between paradigmatic and syntagmatic association in measuring vocabulary depth. As previously discussed, it would be theoretically and psychometrically more appropriate to have paradigmatic and syntagmatic associations scores reported and analyzed separately, than to aggregate them to form a single score to index vocabulary depth, to explore the contribution of depth/network knowledge to reading comprehension.

2.7. The present study and research questions

The foregoing review of literature suggests that different dimensions of word knowledge (size *versus* depth) and their respective accessibility presumably play distinct roles in reading comprehension; their contributions to reading comprehension may be also a function of the levels of comprehension in question. However, little if any research has examined the complex interplay of those knowledge availability and accessibility variables in EFL reading comprehension - and L2 reading comprehension in general - with consideration of different levels of comprehension.

To recap earlier discussion, several major issues warrant further research. First, despite previous studies' attention to different dimensions of vocabulary knowledge (e.g., size and depth), much still remains unclear regarding their relationship and their respective contributions to reading comprehension (Li & Kirby, 2015; Qian, 1999, 2002; Vermeer, 2001; Zhang & Yang, 2016). Moreover, when vocabulary depth was operationalized as network knowledge and measured with a task of Read's (1998) WAT type, an aggregated score for paradigmatic and syntagmatic relationships usually was produced to predict reading comprehension (e.g., Horiba, 2012; Qian & Schedl, 2004; Zhang & Yang, 2016), which possibly could have obscured any distinct contributions of those association relationships to reading comprehension.

Second, accessibility theoretically pertains to activation of both form-meaning connections (i.e., vocabulary size) and lexical network knowledge (i.e., vocabulary depth), and reasonably it should be considered a separate layer of language skill other than the availability of knowledge itself. However, accessibility rarely has been framed this way in previous studies. When it was so framed (e.g., Tannenbaum et al., 2006), findings on its relationships with vocabulary size and depth on one hand and their relationships with reading comprehension on the other sometimes warranted more careful interpretations because of construct measurement concerns.

Third, while a few recent studies suggested that different dimensions of vocabulary knowledge may contribute differentially to reading comprehension depending on levels of comprehension in question (e.g., literal *versus* inferential) (e.g., Li & Kirby, 2015; Zhang & Yang, 2016), research overall is limited, not to mention that those studies did not consider accessibility of knowledge and its concurrent functioning for different levels of comprehension.

Finally, despite strong attention to fluency of recognizing individual words - and less so to access of knowledge of semantic relationships between words - in reading comprehension, the ‘time-controlled’ methods used in those studies, such as timed word decoding, often could not indicate reliably the target skill among L2 learners (e.g., Cremer & Schoonen, 2013). To address accessibility, it seems more desirable to employ computerized measures involving response time or latency (as well as accuracy).

To fill gaps in previous research and enrich our knowledge about the lexical basis of L2 reading comprehension, the present study aims to examine how different dimensions of L2 word knowledge and their respective accessibility function concurrently in explaining different levels of L2 reading comprehension in adolescent EFL learners in Korea. Specific research questions are the following:

1. What is the construct structure of lexical competence represented by different dimensions of word knowledge and their respective accessibility?
2. What are the relationships of dimensions of word knowledge and their respective accessibility with reading comprehension?

3. Do the contributions of different dimensions of word knowledge and their respective accessibility to reading comprehension vary as a function of the level of comprehension?

CHAPTER THREE

METHOD

3.1. Participants

Participants were 116 11th-grade students from five classes in a girls' private high-school in Seoul, South Korea. In this school, 11th-graders had five 50-minute English classes per week, including reading, listening, speaking, and writing although most class time was used in reading and grammar. One English teacher in this school evidenced that students' overall English proficiency had been slightly below average in a nationwide English proficiency test. Students had studied English since 3rd grade, under the Korean national curriculum, in which English language learning was highlighted as a critical global communication skill (M. Jeon, 2009).

3.2. Instruments

All instruments administered in this study are summarized in Appendix J. Details of each instrument are provided below. Task reliability (Crobach's α) is presented in Table 2.

3.2.1. Availability of word knowledge

3.2.1.1. Vocabulary breadth

To measure participants' vocabulary size, Nation's Vocabulary Size Test was used (VST; see I. S. P. Nation & Beglar, 2007 for a description of the test). Nation provided bilingual versions of the test including a Korean version (download available at http:// www.victoria.ac.nz/lals/staff/ paul-nation.aspx), and suggested a bilingual version is preferable to its corresponding monolingual version (English only) especially for lower-proficiency learners. For this study, the Korean bilingual version of VST was adopted.

VST was designed as a test of written receptive vocabulary size. It included items ranging from the first to the 20th 1000-word families of English based on the British National

Corpus (BNC). Items were given in a multiple-choice format with the target word presented in a short, nondefining context. Students were asked to choose the meaning of the target word from four options presented in Korean.

Considering participants' English proficiency based on consultations with two English teachers at the school, only the 2nd, 3rd, and 4th 1000-word families were included. First and fifth (and above) 1000-word families were not included, due to item difficulty. The test was paper-based and the total number of items was 30. At one point per correct answer, maximum possible score was 30 (Appendix A).

3.2.1.2. Vocabulary depth

To measure participants' depth of vocabulary knowledge, I modified Read's (1998) Word Associates Test (WAT). In Read's WAT, a target word (e.g., *sudden*) is presented with eight words in two different boxes, paradigmatic associates on the left, syntagmatic associates on the right. Among these eight words, test-takers are asked to select four associates of the target word. The four words in the left box are adjectives, and the associates are either synonyms of the target word or represent one aspect of the various meanings of the target word. The right box includes four nouns, and the associates are ones that can collocate with the target adjective. Each box considers a different aspect of word association (paradigmatic *vs*. syntagmatic), but WAT neither assesses them separately nor collects a separate score for each. However, for this study, I designed two separate tests for the two different types of word association (paradigmatic vs. syntagmatic) following the format of the items in the WAT.

In designing the paradigmatic and syntagmatic associates tests, an important consideration was that the words presented in the tests, whether target words, associates or answers, or distractors, would need to be known or familiar to the participants. If one encounters a target word one does not know, one would be unable to choose its associates, despite one's familiarity with all words presented in the options. Likewise, students need to be familiar with all words in the options as well including both associates and distractors.

To deal with this, discretion was given to the process of selecting words to be used in the test. To ensure words presented in the test were as familiar as possible to the students, three different sources were used. First, I decided to use only the first and second 1000-word families of BNC as a 'word reservoir' for the test development. BNC is organized based on word frequency, which indicates that the first 2000 words are the most frequently used words in this corpus. Although frequently-used words are not necessarily 'easy' words, particularly for EFL students, easy words are likely to be included in this list.

Second, I referred to Basic English Vocabulary (BEV) List documented by the Korean Ministry of Education. This is a guideline vocabulary list included in the National Curriculum of Korea published in 2015 for the purpose of teaching and textbook development. In total, 1800 words are listed, based on school levels including elementary-, middle-, and high-. I ensured all words in the association tests were included in this list.

Third, I consulted with two English teachers in Korea, both currently teaching 11^{th} graders. I asked them to review the words in the tests to see if any of them could be unfamiliar to most of their students. If any word that met the first two conditions was indicated as 'might-be-difficult' by either one of the teachers, the item was replaced.

More specifically, in developing the paradigmatic associates test (PAT), I first listed 2000 words collected from the BNC in random order (originally alphabetically ordered). The first 40 adjectives in the randomized list were selected as candidate target words to be used in the test.

Those words all appeared in the BEV list. Eventually, 20 of them judged by both English teachers I consulted as the most familiar to their 11th graders, were selected as target words.

Care also was taken in designing the associates and distractors for the PAT. Paradigmatic associates of a target word are either its synonyms or words that indicate one of its various meanings. To select option words for the 20 target words decided upon, I referred to three online dictionaries [1] for associates and the BNC high-frequency words list (i.e., 2000 words) for distractors. During the designing process, I cross-referenced between BNC and BEV lists to ensure all my choices (for associates and distractors) were in those.

After familiarity-check consultation with the two English teachers, a first draft of the PAT was developed and later sent out as an online version in Qualtrics, an online data collection platform, to 20 native English speakers for a pilot test. Those native speakers were undergraduates in the College of Education at Michigan State University; 12 responded. Average agreement rate on answers for all 20 PAT items reached 98.8%. After minor revision based on the responses (i.e., inconsistent responses for an item) and comments from those students, a final version of the PAT was developed (Appendix B).

The PAT was administered as a paper-based test, with its directions given in Korean. It had 20 items, each having four options with two associates and two distractors. Students were asked to choose either a synonym(s) of a target word (e.g., *sudden)* and/or a meaning(s) of the target word (e.g., *quick* and *surprising*). Participants would need to have both associates selected to earn a point for an item. At one point for each item, the total possible score was 20.

---

[1] https://www.merriam-webster.com/
http://www.dictionary.com/
http://www.thefreedictionary.com/

For the syntagmatic associates test (SAT), I selected adjective+noun collocations as target associations among other collocational contexts such as verb+noun or verb+preposition. This was particularly done to maintain consistency with the PAT regarding the part-of-speech of the target words. In developing items for the SAT, the same procedure as that of the PAT was applied in terms of word selection. First, as candidates for target words, I collected another set of 40 adjectives randomly selected from the BNC high-frequency words list; all those words appeared also in the BEV list. After consultation with the two English teachers, 20 most-familiar words were selected as target words.

To identify nouns that come after the target adjectives and make appropriate collocations (thus should be selected as associates), I used the Corpus of Contemporary American English (COCA, http://corpus.byu.edu/coca/). The website provides "collocates search" showing words occurring near a search word, based on frequency in the corpus. As answers I selected the two nouns most frequently occurring after each target adjective. As in the PAT, all nouns/options in the final draft of the SAT, including both associates and distractors, were listed in both the BNC high-frequency list and the BEV list, and both English teachers confirmed them as familiar to most of their students.

Particularly difficult in developing the SAT was to identify "appropriate" distractors or nouns that make no sense or sound awkward when placed after a target adjective. To illustrate, according to COCA collocates search, *happy* came with *birthday, ending, days,* or *hours,* which all made sense and frequently used in general. However, what about *happy garbage*, or *happy chair*? In daily life, these phrases might be rare usages (even weird), but might be encountered in a fairy tale or a joke. So it was assumed that finding nouns completely without context-dependent interpretations together with an adjective would be virtually impossible. To accommodate this issue, I used the phrase ‘selecting the most appropriate and frequently used collocations in general’ instead of ‘finding the correct answers’ in the direction of the SAT (in Korean).

A draft of the SAT also was pilot-tested on the 12 native English speakers who responded to the PAT. Average agreement rate on the correct answers was 99% across the 20 items. Again, minor revision was made reflecting native speakers’ responses and comments to produce the final version of the SAT to be administered in this study (Appendix C).

The SAT also was paper-based with directions presented in Korean. It had 20 items; each having two answers out of four options. Students were instructed to select words making the most appropriate and frequently used collocations with the target in general (*e.g., change* and *noise* as answers*; doctor* and *school* as distractors*)* when placed after the target adjective (*e.g., sudden*). Same as the PAT, participants would need to have both associates selected to earn a point. The maximum score possible was 20.

3.2.2. Accessibility of word knowledge

To measure how word knowledge was *accessible* to participants or how efficiently they could access the knowledge they possess, I designed different sets of computer-based tests using the software *Paradigm Stimulus Presentation* (Perception Research Systems, 2007) (commercially available at http://www.paradigmexperiments.com). *Paradigm* is a millisecond-accurate stimulus presentation program automatically recording participants’ response time when programmed.

Tests for accessibility of word knowledge comprised three different timed-tests: a lexical decision test, a timed paradigmatic associates test, and a timed syntagmatic associates test. As in the word associates tests (i.e., PAT and SAT), I undertook the same word selection process allowing that all words presented in these three tests were familiar to the participants (i.e., presumed knowledge of words) so that the tests would measure the efficiency of their access to the lexicon, the intended construct to measure, rather than their actual knowledge of those words (i.e., knowledge availability). To accommodate this consideration, three criteria I established for the PAT and SAT (i.e., included in the BNC high-frequency words list and the BEV word list, and confirmed by the two English teachers as being familiar to students) also were applied in the process of developing the accessibility tests described in detail below.

3.2.2.1. Lexical decision test

To measure how rapidly students can access words in their lexicon, I developed a lexical decision test (LDT). Students were asked to answer whether or not a given word existed, and each student's response time to each question, in addition to accuracy, was recorded. The test included 25 extant words and 10 non-extant. The former were selected randomly from the first 2000-word families of BNC. To generate the latter, *The ARC Nonword Database*[2] (Rastle, Harrington, & Coltheart, 2002) was used. Each non-extant word had 3-7 letters so that word length would be similar to those of extant words.

In the LDT, all items were ordered randomly and presented on a computer screen, with a prompt presented in Korean in the middle, reading "Does [target word] exist in English?" Via two identified keys on a keyboard students should choose Yes (existing) or No (non-existing) as quickly as possible. Two sample items were provided before the actual test. Once one hit the answer for an item on the keyboard, the computer recorded the response time (RT) and

---

[2] *The ARC Nonword Database* is available at http://www.cogsci.mq.edu.au/research/resources/nwdb/

automatically showed the next item. If within 3 seconds no answer was indicated for an item, the computer automatically would record an incorrect answer and move to the next item.

To gain scores for LDT RT measure, only RTs of correctly responded items (73% mean accuracy) were considered because response time, or accessibility, was defined as how efficiently one can access the knowledge they already know. (See Appendix D for words in the lexical decision test.)

3.2.2.2. Computer-based word associates tests

To measure how efficiently students could access their knowledge about word associations or semantic network knowledge, I developed two timed-word associates tests, both of which were simpler versions of the paper-based PAT and SAT previously explained. As in the PAT and SAT, all words in these two timed-word-associates tests, including target words, associates, as well as distractors, were selected randomly from the first 2000-word families in BNC and all cross-checked with the BEV list. Again, the two English teachers confirmed all words as being familiar to their students.

For items for the computer-based PAT (CPAT), I designed 25 pairs of associated words. The three online dictionaries previously mentioned were used again to check the relation between words. I also added 10 pairs of non-associated words in the test to serve as fillers. Students were asked to choose whether the two words in each pair were synonymous or semantically related. For example, *laugh:smile* should draw the answer Yes, because they are synonyms. The word pair *risky:long* is an example of the fillers (with a NO answer), because there seems no (or rare) meaning relation between them. When this test was piloted on the same native English speakers who had responded to the PAT and SAT, average agreement rate on the answers was 81.73% for the associated pairs and 98.15% for the fillers. Word pairs resulting in inconsistency in native-speakers' responses were revised or replaced by new pairs.

In CPAT, each student sat in front of a computer and gave their answers by pressing one of the two identified keys as quickly as possible within the 3-second fixed time limit: Yes for associated words; No for fillers. Two sample questions were introduced for practice before the actual test. In total, 35 test items, including 25 associated words and 10 fillers, were randomly ordered and presented. Once one hit an answer, the response time was saved, followed by the next item without delay. If within 3 seconds no answer was entered, the computer automatically moved to the next item. As was done in LDT, RT scores for CPAT was calculated only with the items correctly responded, and they were based on 63% mean accuracy. (See Appendix E for word pairs in the CPAT.)

Similarly, for the computer-based SAT (CSAT), I created 25 adjective+noun collocations and 10 fillers using collocates search in the COCA. Students were asked to indicate whether the given collocations were extant expressions or frequently used in general. For example, *happy hour* is generally accepted, whereas *handy smile* not. When I piloted these items with the same 12 native English speakers, average agreement rate on answers was 94.33% for appropriate collocations and 95% for fillers. Again, where inconsistency was found in native speakers' responses, I revised or replaced the items.

As in the CPAT, in the CSAT each student worked independently with an individual computer, and had two practice items before the actual test. The 25 target collocations and 10 distractor items were presented in random sequence on the computer screen. Students chose either Yes or No on the keyboard as quickly as possible within a fixed time limit (3 secs). Like the LDT and the CPAT, RT scores for CSAT were calculated only with the items correct (70% mean accuracy). (See Appendix F for items in the CSAT.)

3.2.3. Working memory

While this study focuses on the role of vocabulary knowledge in reading comprehension, previous research has revealed other important factors also influencing reading comprehension, one being working memory capacity (WMC), which correlates significantly with L2 reading comprehension (Harrington & Sawyer, 1992; Leeser, 2007; Walter, 2004). To control such impact of WMC, students' WMC also was measured and included as a covariate when modeling the contributions of lexical competencies to reading comprehension.

To measure WMC, I considered two word-memory tasks, a digit span task and an operation span task. The digit span task is one of the traditional and most-widely used test of working memory (Richardson, 2007), and measures for digit span (i.e., forward and backward span tasks) long have been a component of the widely-used Wechsler memory scales (WMS) and Wechsler intelligence scales for adults and children (Wechsler, 1997a, 1997b).

In addition to the digit span task, I included an operation span task (Turner & Engle, 1989) in this study. While the digit span task considers one's capacity to store or rehearse information, the operation word span task involves not only the ability to store information, but also the simultaneous processing of additional information, which is more likely to occur in real life (Conway et al., 2005). Considering this, Turner and Engle (1989) developed an operation span task requiring subjects to solve mathematical operations while trying to remember words, thus engaging processing functions of working memory and eliciting individual differences in task performance.

With this in mind, I developed two different tasks, a visual digit span task (DST) and an operation word span task (OWST) in such a way that the original format (aural/oral) of the digit span task and operation span task was represented visually on the computer screen using the software *Paradigm.* In DST, 24 sequences of numerical digits, with the number of digits increasing across trials, were presented on a computer screen, and students were asked to recall correctly 12 sequences (3 sets of 2-5 digits) forward and 12 sequences (3 sets of 2-5 digits) backward. Within each sequence, digits were chosen at random from 0 to 9 with replacement.

Students began the task with a fixation marker ("+") to appear in the center of the screen for 1000ms. Immediately at the offset of the fixation marker, a stimulus (digit sequence) appeared on the screen, and each digit in a sequence stayed on for a fixed rate of 1000ms. The fixation marker appeared again at the end of the sequence for 1000ms, followed by the on-screen question in Korean: "Is this (digit sequence presented in the center) the number you saw in the given (or reverse) order?" Students were asked to respond whether the digit sequence on the screen was what they recalled in the given (forward digit span) or reverse (backward digit span) order by hitting Yes or No on the keyboard as quickly as possible. Once the decision was made, it triggered the appearance of the next item, starting with the fixation marker again. Students' answers and response times were recorded. A no-key-press within 5 seconds was considered incorrect, and led to the next sequence (preceded by a fixation marker). Two practice items were given before the actual test.

In the OWST, students were asked to solve a simple mathematical operation (e.g., 2*1 +1 = 3; 6/3-2=0) then remember a simple Korean word (e.g., 책 (book), 열쇠 (key), 고양이 (cat)) that followed. Twelve sets in total were presented. Each three sets included two, three, four, and five operation-word [OW] pairs given in ascending order. For example, in the first three sets, two pairs of OW [OW-OW] sequence were presented. In the next 3 sets, three pairs of OW [OW-OW-OW] were introduced. Likewise, the remaining two sets of the three contained four and five OW pairs, respectively.

Again, all OW sequences were preceded by a fixation marker ("+") to appear in the center of the screen for 1000ms. In each operation, students were required to answer whether the equation was correct by pressing Yes or No on the keyboard as quickly as possible, within a 5000ms limit. A key strike prompted presentation of the to-be-remembered target word for 1000ms, followed by the next OW sequence. When the OW sequence was presented in each trial, the fixation marker appeared again at the end of the sequence for 1000ms, followed by the screen question in Korean: "Are these (word sequence presented in the center) the words you saw in a given order?" Students were asked to decide whether the words on the screen were what they recalled from the OW by indicating Yes or No on the keyboard as quickly as possible. Their answers and response time were stored.

Regarding scoring for the OWST, multiple data points could be collected from both the processing component (i.e., math problems) of the task and the storage component (i.e., word recall), such as accuracy on the math problems, response latency spent solving/processing the math problems, and accuracy on word recall and response time spent on the word problem. For this study's purpose, only accuracy and response latency on word recall were considered, following the common procedure of not considering processing performance (i.e., scores in operations) in the working memory span tasks (Conway et al., 2005). Evidence from correlation-based studies has shown that subjects' performance on the processing component typically reveals a ceiling effect, and it usually correlates positively with performance on the recall component. Processing component serves more as a 'secondary' task, to ensure that subjects actually are attending to the task. (See Appendix G for items for the WMC test).

For analysis, I calculated WMC scores based on four measures: two of storage components (total number of correct responses in DST and OWST) and two of response times (response times for DST and OWST). Only reaction times for correct answers were taken into account. To obtain a WMC score, I referred to two studies on L2 reading comprehension (Leeser, 2007; Walter, 2004). I first calculated z-scores for the four components. The z-scores for reaction time were multiplied by -1 to consider that higher (or longer) response times indicated less-good processing. I gained final WMC scores calculated by averaging the four z-scores. Considering how WMC score was calculated, higher WMC scores indicated good processing. It was also expected that there would be a positive association between WMC score and RC.

3.2.4. Grammatical knowledge

Grammatical knowledge, an aspect of linguistic knowledge other than vocabulary knowledge, has been found to be an important factor that explains substantial variance in L2 reading comprehension (Grabe, 2009; Jung, 2009; Zhang, 2012). I developed a grammaticality judgment test (GJT) as a grammatical knowledge measure and later included this as another covariate when the contributions of lexical measures to reading comprehension were modeled. The test included 13 pairs of short and lexically-simple sentences; in each pair, one sentence was grammatically correct, the other grammatically incorrect. The sentences were adapted from DeKeyser's (2000) task where un/grammatical sentences with a wide range of English grammatical structures were provided to measure ESL learners' grammatical knowledge. The original task had 200 sentences under 11 grammatical structures including past tense, plural, third-person singular, present progressive, determiners, pronominalization, particle movement, and others. I selected one to two sentences from each structure for the current study and included two practice items as well. I kept most sentences in their original form, but adjusted proper nouns and some other vocabulary that might be culturally unfamiliar to Korean students.

The GJT was paper-based. Students were asked to judge whether each sentence was grammatically correct or not, and mark Yes or No on the test. A correct answer received one point, and an incorrect answer or a missing response received no point. Maximum possible score was 26 (See Appendix H for the grammaticality judgment test).

3.2.5. Reading comprehension

To measure students' reading comprehension, I used the Level 4 comprehension subtest of the Gates-MacGinitie Reading Test (GMRT-4, Fourth Edition, Form S; MacGinitie, MacGinitie, Maria, & Dreyer, 2000). The GMRT-4 was chosen because originally it considered the type of questions (literal *vs.* inferential) in each passage, which represented the different levels of comprehension this study examined. According to the Technical Report documented by the developers, in developing the Fourth Edition, questions are classified as literal if they can be answered by choosing a restatement of something stated explicitly in the passage. Questions are categorized as inferential if test takers could not answer by choosing restatement.

The GMRT-Level 4 comprised a series of 11 passages. Among those, three were removed after consultation with the two English teachers with whom I worked, mainly due to the difficulty of the passages and the time limit for administration. Finally, the test for this study included 8 passages with a total of 30 multiple-choice questions, 15 tapping literal comprehension and the other 15 tapping inferential comprehension. Students were asked to read silently each paragraph then answer three to six multiple-choice questions following each.

Maximum score was 30. (Because it is copyrighted material, this task is not included in Appendices.)

3.2.6. Word familiarity checklist

To cross-check if words included in the tests except VST (i.e., LDT, PAT, SAT, CPAT, and CSAT), I designed and administered a word familiarity checklist. The checklist included 40 words selected randomly from the words in LDT, PAT, SAT, CPAT, and CSAT. Students were asked to indicate by marking Yes or No whether they knew the given words. Although development of those tests underwent a rigorous process to ensure all words therein were familiar to participants (see 3.2 Instruments), this checklist would help check whether this was the case for most of the students, if not all. Average response rate for Yes for the 40 items was about 95%, which confirmed that the words were overall familiar to the participants (See Appendix I for the word familiarity checklist).

3.3. Data collection procedures

Data collection was administered in a computer lab in the school, which had about 30 computers available for 20-25 students' use at a time. Before those students entered the lab for testing procedures, each computer was set up with necessary programs and operationally tested.

The flowchart for the entire procedure is presented in Figure 2. All paper-based tests except the Word Familiarity Check were administered first, followed by the computer-based tests. In all tests, instructions were first explained to students, and practice questions provided. In those tests, students began with the vocabulary size test (i.e., VST), then sat the two word associates tests (i.e., PAT and SAT), followed by the grammatical knowledge test (i.e., GJT) and the reading comprehension test (i.e., GMRT). Paper-based tests (including instructions) occupied 80 mins.

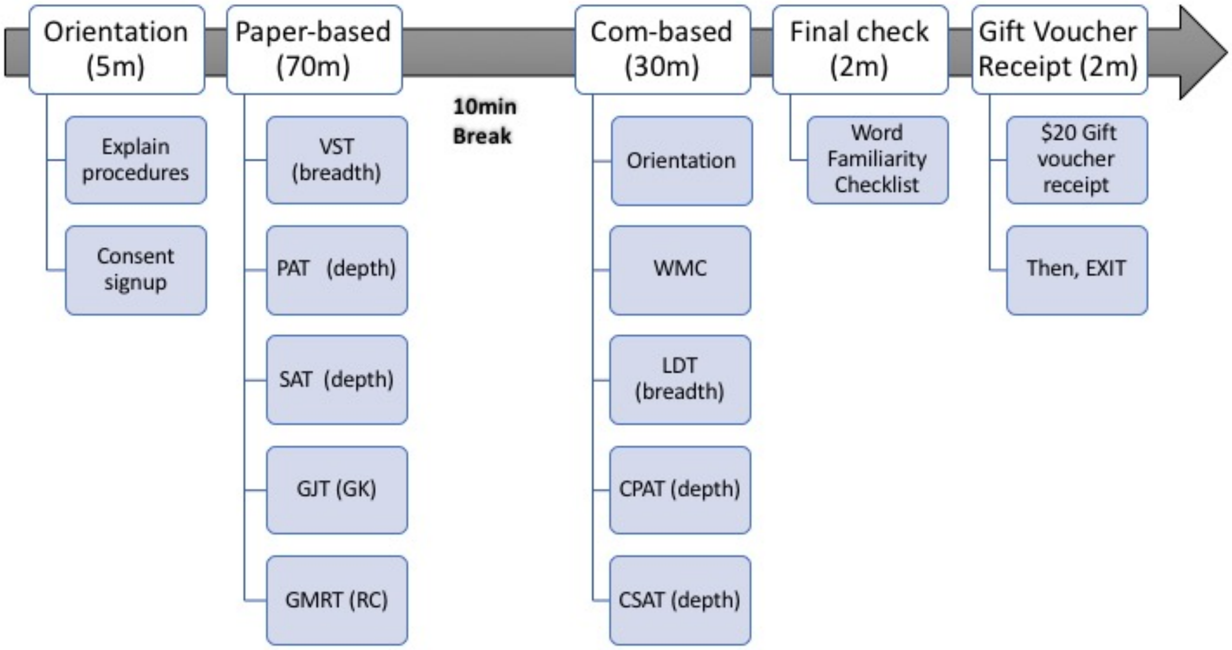

Figure 2: Flowchart for the Procedure of Data Collection.

After a ten-minute break, computer-based tests followed. In the computer-based tests, each student sat at a computer and worked individually. They particularly were cautioned to avoid unnecessary noise and interrupting others. Considering some students might be unfamiliar with computer-based tests, all procedures were explained, and demonstrated via on-screen shots of what they would encounter on the screen they then were at. Students began with the working memory capacity tests (i.e., DST and OWST), followed by the lexical decision test (LDT) and the word associates tests (i.e., CPAT and CSAT). Computer-based tests (including instructions) occupied about 30 minutes. As a final task after completion of all tests, students worked on a word familiarity checklist and left the classroom.

3.4. Data analysis methods and procedures

Before main analyses were conducted, the data set was examined for missing data, outliers, skewness and kurtosis, and basic assumptions for analysis within the general linear model framework. One participant who did not complete all tests was removed from the data set. Outliers, defined as cases whose scores were three standard deviations above or below the group mean, first were sorted out; each of them was inspected in scatter plots. Seven extreme cases were removed as outliers in this process. Thus 108 of 116 cases were kept for all analyses. The resulting skewness and kurtosis values for each variable fell within acceptable ranges ± 2 (Gravetter & Wallnau, 2014). The final data set met the assumptions for correlation-based analyses.

To address the first research question, *what is the construct structure of different dimensions of word knowledge and their accessibility?,* a series of Confirmatory Factor Analyses (CFA) was conducted on Amos 20 (Arbuckle, 2011) with a maximum likelihood estimation method. To judge the goodness of model fit, multiple indices were adopted for all models tested including χ^2 values, Comparative Fit Index (CFI), and Root Mean Square Error of Approximation (RMSEA). Following the recommendations of Hu and Bentler (1999), a cutoff value of .95 in Comparative Fit Index (CFI), and a cutoff value of .06 for the Root Mean Squared Error of Approximation (RMSEA) was used as indicators of goodness of model fit.

For the second and third research questions, *What are the relationships of dimensions of word knowledge and their accessibility with reading comprehension?* and *Do the contributions of different dimensions of word knowledge and their accessibility to reading comprehension vary as a function of the level of comprehension?*, the Structural Equation Modeling (SEM) method was adopted. SEM is a statistical technique that enables testing of a hypothesized model that represents structural relationships among a set of observed (measured) and/or unobserved (latent) variables (Kline, 2011). It allows for evaluation of the unique contribution of each independent variable to a dependent variable while controlling for all other variables. In each model, the same criteria for CFA were applied for the judgement of goodness of fit. Path coefficients for the effects of each dimension of word knowledge and its accessibility to reading comprehension were examined.

CHAPTER FOUR

RESULTS

4.1. Descriptive statistics and bivariate correlations

Means, standard deviations, skewness, kurtosis, maximum score possible, and correlations among variables are presented in Table 1. All predictor variables including grammatical knowledge and working memory capacity significantly correlated with reading comprehension and its two sub-levels. Among those, three accessibility measures (i.e., LDTRT, CPATRT, CSATRT), indicating how efficiently participants activated their word knowledge, negatively correlated with reading comprehension. Availability measures (i.e., VST, PAT, SAT), denoting participants' level of word knowledge, more strongly correlated with RC (r= .680 ~ .781) than did accessibility measures (r=-.290 ~ -.626). All availability and accessibility measures also significantly correlated with each other.

Reliabilities for all variables measured in terms of Cronbach's α were computed with SPSS 21 and reported as satisfactory overall (Table 2). Variables measured in computer-based tests (i.e., LDT, CPAT, CSAT, and WM) had two sets of scores, accuracy and response time. Thus for these variables, reliabilities for accuracy and response time were calculated and presented separately, although accuracy scores were not included in the analysis. Note that reliability for WM accuracy was .37. This low reliability seems to be mainly due to the relative easiness of the tasks involved (Cremer & Schoonen, 2013).

Table 1: Descriptive Statistics and Intercorrelations for All Variables.

	1	2	3	4	5	6	7	8	9	10	11
1. RC											
2. RC_LIT	.910***										
3. RC_INF	.920***	.673***									
4. VST	.680***	.697***	.551***								
5. PAT	.778***	.738***	.686***	.739***							
6. SAT	.781***	.743***	.688***	.756***	.857***						
7. LDTRT	-.290**	-.310***	-.223*	-.279**	-.299**	-.288**					
8. CPATRT	-.343***	-.360***	-.270**	-.223*	-.317***	-.333***	.463***				
9. CSATRT	-.626***	-.598***	-.549***	-.451***	-.534***	-.574***	.484***	.709***			
10. GK	.591***	.591***	.493***	.559***	.574***	.634***	-.106	-.200*	-.383***		
11. WM	.378***	.394***	.300**	.272**	.283**	.292**	-.221*	-.222*	-.340***	.179	
M	15.68	8.52	7.16	20.34	15.09	14.82	861.87	2031.22	1681.72	18.81	.00
SD	5.51	2.93	3.10	3.67	3.60	4.30	146.24	352.00	329.45	3.50	.56
Skewness	.06	-.146	.177	-1.075	-1.221	-1.194	.528	.269	.55	-.37	-.837
Kurtosis	-.658	-.488	-.698	1.641	1.19	.756	.195	.113	.461	-.285	.604
Maximum	30	15	15	30	20	20	---	---	---	26	---

Note. RC = reading comprehension; RC_LIT = literal comprehension; RC_INF = inferential comprehension; VST = vocabulary size test; PAT = paradigmatic associate test; SAT = syntagmatic associate test; LDTRT = response time for lexical decision test; CPATRT = response time for computerized-paradigmatic associate test; CSATRT = response time for computerized-syntagmatic associate test; GK = grammatical knowledge; WM = working memory capacity score. Response time was caulcauted based on test items with correct responses and represented in millisecond. *$p < .05$, **$p < .01$, ***$p < .001$.

Table 2: Internal Consistency Reliability for All Variables Measured.

		Cronbach's α
	Paper-based	
Accuracy	RC	.83
	RC_LIT	.68
	RC_INF	.74
	VST	.69
	PAT	.83
	SAT	.86
	GK	.66
	Computer-based	
Response time	LDTRT	.75
	CPATRT	.80
	CSATRT	.85
	WMRT	.84
Accuracy	LDT	.80
	CPAT	.77
	CSAT	.73
	WM	.37

Note. RC = reading comprehension; RC_LIT = literal comprehension; RC_INF = inferential comprehension; VST = vocabulary size test; PAT = paradigmatic associate test; SAT = syntagmatic associate test; GK = grammatical knowledge; LDTRT = response time for lexical decision test; CPATRT = response time for computerized-paradigmatic associate test; CSATRT = response time for computerized-syntagmatic associate test; WMRT_ response time for working memory capacity; LDT = lexical decision test; CPAT = for computerized-paradigmatic associate test; CSAT = computerized-syntagmatic associate test; WM = working memory capacity test. Reliability was calculated with Cronbach's α.

4.2. Evaluation of confirmatory factor analyses

To answer the first research question, a series of six confirmatory factor analyses (CFA) was conducted to examine the factor structure of different dimensions of word knowledge and their respective accessibility measured in the six lexical competence measures (i.e., VST, PAT, SAT, LDTRT, CPATRT, and CSATRT).

4.2.1. Step 1: Testing a single factor with multiple indicators model

To examine whether the six competence measures were manifestations of a single underlying construct, the first model specified the six measures as indicators of a single latent factor, labeled as Lexical Competence (LC) (Model 1 in Table 3; Figure 3). The model did not show a good fit with $\chi^2(9) = 84.60$ ($p < .001$), CFI= .802, and RMSEA= .280.

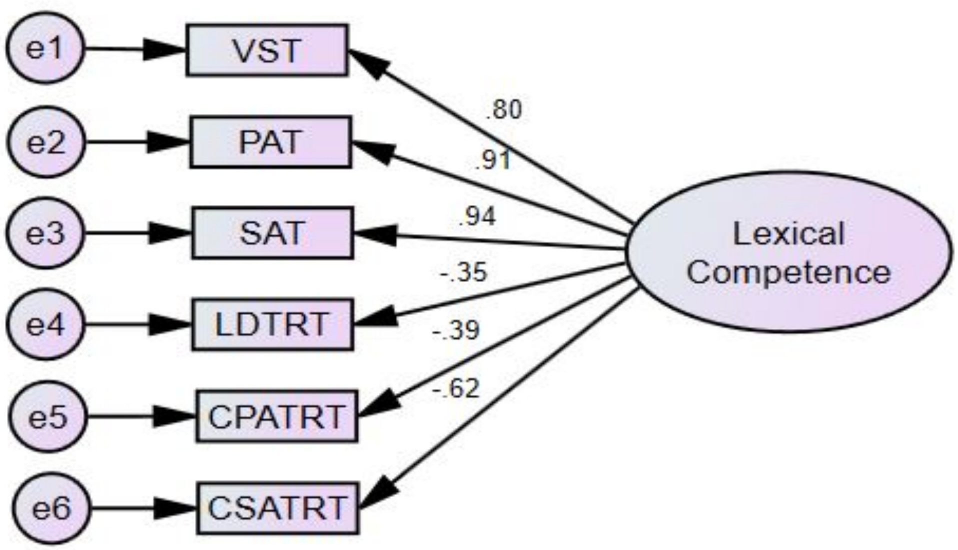

Figure 3: Model 1: One Factor Model of Lexical Competence.
Note. All factor loadings were significant (p<.05). VST = vocabulary size test; PAT = paradigmatic associate test; SAT = syntagmatic associate test; LDTRT = response time for lexical decision test; CPATRT = response time for computerized-paradigmatic associate test; CSATRT = response time for computerized-syntagmatic associate test.

4.2.2. Step 2: Testing two-factor models

The poor fit of Model 1 led to the next step testing two two-factor models (Models 2-3 in Table 3). As indicated in the literature review section, previous studies have distinguished between vocabulary breadth (or size) and depth, respectively, as the quantity and quality aspect of word knowledge (Anderson & Freebody, 1981; Read, 2004; Schmitt, 2014). However, whether these two types of word knowledge were indeed distinct concepts has been in question because of its strong correlation (Vermeer, 2001); sometimes, no significant unique contribution of depth to reading comprehension was found when breadth was controlled for (Horiba, 2012). Thus the need arose to test whether tasks employed in this study indeed, as intended, measured distinguishable constructs of breadth and depth.

Thus, in the first two-factor model (Model 2; Figure 4), it was examined whether vocabulary breadth measure (i.e., VST) and its accessibility measure (i.e., LDTRT) would load on a factor labeled as Breadth Competence, which referred to the knowledge and skills pertaining to individual word meanings, and whether vocabulary depth measures (i.e., PAT and SAT) and its accessibility measures (i.e., CPATRT and CSATRT) would load on a separate factor named Depth Competence, which referred to the knowledge and skills related to word meaning relations. This two-factor model did not indicate a good fit ($\chi^2(8) = 84.597$ ($p < .001$), CFI= 0.799, and RMSEA= .299), suggesting it would not be a good representation to put an availability measure and its corresponding accessibility measure together under a construct.

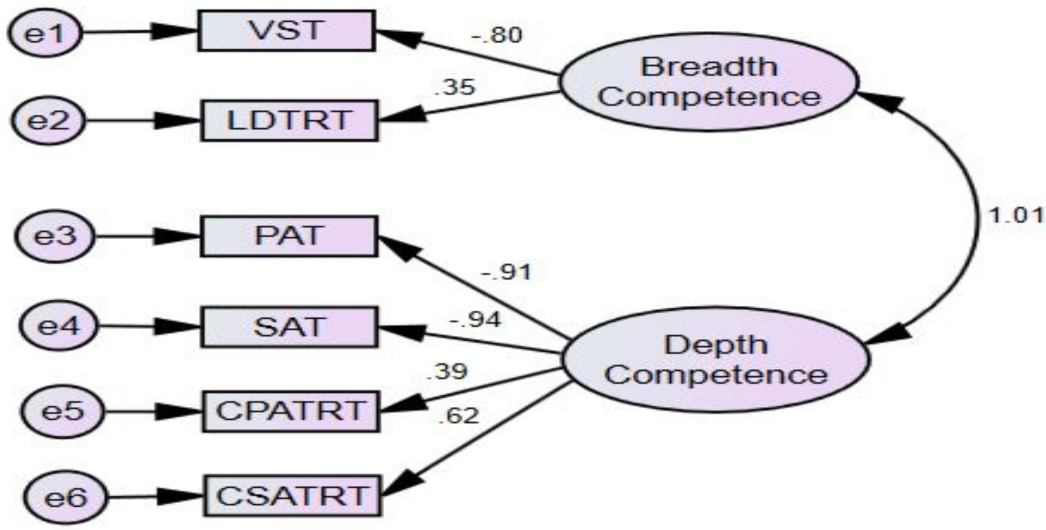

Figure 4: Model 2: Two-Factor Model of Breadth Competence and Depth Competence.
Note. VST = vocabulary size test; PAT = paradigmatic associate test; SAT = syntagmatic associate test; LDTRT = response time for lexical decision test; CPATRT = response time for computerized-paradigmatic associate test; CSATRT = response time for computerized-syntagmatic associate test.

In addition to the breadth *versus* depth distinction made in the literature, previous research also has suggested a distinction between knowledge one has (i.e., availability) and the efficiency of one's access to that knowledge (i.e., accessibility) (Cremer & Schoonen, 2013; Daller et al., 2007). This led to another two-factor model being tested specifying vocabulary knowledge/availability measures (i.e., VST, PAT, CAT) and their accessibility measures (i.e., LDTRT, CPATRT, CSATRT) as distinct yet covarying latent factors, labeled as Availability and Accessibility, respectively (Model 3 in Table 3; Figure 5). This model provided a very good fit to the data ($\chi^2(8) = 9.269$ ($p = .320$), CFI= .997, and RMSEA= .039). The correlation between Availability and Accessibility was r=-.601, p<.001. The result indicated that knowledge learners have about words (i.e., availability; knowing word meanings and meaning relations) and the efficiency of their access to that knowledge (i.e., accessibility) were negatively correlated but distinct constructs.

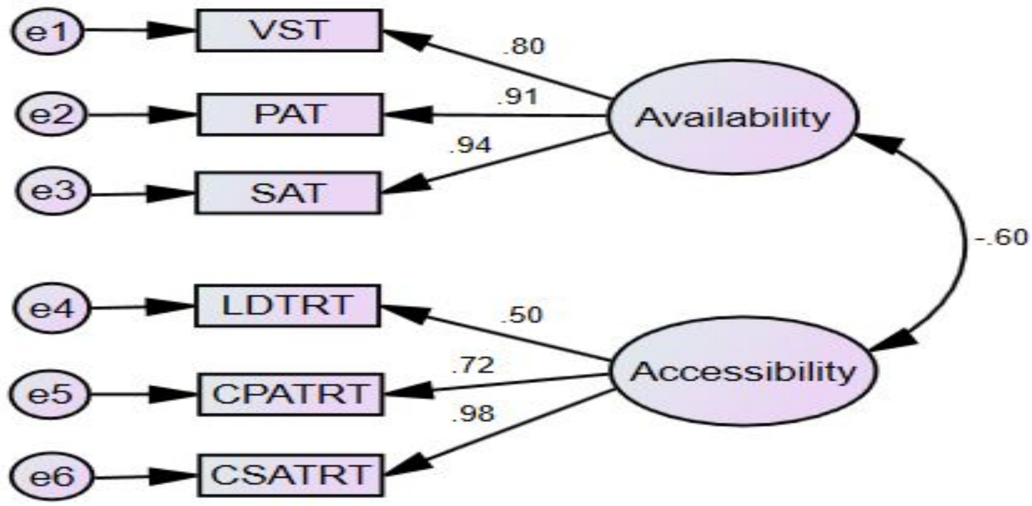

Figure 5: Model 3: Two-Factor Model of Availability and Accessibility of Word Knowledge.
Note. All factor loadings and the correlation were significant ($p<.05$). VST = vocabulary size test; PAT = paradigmatic associate test; SAT = syntagmatic associate test; LDTRT = response time for lexical decision test; CPATRT = response time for computerized-paradigmatic associate test; CSATRT = response time for computerized-syntagmatic associate test.

4.2.3. Step 3: Testing second-order factor models

The result of Model 3, that is, availability and accessibility of word knowledge were separate constructs, suggested that Model 2 could be reconceptualized in a way that Breadth and Depth Competence had their availability and accessibility indicators separately specified. In other words, four different competence measures – knowledge of individual word meanings and its accessibility, and knowledge of word meaning relations and its accessibility – could be specified independently in a model.

To this end, two second-order models subsequently were conceptualized and tested. The need for second-order models arose from the fact that availability (i.e., PAT and SAT) and accessibility (i.e., CPATRT and CSATRT) of depth knowledge (i.e., network knowledge or knowledge of word meaning relationships) were measured with two different instruments for each in this study (whereas availability and accessibility of breadth knowledge were measured with a single indicator for each). Thus, before testing a second-order model with all six different competence measures, a two-factor model was tested to examine whether PAT and SAT would load on a factor of availability of depth knowledge (Depth Availability) and CPATRT and CSATRT on the other factor of accessibility of depth knowledge (Depth Accessibility) with the two factors covarying (Model 4 in Table 3; Figure 6). This model showed a very good fit ($\chi^2(1) = .036$ ($p = .850$), CFI= 1.000, and RMSEA= .000). Correlation between the two latent factors was $r = -.54$, $p = .001$.

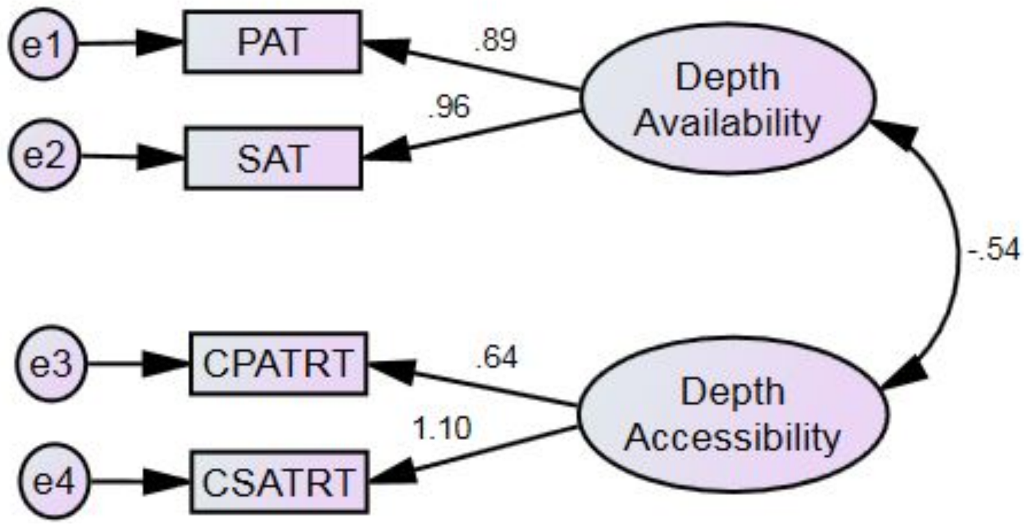

Figure 6: Model 4: Two-Factor Model of Depth Availability and Depth Accessibility.
Note. [3]All factor loadings and the correlation were significant (p<.01). PAT = paradigmatic associate test; SAT = syntagmatic associate test; CPATRT = response time for computerized-paradigmatic associate test; CSATRT = response time for computerized-syntagmatic associate test.

Given that Model 4 demonstrated a good fit, two higher-order structures were imposed subsequently with Model 4 embedded in each (Models 5-6 in Table 3). The first second-order model (Model 5 in Table 3; Figure 7) specified that knowledge of individual word meanings

[3] Note that the factor loading of CSATRT on Depth Accessibility is greater than 1. Although standardized coefficients typically are smaller than 1, they can be larger than 1 in magnitude when factors are correlated (Jöreskog, 1999).

(i.e., VST) and its accessibility (i.e., LDTRT), and Depth Availability (first-order factor represented by PAT and SAT) and Depth Accessibility (first-order factor by CPAT and CSAT) would load on a higher-order factor labeled as Lexical Competence. This model revealed a poor fit ($\chi^2(7) = 23.945$ ($p = .001$), CFI= .959, and RMSEA= .150). The second second-order model (Model 6 in Table 3; Figure 8) specified Availability and Accessibility as two distinct yet covarying second-order factors, and that vocabulary knowledge measures (i.e., VST, PAT/CAT) and their accessibility measures (i.e., LDTRT, CPATRT/CSATRT) would load on each of the second-order factors, respectively. This model showed a very-good fit ($\chi^2(6) = 8.769$ ($p = .187$), CFI= .993, and RMSEA= .066).

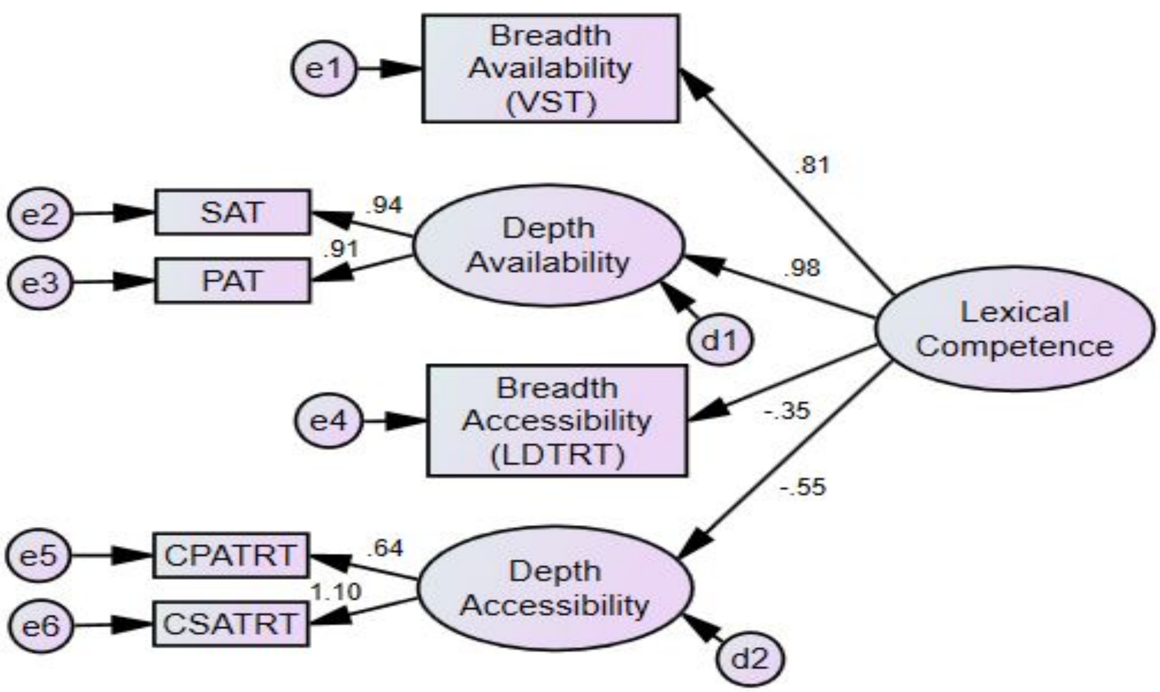

Figure 7: Model 5: Second-Order Factor Model of Lexical Competence.
Note. All factor loadings were significant (p<.01). VST = vocabulary size test; PAT = paradigmatic associate test; SAT = syntagmatic associate test; LDTRT = response time for lexical decision test; CPATRT = response time for computerized-paradigmatic associate test; CSATRT = response time for computerized-syntagmatic associate test.

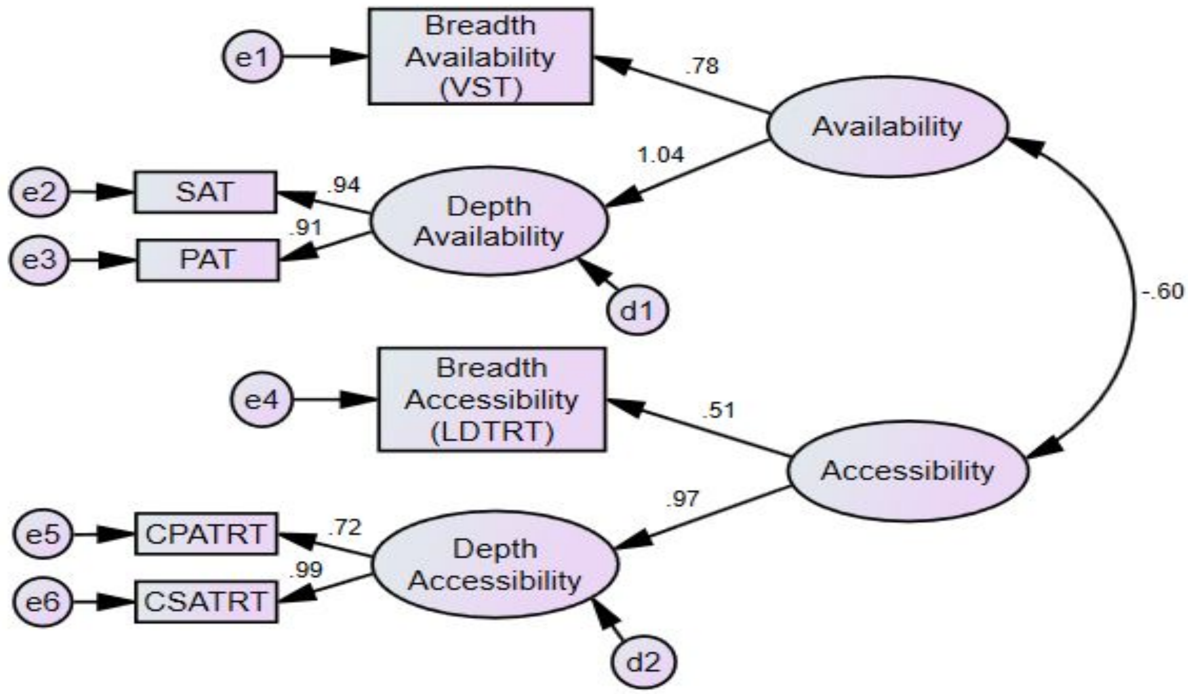

Figure 8: Model 6: Second-Order Factor Model of Availability and Accessibility.
Note. All factor loadings and the correlation were significant (p<.01). VST = vocabulary size test; PAT = paradigmatic associate test; SAT = syntagmatic associate test; LDTRT = response time for lexical decision test; CPATRT = response time for computerized-paradigmatic associate test; CSATRT = response time for computerized-syntagmatic associate test.

Summarized model fit indexes for all six models are presented in Table 3. As the summary shows, Model 3 with two factors of availability and accessibility and Model 6 with availability and accessibility as two distinct yet covarying second-order factors were identified with good fits. This indicated that availability and accessibility of word knowledge, that is, whether one has the knowledge about word meanings and meaning relations and how effectively one can activate the knowledge, are two distinct constructs which should be considered separately in conceptualization of lexical competence. Findings also suggested that lexical competence could be represented best in the second-order model with the availability and accessibility as higher-order lexical abilities, under each of which both breadth and depth knowledge were specified.

Table 3: Model Fit Indices for Confirmatory Factor Analysis (CFA).

	Model	χ^2	*df*	*p*	CFI	RMSEA
1	One-Factor Model: Lexical Competence	84.600	9	<.001	.802	.280
2	Two-Factor Model: Breadth Competence, Depth Competence	84.597	8	<.001	.799	.299
3	Two-Factor Model: Availability, Accessibility	9.269	8	.320	.997	.039
4	Two-Factor Model: Depth Availability, Depth Accessibility	.036	1	.850	1.000	.000
5	2nd-Order Factor Model: Breadth Availability, Breadth Accessibility, Depth Availability, Depth Accessibility under Lexical Competence	23.945	7	.001	.959	.150
6	2nd-Order Factor Model: Breadth Availability, Breadth Accessibility, Depth Availability, Depth Accessibility under Availability, Accessibility	8.769	6	.187	.993	.066

4.3. Structural model predicting reading comprehension as a whole

To investigate the relationships of dimensions of word knowledge and their accessibility with reading comprehension (RQ2), a series of structural equation modeling (SEM) analyses was performed. While Models 3 and 6 both had good model fits, they were not the models subsequently adopted to predict reading comprehension. Given the empirical and practical needs from the interest of this dissertation in how distinct lexical competencies pertaining to availability *vs*. accessibility as well as individual word meanings *vs*. meaning relationships between words contribute to reading comprehension, structural models specifying the four competencies separately without a second-order factor were needed.

In the conceptual model that addressed RQ2, the two availability variables (i.e., breadth [VST] and depth [PAT and SAT]) were allowed to covary, as was also the case for (the residuals of) the two accessibility variables (i.e., breadth accessibility [LDTRT] and depth accessibility [CPAT and CSAT]). The latent reading comprehension factor (RC) was indicated by comprehension scores on literal questions and inferential questions. As shown in Figure 9, all latent and non-latent lexical variables were hypothesized to predict RC. In addition, working memory capacity (WM) and grammatical knowledge (GK) also were included in the model to predict RC. WM also was set to predict accessibility variables.

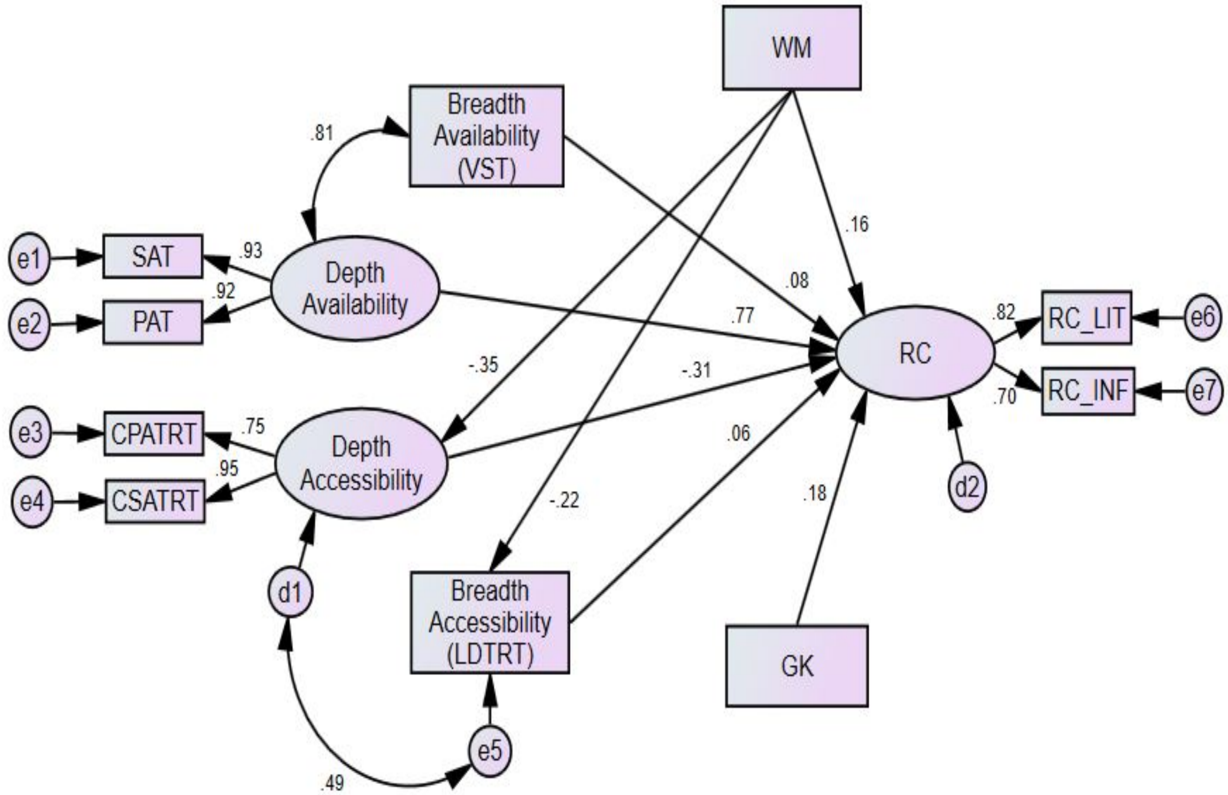

Figure 9: Standardized Structural Regression Coefficients among Working Memory Capacity, Grammatical Knowledge, Word Knowledge, and Reading Comprehension.
Note. RC = latent variable of reading comprehension; RC_LIT = literal comprehension; RC_INF = inferential comprehension; VST = vocabulary size test; PAT = paradigmatic associate test; SAT = syntagmatic associate test; LDTRT = response time for lexical decision test; CPATRT = response time for computerized-paradigmatic associate test; CSATRT = response time for computerized-syntagmatic associate test; WM = working memory capacity; GK = grammatical knowledge.

The initial model (Figure 9) did not yield a good fit to the data ($\chi^2(29) = 123.078$ ($p < .001$), CFI= 0.854, and RMSEA= .174). To find what provided a poor fit, modification indices (MI), produced by Amos for modification suggestions, were examined. Subsequently, two modifications were undertaken while all others were discarded on theoretical grounds (e.g., covariance between GK and residuals of CSATRT, or a predictive path from GK to SAT). They included removing GK as a predictor or RC and allowing Depth Availability (PAT and SAT) to predict Depth Accessibility (CPAT and CAST). The latter modification was conducted because a denser lexical network with knowledge of more word meaning relationships (Meara, 2009) reasonably would improve the efficiency of accessing those meaning relationships.

The modified model with the two aforementioned modifications revealed a good fit with $\chi^2(20) = 35.427$ (p =.018), CFI = .974, and RMSEA= .085 (Figure 10). Approximately 90.4% of variance was explained in reading comprehension. Taken together, this model was selected as the final fitting model to examine the predictive relations of lexical competence measures with reading comprehension.

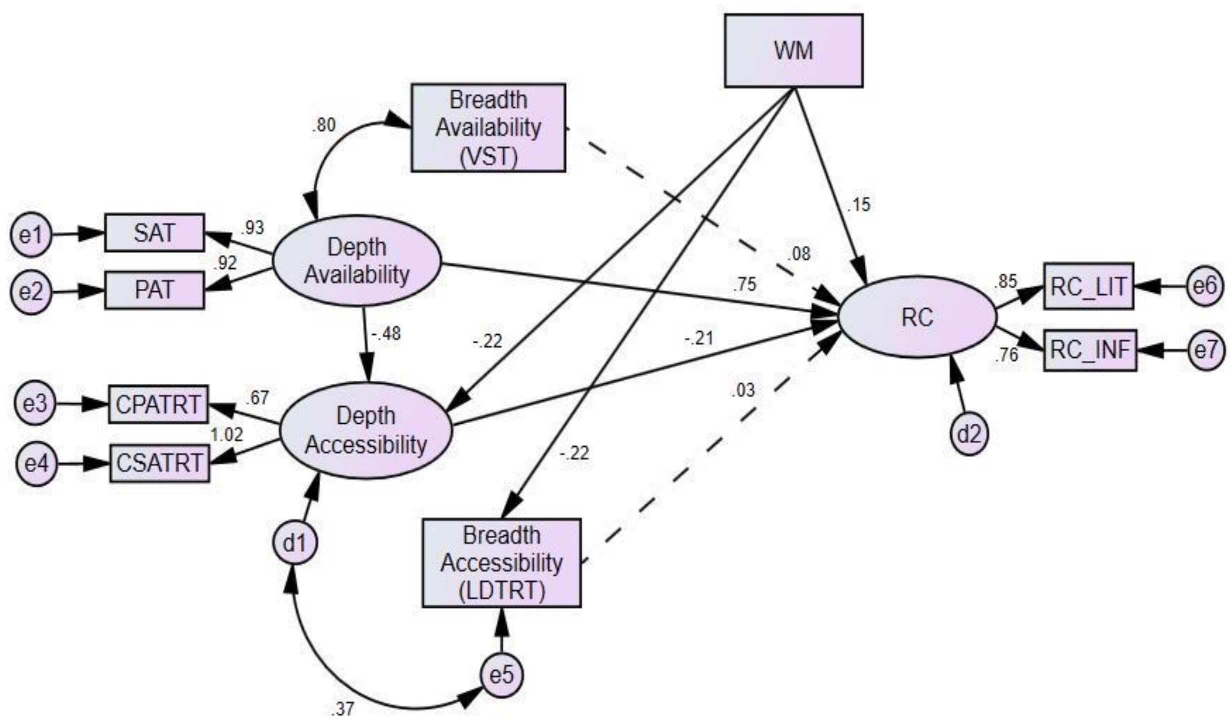

Figure 10: Standardized Structural Regression Coefficients among Working Memory Capacity, Availability and Accessibility of Word Knowledge, and Reading Comprehension.
Note. RC = latent variable of reading comprehension; RC_LIT = literal comprehension; RC_INF = inferential comprehension; VST = vocabulary size test; PAT = paradigmatic associate test; SAT = syntagmatic associate test; LDTRT = response time for lexical decision test; CPATRT = response time for computerized-paradigmatic associate test; CSATRT = response time for computerized-syntagmatic associate test; WM = working memory capacity. Solid lines indicate statistically significant paths and dotted lines statistically not-significant paths.

Parameter estimates of this model are presented in Table 4. Availability of breadth knowledge correlated highly with availability of depth knowledge (r=.80, p<.001) and accessibility of breadth knowledge also correlated positively with accessibility of depth knowledge but with a smaller magnitude (r=.37, p< .001). As Table 4 shows, working memory capacity significantly predicted reading comprehension (β=.15, p= .019). It also predicted significantly both breadth and depth accessibility variables but with negative directions (β=-.22, p= .019; β=-.22, p= .007, respectively).

In relation to reading comprehension, both availability and accessibility of depth knowledge contributed significantly to reading comprehension (β=.75, p<.001 for availability, β=-21, p< .05 for accessibility). Neither availability nor accessibility of breadth knowledge added any unique contribution to reading comprehension (β= .08, p= .469 for availability, β= .03, p= .644 for accessibility) once availability and accessibility of depth knowledge as well as working memory capacity were taken into consideration.

Taken together, these findings indicated that knowing meaning relations among words and accessing them efficiently played a more predictive role in explaining reading comprehension than did knowing individual word meanings and efficiently activating them. In addition, between availability and accessibility of depth knowledge, the β values also suggested that availability of depth knowledge was more contributive to reading comprehension than was accessibility of depth knowledge when other variables were controlled for. This implied that for participants in the present study, knowledge of meaning relations among words was more important for their text comprehension than was the ability to access the knowledge as quickly as possible.

Table 4: Parameter Estimates of the Structural Equation Model Representing the Relationship of Availability and Accessibility of Word Knowledge with Reading Comprehension.

Paths			*β*	S.E.	C.R. (z)	*p*
Structural Model (Path Coefficients)						
Breadth Accessibility	<---	WM	-.221	24.803	-2.342	.019
Depth Accessibility	<---	WM	-.217	45.889	-2.675	.007
Depth Accessibility	<---	Depth Availability	-.476	7.570	-5.972	<.001
RC	<---	Depth Availability	.753	.095	5.765	<.001
RC	<---	Depth Accessibility	-.207	.001	-2.409	.016
RC	<---	Breadth Availability	.080	.072	.725	.469
RC	<---	Breadth Accessibility	.031	.001	.462	.644
RC	<---	WM	.146	.270	2.344	.019
Measurement Model (Factor Loadings)						
PAT	<---	Depth Availability	.919			
SAT	<---	Depth Availability	.934	.076	16.061	<.001
CSATRT	<---	Depth Accessibility	1.020			
CPATRT	<---	Depth Accessibility	.674	.118	6.202	<.001
RC_LIT	<---	RC	.854			
RC_INF	<---	RC	.760	.103	9.192	<.001

4.4. Structural model predicting different levels of reading comprehension

To answer the third research question, that is, whether the contributions of different dimensions of word knowledge and their accessibility to reading comprehension vary as a function of the level of comprehension, another SEM analysis was conducted. The same structural model shown in Figure 7 was fitted to the data with the exception that the two levels of reading comprehension, that is, literal and inferential, were modeled separately (and with their residuals allowed to covary) instead of serving as indicators of a latent variable of reading comprehension as in the previous SEM analyses (see Figure 10).

This model yielded a good fit ($\chi^2(16) = 30.188$ (p =.017), CFI= .976, and RMSEA= .091) (Figure 11). Standardized structural regression coefficients are presented in Table 5 below. Results revealed that both availability and accessibility of depth knowledge significantly contributed to literal comprehension (β=.55, p< .001 for availability, β= -.16, p< .05 for accessibility) when availability and accessibility of breadth knoweldge and working memory capacity were accounted for. In contrast, neither availability nor accessibility of breadth knowledge was a significant predictor (β=.17, p= .12 for availability, β= .001, p= .99 for accessibility) when other impacts were controlled for. Similarly, availability (β=.73, p< .001) and accessibility (β= -.18, p= .05) of depth knowledge also significantly and uniquely contributed to inferential comprehension; and availability and accessibility of breadth knowledge were not significantly and uniquely contributive (β= -.11, p= .35 for availability, β= .06, p= .39 for accessibility). The model explained approximately 65% and 56% of total variance in literal and inferential comprehension, respectively.

Overall patterns for the two different levels of comprehension, which showed a stronger effect of vocabulary depth (and within vocabulary depth, availability), seem very similar to that obtained in the previous SEM analysis where reading comprehension was modeled as a latent variable (see Table 4 and Figure 10). It seemed that knowing meaning relations among words and activating the knowledge efficiently played a more significant role than knowing and quickly accessing individual words, in explaining both literal and inferential comprehension.

Despite the above-shared pattern, it should be noted that lexical competence pertaining to meaning relations among words, particularly knowledge availability (β= .55 for literal comprehension, β= .73 for inferential comprehension), had stronger impact on inferential comprehension than did literal comprehension as seen in the magnitude of the regression coefficients.

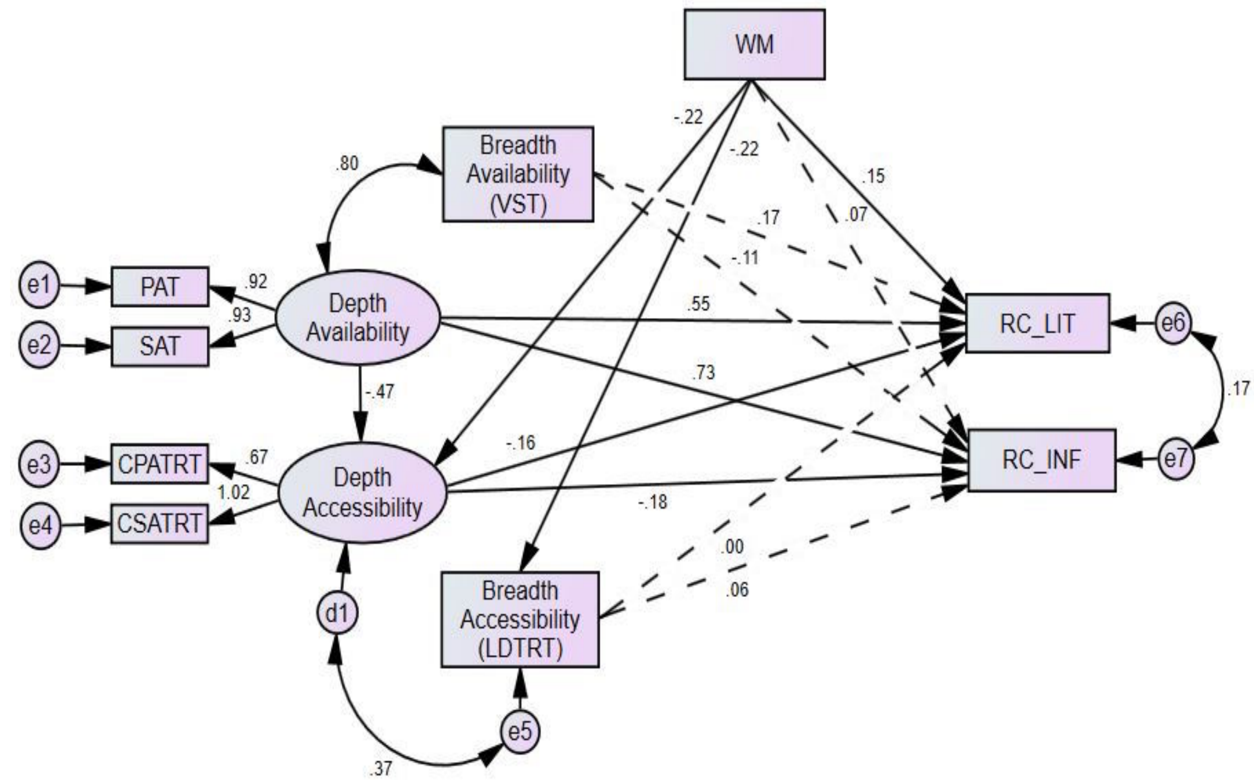

Figure 11: Standardized Structural Regression Weights among Working Memory Capacity, Availability and Accessibility of Word Knowledge, and Levels of Reading Comprehension.
Note. RC_LIT = literal comprehension; RC_INF = inferential comprehension; VST = vocabulary size test; PAT = paradigmatic associate test; SAT = syntagmatic associate test; LDTRT = response time for lexical decision test; CPATRT = response time for computerized-paradigmatic associate test; CSATRT = response time for computerized-syntagmatic associate test; WM = working memory capacity. Solid lines indicate statistically significant paths and dotted lines statistically not-significant paths.

Table 5: Standardized Regression Coefficients among Working Memory Capacity, Availability of Accessibility of Word Knowledge, and Levels of Reading Comprehension.

Paths			β	S.E.	C.R. (z)	*p*
Depth Accessibility	<---	Depth Availability	-.474	7.570	-5.986	<.001
Breadth Accessibility	<---	WM	-.221	24.803	-2.342	.019
Depth Accessibility	<---	WM	-.215	45.845	-2.673	.008
RC_LIT	<---	Breadth Availability	.169	.082	1.573	.116
RC_LIT	<---	Breadth Accessibility	.001	.001	.016	.987
RC_LIT	<---	Depth Availability	.554	.103	4.530	<.001
RC_LIT	<---	Depth Accessibility	-.163	.001	-2.045	.041
RC_LIT	<---	WM	.147	.307	2.408	.016
RC_INF	<---	Breadth Availability	-.114	.102	-.926	.354
RC_INF	<---	Breadth Accessibility	.064	.002	.863	.388
RC_INF	<---	Depth Availability	.734	.129	5.230	<.001
RC_INF	<---	Depth Accessibility	-.176	.001	-1.963	.050
RC_INF	<---	WM	-.074	.379	1.064	.287

CHAPTER FIVE

DISCUSSION

The primary focus of the present study was to examine the relationships of dimensions of word knowledge and their accessibility with reading comprehension in adolescent Korean EFL learners. Three research questions were addressed; (1) What is the construct structure of lexical competence represented by different dimensions of word knowledge and their respective accessibility? (2) What are the relationships of dimensions of word knowledge and their respective accessibility with reading comprehension? and (3) Do the contributions of different dimensions of word knowledge and their respective accessibility to reading comprehension vary as a function of the level of comprehension?

5.1. Construct structure of lexical competence

To answer the first question, a series of confirmatory factor analyses was conducted. CFAs identified with good fits two models that included all six measured competencies: a two-factor model of Availability (VST, PAT, and SAT) and Accessibility (LDTRT, CPAT, and CSAT), and a second-order model comprising two higher-order factors of Availability (Breadth Availability [VST] and a first-order factor of Depth Availability [PAT and SAT]) and Accessibility (Breadth Accessibility [LDTRT] and a first-order factor of Depth Accessibility [CPAT and CSAT]).

Taken together, the results of the CFAs (see Table 3) support for two claims. First, accessibility of word knowledge is a distinct lexical competence functioning separately, parallel to availability of word knowledge. In other words, it is one thing to have knowledge of individual word meanings and meaning relations, but it is another to be able to activate the knowledge and access it in an efficient manner. Second, within availability and accessibility, breadth and depth knowledge are conceptualized separately. That is, knowledge of individual word meanings (i.e., breadth availability) is related to but distinct from knowledge of meaning relations among words (i.e., depth availability). More importantly, the ability to efficiently access the knowledge of individual word meanings (i.e., breadth accessibility) is also a distinct, albeit related, concept from the ability to access efficiently the knowledge of meaning relations among words (i.e., depth accessibility).

The identification of distinction between availability of breadth and depth knowledge is not unexpected, and confirms previous conceptualizations about breadth/size *versus* depth distinction in the literature (Nagy & Anderson, 1984; Schmitt, 2014). Over and beyond the distinction identified at the knowledge (availability) level, the CFA results further identified a distinction between breadth and depth from the perspective of knowledge accessibility, which rarely has been demonstrated in previous research. As discussed in the Literature Review section, previously researchers conceptualized accessibility (e.g., fluency) either as a separate dimension of word knowledge parallel to vocabulary size and depth (Chapelle, 1994; Tannenbaum et al., 2006) or as a sub-component of depth of word knowledge (Schmitt, 2014). However, the poor fits of CFA models like Model 2 (see Table 3 and Figure 4), which had knowledge availability and accessibility measures loaded on a same factor, suggested it is inappropriate to conceptualize accessibility as a certain type of knowledge (e.g., lexical fluency as a type of vocabulary depth knowledge; see Schmitt, 2014). Rather, accessibility stands as a lexical skill indicating learners' quick and efficient activation of, or access to, knowledge they have, be it about individual word meanings or meaning relations among words in the mental lexicon, for meeting language use goals like reading comprehension; thus it is distinct from learners' knowledge itself.

Previously, some researchers seemed to tend to equate dimensions of word knowledge with dimensions of lexical competence, assuming that lexical competence, which embodies the ability of efficient access to knowledge as well as knowledge itself, can be achieved by gaining word knowledge (e.g., Meara, 1996; Henriksen, 1999; Haastrup & Henrikson, 2000). This appears to explain why the focus of previous research has been predominantly on what the dimensions of word knowledge are and how different knowledge dimensions – breadth or depth – can be developed or measured. It also seems to explain why previous research has paid particular attention to the importance of L2 learners' vocabulary knowledge for their general language proficiency development in general and reading proficiency development in particular (see discussion in the next part). Instructionally, it seems to agree also with the traditional emphasis of classroom teaching on ensuring that learners accumulate knowledge about a large number of words (and word relationships like synonyms and collocations). Consequently, attention has lacked regarding the construct of accessibility. Given that the purpose of leaning vocabulary is not merely to increase the number of words one knows, but also to improve competence in accessing those words efficiently to meet the goal of various language use situations like reading comprehension, the distinction between availability and accessibility, which was tested to be valid in this study, should be an additional conceptualization to enrich the oft-mentioned distinction between breadth/size and depth in the L2 literature.

5.2. Contribution of different lexical competences to reading comprehension

Based on the results of CFAs, structural equation modeling analysis was conducted to examine the unique contribution of the four different lexical competences (i.e., breadth and depth availability, and breadth and depth accessibility) to reading comprehension, which addressed the second research question. The results indicated that depth knowledge, which was true for both availability (PAT and SAT) and accessibility (CPAT and CSAT), was more important than breadth (VST and LDTRT), in predicting reading comprehension. Both availability and accessibility of depth knowledge made a unique and significant contribution to reading comprehension, whereas those of breadth knowledge did not play significant roles when the impact of depth and working memory was controlled for.

A significant and unique contribution of depth in this study is in line with those of some previous L2 studies adopting measures of word-associates format for assessing learners' vocabulary depth knowledge (Li & Kirby, 2005; Mehrpour et al., 2010; Qian, 1999, 2002; Zhang & Yang, 2016). The significant role of depth knowledge, including both availability and accessibility, is reasonable given that comprehension process not only involves learners' attention to a series of individual words in the text, but also entails identification of the meaning of the text as a connected whole to construct a coherent mental representation of text being read (i.e., situation model, Kintsch, 1988; Van Dijk & Kintsch, 1983). It would make sense that depth knowledge concerning meaning relations among words plays a more significant role than does breadth knowledge in building connections between propositions or discourses, which subsequently would improve comprehension. Efficient processing of meaning relations further would facilitate the comprehension process (Nagy & Scott, 2000).

As previously elaborated in Chapter 2, Perfetti highlights a lexical basis reading assuming knowledge of written word forms and meanings is central to reading comprehension. In a more recent work from Perfetti and his colleague (Perfetti & Stafura, 2014), they further looked into the role of word knowledge focusing on its interaction with text representation, which has implications on the critical role of depth of word knowledge in reading comprehension evidenced in the current study.

Drawing on research and theory of the modern study of reading comprehension, Perfetti and Stafura (2014) proposed the *Reading Systems Framework*, a general framework of reading systems placing word knowledge in the center of reading comprehension. They pointed out that within this framework, the processes of a word being integrated as part of the situation model established (i.e., word-to-text integration) serve as a key model for local comprehension processes. In other words, comprehension processes involve updating reader's mental representation of what is being described (i.e., situation model) by fitting a word or phrase into its context. If the integration occurs easily and effortlessly, comprehension takes place successfully, whereas if the word does not fit as easily into the context for some reasons, comprehension might be less effective, if not break down. One successful case of word-to-text integration the researchers demonstrated is when a word or phrase in one sentence is in an (implicit) co-referential relation with a word or phrase in a following sentence, which is what they called *the paraphrase effect*. They also found that skilled comprehenders were more capable of producing the paraphrase effect than less skilled comprehenders, who were described as showing 'sluggish' word-to-text integration (Perfetti et al., 2008).

Although Perfetti and Stafura did not directly discuss the depth aspect of word knowledge in terms of what types of word knowledge are responsible for the success of word-to-integration processes, it does seem that word-to-text integration, particularly the paraphrase effect, depends in part on knowledge of word meaning relations or network knowledge as well as knowledge of individual word meanings. The co-referential relation does not merely indicate the same thing or event referred across sentences by mentioning it in another way; it rather means that the paraphrase "fine-tunes the mental model by identifying a correlate or consequence of" (Perfetti & Stafura, 2014, p. 29) the event that was established in the previous sentence. In other words, knowledge of semantic relations among words can facilitate the word-to-text integration by allowing readers to link the sentences semantically and update the situation properly. This helps maintain coherence of the discourse and subsequently improves the ensuing comprehension process involved. Individual differences in the lexicon that supports the use of depth of word knowledge thus seem to lead to individual differences in comprehension skills found in this study.

To understand the bigger picture of the role of work knowledge (both breadth and depth) in comprehension, it might help return to Meara's representation of vocabulary size and organization. In this representation, breadth corresponds to the number of nodes in the network, while links between nodes represent depth. When adding a new node, that is, when one increases breadth, it either stands alone in the network or is linked to one or more existing ones as increasing depth. Of interest here is that breadth seems a "necessary but not sufficient" condition for depth. That is, depth can be established within the existing words in the network with or without a new word and, more importantly once it is established, the individual meanings of the words involved are necessarily there. In other words, depth knowledge not only concerns semantic word relations, but also necessarily involves individual word meaning, which is breadth knowledge, for lexical network to be established. Given that the comprehension process is facilitated by a dense, complex lexical network, the highlighted role of depth knowledge could be well explained.

While a significant and unique contribution of depth (availability as well as accessibility) to reading comprehension is reasonable, it was a surprise that in this study breadth knowledge (both availability and accessibility) was not found to be a significant predictor when depth and working memory capacity were accounted for. Considering that some previous studies have reported a significant contribution, not to mention sometimes a stronger one, of vocabulary breadth than vocabulary depth (Farvardin & Koosha, 2011; Huang, 2006), no significant unique contribution of breadth knowledge in this study was not usual. One speculation could be from the instruments for breadth knowledge. In this study, while availability (i.e., PAT, SAT) and accessibility (i.e., CPATRT, CSATRT) of depth knowledge were measured with two different instruments for each, only one instrument was used to measure each of availability (i.e., VST) and accessibility (i.e., LDTRT) of breadth knowledge. Employing a single indicator for measures of breadth knowledge would be psychometrically undesirable (Kline, 2010) in terms of lack of consideration of some error variances accounted for in the measurement model. Thus the results could have been different had there been a balance in the number of measures for breadth and depth knowledge allowing a range of error variances for both.

In addition, task reliabilities for measures of breadth knowledge (α= .69 for VST, α= .75 for LDTRT) were relatively low compared to those of depth knowledge (α= .83, α= .86 for PAT, SAT; α= .80, α= .85 for CPATRT, CSATRT), which might have impacted a less predictive role of breadth knowledge. In addition to having only a single indicator for breadth measures mentioned above, tasks used in this study for availability and accessibility of breadth knowledge might not successfully have represented individual differences in breadth knowledge.

Another interesting finding revealed in the SEM analysis was that within vocabulary depth measures, the unique contribution of accessibility of depth knowledge (CPAT and CSAT) appeared smaller than that of availability of depth knowledge (PAT and SAT). The result is in agreement with the findings of previous studies showing the important role of semantic processing skills in reading comprehension over and above semantic word knowledge (Cremer & Schoonen, 2014; Nation & Snowling, 2004; Oakhill et al., 2015). This suggests that a critical factor in reading comprehension is not only knowledge about meaning relations among words (i.e., availability of depth knowledge), but also is the ability to efficiently access and retrieve the knowledge.

The stronger impact of availability than of accessibility of depth knowledge might relate to participants' low level of language proficiency. As mentioned in the Method section, although students' language proficiency was not measured separately in this study, an English teacher in this school informed that students' overall English proficiency had been slightly below average in a nationwide English proficiency test. While they had learned English for more than eight years as a formal school education (from 3rd to 11th grades), their language proficiency might still be low given the limited language skills provided within the curriculum. This may suggest that knowledge the students possess perhaps overweighs their efficiency of retrieving it. If students with low-proficiency do not have a sufficient knowledge base available for comprehending a text, efficiency in accessing that limited knowledge would not be likely to help improve comprehension. As Nation and Snowling (1999) mentioned, semantic access depends on representations, organizations, and connections within the semantic system, which indicates that for speed or efficiency of semantic access to occur, one should have a certain level of semantic network knowledge available in one's lexicon. In this sense, for texts used in the comprehension tasks and for the group of participants present, word knowledge availability, rather than accessibility, is still the key.

5.3. Contribution of different lexical competences to different levels of comprehension

For the third research questions, the same SEM was performed, but with the two levels of reading comprehension modeled separately. In both literal and inferential comprehension, a pattern similar to that obtained in the previous SEM analysis was revealed. That is, availability and accessibility of depth knowledge seemed to have a stronger impact than those of breadth when controlling for other variables. Breadth knowledge, both availability and accessibility, did not show significant unique predictive power in both levels of comprehension. Accessibility of depth knowledge was also a unique predictor in both levels of comprehension. In other words, contributions of different dimensions of word knowledge and their accessibility seemed not to vary depending on levels of comprehension.

Research has demonstrated that the influence of vocabulary breadth and depth can vary based on levels of comprehension (Cain & Oakhill, 2014; Oakhill et al., 2015; Zhang & Yang, 2017), but that was not the case in this study. One explanation for the discrepant findings might be the nature of the items used to measure literal and inferential comprehension. In this study, passages in the Level 4 test in the Gates-MacGinitie Reading Test (Fourth Edition) were adopted. In developing the test, questions were classified as literal if they could be answered by choosing a restatement of something stated explicitly in the passage, whereas those that could not be answered by choosing a restatement were classed as inferential. However, a close examination of some questions in the passages adopted in this study suggested they did not seem to be appropriately classified. For example, in a passage about a young cactus sprouts, a question classified as inferential asked what makes the seed sprout. Students were supposed to choose '*wet weather'* as the answer out of distractors including *the sunlight, the paloverde tree,* and *the cool of the evening*. It seems that the clue of the answer *wet weather* came from a sentence in the passage '*After many dry days, a heavy rain falls on the desert'*. Although this sentence did not literally state 'wet weather' in the sentence, the level of inference needed to select the correct answer was somewhat low. Such items might be considered even tapping literal comprehension with a low level of interpretation needed. In other words, the inclusion of items like the example given above might have blurred the intended distinction between literal and inferential comprehension; and consequently, the patterns revealed in the SEM analysis were largely similar between the two levels of comprehension.

On the other hand, it is noted that the β value of availability of depth knowledge appeared stronger for inferential comprehension than for literal comprehension, indicating a possibly greater involvement of semantic network knowledge in inferential comprehension. To infer while reading, readers need to make sense of implicit information from information explicitly stated in the text by actively engaging their background knowledge. If rich semantic network knowledge helps activate not only word meanings but also relationships between the concepts represented in those words (Oakhill et al., 2015) where one's world knowledge is likely to be involved (e.g., Michigan, December, and cancellation of flights), it would make sense that participants in this study might have needed to depend more on their semantic network knowledge for inferential than for literal comprehension.

CHAPTER SIX

CONCLUSION

6.1. Summary of main findings

This study investigated the relationships of different dimensions of word knowledge and their respective accessibility with (different levels of) reading comprehension in Korean adolescent EFL learners. To achieve this research goal, it defined lexical competence by distinguishing between vocabulary breadth and depth knowledge as well as between knowledge availability and accessibility; it also differentiated between literal and inferential comprehension. A number of tasks were employed to measure participants' lexical competences. The following key research findings were revealed.

First, CFA analyses confirmed that knowledge of individual word meanings (breadth availability) and the ability to activate that knowledge efficiently (breadth accessibility), and knowledge of meaning relations among words (depth availability) and the ability to access that knowledge efficiently (depth accessibility) are distinctive constructs under the conceptualization of lexical competence.

Second, availability and accessibility of depth knowledge were the only predictors that made significant contributions to reading comprehension when the impacts of those of breadth knowledge and working memory were controlled for. Within vocabulary depth, availability played a more important role than did accessibility. Availability and accessibility of breadth knowledge made no unique contribution to reading comprehension over and beyond those of depth knowledge.

Third, the patterns of predictive roles of each lexical competence in reading comprehension did not seem substantially different between literal and inferential comprehension. Similarly, for reading comprehension without considering specific comprehension levels, a stronger impact was found of depth knowledge, both availability and accessibility, than of vocabulary breadth, and of availability of depth knowledge than of accessibility. On the other hand, the involvement of depth availability seemed greater in inferential comprehension than in literal comprehension.

6.2. Theoretical and pedagogical implications of the study

This study's findings have some important implications. Theoretically, this study contributes to current understandings about the lexical basis of L2 reading comprehension by unpacking the construct of lexical competence through the lens of distinctions between breadth and depth, as well as between knowledge availability and accessibility. The finding about a unique contribution of depth accessibility to reading comprehension, over and above other lexical competences, is particularly notable, given that it rarely was measured and examined in the context of reading comprehension in the previous literature. In addition, the finding about a greater involvement of depth availability in inferential than in literal comprehension, also enriches knowledge of the lexical basis of L2 reading comprehension.

This study also has important instructional implications for L2 vocabulary and reading classrooms. The unique contribution of depth knowledge to reading comprehension found in this study indicates that vocabulary teaching and learning should go beyond simply attending to individual word meanings. Vocabulary instruction should go toward the way in which students are supported to establish a semantic network in their lexicon and continue developing, expanding, and consolidating the lexicon as they learn new words.

One way of fostering development of semantic network for students is to have them learn words in a meaningful reading context such as in a text (August, Carlo, Dressler, & Snow, 2005;

Carlo et al., 2004), where words naturally connect with other words based on coherence and cohesion. This provides students with a rich context in which they can see what words are used in a particular context and how word meanings are linked both within a sentence or across sentences. This allows students to practice making connections among words, sentences, and propositions, which eventually help enhance their comprehension of the passage being read. Given that the major goal of learning words is to use them efficiently in natural language-use tasks including reading comprehension, teaching and learning words should not be isolated from such meaningful contexts.

In addition to development of (availability of) depth knowledge, how to address learners' efficient use of word knowledge in reading comprehension from the pedagogical perspective is certainly important to pay attention to, given the unique contribution of accessibility of depth knowledge to reading comprehension in the study. Because efficient processing consumes fewer cognitive resources, the more efficient knowledge activation occurs, the more attentional resources are left for other language purposes (Segalowitz, 2003), such as comprehension. This gives implications on instructions and assessments for vocabulary and reading. Learners' practices for quick and accurate vocabulary use and comprehension should be paid attention to, in addition to focus on building up semantic network knowledge.

The findings of the study also shed light on the assessment of vocabulary regarding how vocabulary tests should be designed and implemented. Given this study's identification of the importance of vocabulary depth for reading comprehension, including both availability and accessibility, vocabulary assessment should focus not only on word meanings, but also on meaning relations among the words. Test items asking simple form-meaning correspondence without being presented in context are unlikely to promote positive washback on teaching and learning, in the sense it does not help learners build rich semantic networks in their lexicon.

More importantly, vocabulary assessment needs to be implemented in an environment where learners' accessibility of word knowledge is brought into play. One practical way of this would be in the timed-test context. Read (2004) pointed out the limitations of confining the assessment of vocabulary to declarative knowledge, that is, knowledge of word meaning or meaning relations. Given that the goal of vocabulary learning is not merely to know about words, but to be able to access that knowledge efficiently and use it in a variety of contexts, vocabulary assessment needs to conceptualize procedural knowledge (or accessibility) as an important component, perhaps of even more importance than declarative knowledge.

6.3. Limitations and suggestions for future research

An important caveat in this study is the instrument that measured reading comprehension. As previously indicated, although I used a standardized test originally designed to have both literal and inferential questions, some of the questions classified as inferential seemed to require only low-level inference, which might have obscured the intended distinction between literal and inferential; consequently, the patterns of relationships of the lexical measures were largely similar between the two levels of comprehension. Considering the rather short length of each passage, perhaps not the best context in which to address inferential questions, more precise classification between literal and inferential questions should be made in designing items. Further, due to time constraints, I administered 30 items out of the original 48, including 15 for literal comprehension, the other 15 for inferential comprehension. Those 15 might not be large enough to represent each target level of comprehension in the SEM analysis for the purpose of comparison between the two levels of comprehension. Future research should consider more precise classifications between literal and inferential questions with a larger number of items for each level of comprehension.

Another limitation concerns the instruments for vocabulary breadth. While availability (i.e., PAT, SAT) and accessibility (i.e., CPATRT, CSATRT) of depth knowledge were measured with two different instruments for each, only one instrument was used to measure each of availability (i.e., VST) and accessibility (i.e., LDTRT) of vocabulary breadth. Typically in SEM analysis, the recommendation is to have at least two measures for a construct, although a single indicator is not impossible. Regarding a single indicator, Kline (2010) stated that if there is a single measure for a construct, it is especially critical for that single indicator to have good psychometric properties (Kline, 2010). In addition, as also indicated, reliabilities for the tasks for breadth knowledge were not as high as those for depth knowledge. Particularly for the accessibility task (i.e., LDTRT), the easiness of the test might have influenced low reliability. Given that this study did not adopt a full latent SEM analysis, where relationships among the variables can be examined with measurement error accounted for and the reliability of the measures might not impact the relationships, the relatively-low reliability for breadth knowledge with using a single indicator might have played a role in drawing the current results. Future studies could be psychometrically improved by adding more indicators and increasing item reliability.

As indicated earlier in the Method section, one of the challenges in designing this study was to ensure words presented in the tests (all the tests but VST) would need to be known or familiar to participants. Although much discretion was given to the process of selecting words to be used, the expected familiarity was not obtained in some tasks. For example, in CPAT, the mean accuracy was 63%, which was lower than expected. Considering RT scores in this study were calculated only with the items correctly responded to, high scores in item accuracy were needed. One possible solution for this in future research would be to conduct a pilot study to validate items, and with a sizable number of L2 participants.

Lastly, it was noted that a greater involvement of vocabulary depth availability than accessibility might come from the low language proficiency of students participating in this study. To explore how availability and accessibility of word knowledge interact with each other and might be affected by language-proficiency level or participants' developmental stage, additional data from longitudinal research are necessary. Longitudinal data also would allow causal claims to be made from the structural models tested, which could not be gained from the current study.

APPENDICES

Appendix A Vocabulary Size Test (VST)

Student ID _______________ Name_________________

Circle the correct meaning of the given word from the four options below. (Directions were presented in Korean.)

Example) peel: Shall I **peel** it?
- a. 물에 담그다
- (b.) 껍질을 벗기다
- c. 희게 만들다
- d. 잘게 썰다

1. maintain: Can they **maintain** it?
   - a. 유지하다
   - b. 확대시키다
   - c. 더 나은 것을 얻다
   - d. 얻다

2. stone: He sat on a **stone**
   - a. 돌
   - b. 의자
   - c. 양탄자, 깔개
   - d. 나무

3. upset: I am **upset**.
   - a. 피곤한
   - b. 유명한
   - c. 부유한
   - d. 불행한

4. drawer: The **drawer** was empty.
   - a. 서랍
   - b. 차고
   - c. 냉장고
   - d. 동물의 집

5. patience: He has no **patience**.
   - a. 인내심
   - b. 여유시간
   - c. 믿음
   - d. 정직함

6. nil: His mark for that question was **nil**.
   - a. 매우 나쁜
   - b. 영점인
   - c. 매우 좋은
   - d. 중간인

7. pub: They went to the **pub**.
   - a. 술집
   - b. 도박장
   - c. 쇼핑센터
   - d. 수영장

8. circle: Make a **circle**.
   a. 스케치
   b. 빈 공간
   c. 원(형)
   d. 큰 구멍

9. microphone: Please use the **microphone**.
   a. 전자레인지
   b. 마이크
   c. 현미경
   d. 휴대폰

10. pro: He's a **pro**.
   a. 사립탐정
   b. 미련한 사람
   c. 기자
   d. 직업으로 하는 운동선수

11. soldier: He is a **soldier**.
   a. 사업가
   b. 학생
   c. 금속 공예사
   d. 군인

12. restore: It has been **restored**.
   a. 다시 말해졌다
   b. 다른 사람에게 주어졌다
   c. 더 낮은 가격이 주어졌다
   d. 다시 새것처럼 만들어졌다

13. jug: He was holding a **jug**.
   a. 주전자
   b. 격식 없는 대화
   c. (챙이 달린) 모자
   d. 폭탄

14. scrub: He is **scrubbing** it.
   a. 가는 선을 새기로 있다
   b. 고치고 있다
   c. 깨끗하게 하려고 세게 문지르고 있다
   d. 간단한 그림을 그리고 있다

15. dinosaur: The children were pretending to be **dinosaurs**.
   a. 해적
   b. 요정
   c. 용
   d. 공룡

16. strap: He broke the **strap**.
   a. 약속
   b. 뚜껑
   c. 접시
   d. 끈

17. pave: It was **paved**.
   a. 막혔다
   b. 나눠져 있었다
   c. 금테가 둘러있었다
   d. (도로가) 포장되어 있었다

18. dash: They **dashed** over it.
a. 돌진하였다
b. 천천히 움직였다
c. 싸웠다
d. 빨리 보았다

19. rove: He couldn't stop **roving**.
a. 술 마시는 것
b. 배회하는 것
c. 허밍 하는것
d. 열심히 일하는 것

20. lonesome: He felt **lonesome**.
a. 은혜를 모르는
b. 피곤한
c. 외로운
d. 힘이 넘치는

21. compound: They made a new **compound**.
a. 동의
b. 혼합물
c. 사업하는 팀
d. 추측

22. latter: I agree with the **latter**.
a. 교회에서 온 남자
b. 주어진 이유
c. 마지막 것 (후자)
d. 대답

23. candid: Please be **candid**.
a. 조심하다
b. 동정하다
c. 공평하다
d. 솔직하다

24. tummy: Look at my **tummy**.
a. 두건
b. (사람의) 배
c. 털이 있는 작은 동물
d. 엄지 손가락

25. quiz: We made a **quiz**.
a. 화살 통
b. 심각한 실수
c. 간단한 시험
d. 새장

26. input: We need more **input**.
a. (정보, 자원) 투입
b. 노동자
c. 나무에 난 구멍을 메우는 것
d. 돈

27. crab: Do you like **crabs**?
a. 게
b. 팬케이크
c. 딱딱한 칼라 (깃)
d. 귀뚜라미

28. vocabulary: You will need more **vocabulary**.
   a. 단어
   b. 규칙
   c. 돈
   d. 총

29. remedy: We found a good **remedy**.
   a. 해결책
   b. 대중식당
   c. 조리법
   d. 숫자체계

30. allege: They **alleged** it.
   a. 우겨댔다
   b. 남의 생각을 도용했다
   c. 증거를 댔다
   d. 주장에 강하게 반대했다.

Appendix B Paradigmatic Associates Test (PAT)

This is a test of how well you know the meaning of adjectives that are commonly used in English. Choose two words that you think are relevant to the target word. The words you choose may help to explain the meaning of the given word. (i.e., synonym) For example, in the box below, "**sudden**" means "happening quickly and unexpectedly", so the correct answers are "quick" and "surprising". (Directions were presented in Korean)

Sudden			
① beautiful	✓② quick	✓③ surprising	④ thirsty

==

1. bright

① sunny	② famous	③ expensive	④ shining

2. calm

① small	② quiet	③ peaceful	④ tired

3. fresh

① expected	② cool	③ same	④ new

4. empty

① free	② heavy	③ blank	④ boring

5. common

① full	② light	③ familiar	④ typical

6. final

① last	② healthy	③ end	④ laughing

7. helpful

① useful	② sudden	③ often	④ important

8. recent

① current	② small	③ modern	④ glad

9. busy

① active	② sweet	③ soft	④ hardworking

10. angry

① mad	② simple	③ annoyed	④ round

11. tight

① narrow	② rough	③ uncomfortable	④ wet

12. dense

① straight	② thick	③ close	④ right

13. different

① various	② unlike	③ proud	④ real

14. difficult

① local	② hard	③ exact	④ tough

15. formal

① fast	② loud	③ official	④ serious

16. gentle

① possible	② kind	③ mild	④ high

17. painful

① sore	② shy	③ terrible	④ valuable

18. favorite

① popular	② early	③ well-liked	④ worried

19. relaxed

① easy	② normal	③ poor	④ informal

20. public

① delicious	② general	③ open	④ windy

Appendix C Syntagmatic Associates Test (SAT)

This is a test of how well you know about words that generally go together (i.e., collocation) in English. Choose two words that may come after the given word in a phrase or a sentence. For example, in the box below, we don't normally say " a sudden doctor" or "a sudden school", but we often say "a sudden change" or "a sudden noise", so "change" and "noise" are the correct answers. (Directions were presented in Korean)

Sudden			
✓① change	② doctor	✓③ noise	④ school

==

1. effective

① way	② face	③ taste	④ system

2. mental

① illness	② brand	③ health	④ chance

3. public

① frost	② school	③ opinion	④ throat

4. dry

① count	② mouth	③ season	④ height

5. complex

① bath	② cook	③ patterns	④ problem

6. broad

① night	② river	③ shoulders	④ mother

7. deep

① breath	② future	③ water	④ weather

8. necessary

① time	② money	③ nobody	④ steel

9. dull

① color	② knife	③ vegetable	④ left

10. direct

① century	② flight	③ heat	④ salad

11. violent

① crime	② rice	③ total	④ attack

12. boring

① oil	② limit	③ job	④ story

13. brown

① hair	② shopping	③ sugar	④ cost

14. official

① stranger	② document	③ language	④ adult

15. major

① decision	② role	③ tomorrow	④ sky

16. soft

① touch	② maximum	③ voice	④ corner

17. automatic

① heaven	② machine	③ skin	④ process

18. serious

① issue	② situation	③ wind	④ tissue

19. familiar

① husband	② stop	③ topic	④ place

20. voluntary

① basis	② service	③ sea	④ arrow

Appendix D Lexical Decision Test (LDT)

- Items were randomly ordered and presented in a computer screen with a prompt "Does this word exist in English?" in Korean
- Students were asked to choose either yes or no as quickly as possible.

Existing words: 25 items

1. any
2. low
3. off
4. say
5. come
6. draw
7. drop
8. form
9. rain
10. wash
11. aware
12. class
13. clear
14. clock
15. ready
16. shape
17. south
18. state
19. thing
20. matter
21. minute
22. record
23. company
24. explain
25. absolute

Non-words: 10 items

1. olk
2. swal
3. peem
4. atus
5. knilc
6. boash
7. tharm
8. voped
9. glarms
10. dwurlde

Appendix E Computer-based Paradigmatic Associates Test (CPAT)

- Target words and distractors were randomly selected and presented in a computer screen with a prompt “ Are they synonymous or semantically related?” in Korean
- Students were asked to choose either yes or no as quickly as possible.

Paradigmatic associates: 25 items

1. angry : annoyed
2. small : little
3. different : various
4. quick : active
5. rich : full
6. slow : lazy
7. exact : perfect
8. short : brief
9. sad : sorry
10. safe : protected
11. final : closing
12. strong : powerful
13. strange : curious
14. heavy : huge
15. important : great
16. shy : nervous
17. difficult : tough
18. usual : familiar
19. wet : rainy
20. fair : equal
21. cheap : poor
22. young : new
23. normal : natural
24. easy : clear
25. lovely : beautiful

Distractors (random relations): 10 items

1. deep : happy
2. dirty : minor
3. early : delicious
4. empty : serious
5. fast : harm
6. general : old
7. terrible : brave
8. warm : famous
9. favorite : rare
10. close : hard

Appendix F Computer-based Syntagmatic Associates Test (CSAT)

- Target words and distractors were randomly selected and presented in a computer screen with a prompt "Do the two words go together in English in general?" in Korean
- Students were asked to choose either yes or no as quickly as possible.

Syntagmatic associates: 25 items

1. nice weather
2. stupid question
3. hard work
4. healthy food
5. bad news
6. simple task
7. late night
8. serious problem
9. thick hair
10. large number
11. pretty woman
12. cold water
13. wrong person
14. special need
15. clean house
16. clear sound
17. busy life
18. open market
19. local time
20. middle class
21. near future
22. same place
23. dark room
24. polite way
25. native language

Distractors (random relations): 10 items

1. thirsty hand
2. front family
3. delicious hole
4. familiar health
5. blank leave
6. plain police
7. previous adult
8. wise park
9. basic west
10. sharp ocean

Appendix G Working Memory Capacity Stimuli List

Table 6: Forward Digit Span Task (FDST).

		Digit Sequence Presented as Stimuli	Digit Sequence Presented as an Answer	Correct Answer
Practice item #1		3,1	13	Yes
Practice item #2		5,2,8	529	No
2 digits	Item #1	3,2	23	No
	Item #2	8,7	87	Yes
	Item #3	4,0	50	NO
3 digits	Item #4	4,7,3	473	Yes
	Item #5	2,8,5	925	No
	Item #6	6,1,0	610	Yes
4 digits	Item #7	1,6,3,0	1630	Yes
	Item #8	3,0,6,7	3067	Yes
	Item #9	8,6,7,3	8346	NO
5 digits	Item #10	6,3,8,7,0	63870	Yes
	Item #11	8,4,2,5,6	84256	Yes
	Item #12	7,6,3,9,8	76398	Yes

Table 7: Backward Digit Span Task (BDST).

		Digit Sequence Presented as Stimuli	Digit Sequence Presented as an Answer	Correct Answer
Practice item #1		5,6	65	Yes
Practice item #2		9,4,6	652	No
2 digits	Item #1	2,6	62	Yes
	Item #2	8,7	75	No
	Item #3	3,2	23	Yes
3 digits	Item #4	7,9,2	297	Yes
	Item #5	6,4,1	146	Yes
	Item #6	3,4,6	346	No
4 digits	Item #7	8,5,6,1	1658	Yes
	Item #8	1,9,5,2	2591	Yes
	Item #9	5,2,0,6	6205	NO
5 digits	Item #10	6,5,9,0,9	90956	Yes
	Item #11	2,9,0,8,7	78915	No
	Item #12	7,6,5,2,1	12567	Yes

Table 8: Operation Word Span Task (OWST).

	Operation	To-be-rememberd Word (English)	Word Sequence Presented (English)	Correct Answer
Practice item #1	(2*3)+2 = 8	과일 (FRUIT)	과일 주전자 (FRUIT JUG)	Yes
	(6/3)-1 = 3	주전자 (JUG)		
Practice item #2	(4*1)+5 = 9	컵 (CUP)	컵 사과 (CUP APPLE)	No
	(9/3)-1 = 3	딸기 (STRAWBERRY)		
Item #1	(8/4)-1 = 1	중국 (CHINA)	중국 엄마 (CHINA MOM)	Yes
	(9/9)+4 = 5	엄마 (MOM)		
Item #2	(4/1)-3 = 5	목 (NECK)	가슴 흰색 (CHEST WHITE)	No
	(7*2)-1 = 13	흰색 (WHITE)		
Item #3	(8*1)-4 = 4	고기 (MEAT)	고기 마늘 (MEAT GARLIC)	Yes
	(2*5)-3 = 2	마늘 (GARLIC)		
Item #4	(1*5)+6 = 11	호랑이 (TIGER)	호랑이 열쇠 진달래 (TIGER KEY FLOWER)	Yes
	(4/2)-1 = 1	열쇠 (KEY)		
	(3*1)+5 = 1	진달래 (FLOWER)		
Item #5	(3*2)-4 = 9	테이블 (TABLE)	상자 정문 생신 (BOX GATE BIRTHDAY)	No
	(6/3)+5 = 7	정문 (GATE)		
	(2*4)-3 = 5	생신 (BIRTHDAY)		
Item #6	(8/4)+3 = 5	거실 (LIVINGROOM)	거실 화장품 소설가 (LIVININGROOM LOTION NOVELIST)	Yes
	(10/5)-1 = 5	화장품 (LOTION)		
	(8/8)+4 = 5	소설가 (NOVELIST)		
Item #7	(9/3)-3 = 0	가슴 (CHEST)	머리 감기 젓가락 감자 (HEAD COLD CHOPSTICS POTATO)	No
	(9/3)+5 = 12	감기 (COLD)		
	(1*4)-1 = 3	젓가락 (CHOPSTICS)		
	(6/2)+4 = 7	당근 (CARROT)		
Item #8	(2*3)-5 = 1	음료수 (DRINK)	음료수 넥타이 왼쪽 골목 (DRINK TIE LEFT ALLEY)	Yes
	(3/1)+4 = 10	넥타이 (TIE)		
	(2*8)+5 = 21	왼쪽 (LEFT)		
	(3*7)+4 = 25	골목 (ALLEY)		

Table 8 (cont'd).

Item #9	(2/2)+4 = 5 (6*4)-4 = 12 (3*4)+5 = 17 (6*1)+1 = 20	빵 (BREAD) 형제 (BROTHER) 요일 (DAY) 결혼식 (WEDDING)	빵 형제 가수 결혼식 (BREAD BROTHER MUSICIAN WEDDING)	No
Item #10	(4/2)+5 = 7 (8/4)+8 = 5 (3*1)+8 = 11 (6/3)+5 = 12 (6*2)+8 = 20	송이 (BUNCH) 미술 (ART) 한글 (KOERAN) 고양이 (CAT) 가을 (FALL)	송이 미술 한글 고양이 가을 (BUNCH ART KOREAN CAT FALL)	Yes
Item #11	(4*5)-6 = 7 (3/1)+4 = 7 (4*2)-2 = 2 (2*9)+3 = 21 (5*3)+3 = 13	수돗물 (TAP(WATER)) 가족 (FAMILY) 아들 (SON) 독서 (READING) 숙제 (HOMEWORK)	수돗물 가족 아들 독서 숙제 (TAP FAMILY SON READING HOMEWORK)	Yes
Item #12	(9/3)+4 = 7 (6/2)+2 = 5 (2/1)-1 = 4 (4*3)+2 = 14 (9/3)-1 = 2	잔 (GLASS) 어린이 (CHILD) 음악가 (MUSICIAN) 떡 (RICECAKE) 목소리 (VOICE)	잔 어린이 음악가 떡 목소리 (GLASS CHILD MUSICIAN RICECAKE VOICE)	Yes

Appendix H Grammaticality Judgment Test (GJT)

Please indicate whether each sentence is grammatically correct by checking Yes or No.

1.	Every Friday my dad washes his car.	Yes	No
2.	The boy lost two teeth in the fight.	Yes	No
3.	The woman asked the policeman a question.	Yes	No
4.	The boy carrots feeds the rabbits.	Yes	No
5.	The lady clean her house every morning.	Yes	No
6.	The girls want watching TV.	Yes	No
7.	The little boy is speaking to a policeman.	Yes	No
8.	Does Sumi use her computer?	Yes	No
9.	Jane is wear the dress I gave her.	Yes	No
10.	Mike wrote the letter but didn't send it.	Yes	No
11.	Tom is reading a book in his room.	Yes	No
12.	The lady fill a box with apples.	Yes	No
13.	When do they leave for Korea?	Yes	No
14.	Two mouses ran into the house.	Yes	No
15.	Can John ride a bike?	Yes	No
16.	The boy is helping the man build a house.	Yes	No
17.	The man climbed the stair up carefully.	Yes	No
18.	Mary looked at the flowers but didn't buy them.	Yes	No
19.	Did Sumi stayed at home last night?	Yes	No
20.	The children with the dog play.	Yes	No
21.	What they sell at the store?	Yes	No
22.	They carried a long conversation on.	Yes	No
23.	The students went to the movies.	Yes	No
24.	Is Sally waiting in the car?	Yes	No
25.	The girls enjoy watching TV.	Yes	No
26.	Last night the old lady died in her sleep.	Yes	No

Appendix I Word Familiarity Checklist

Word Check List

Please indicate whether you know the word by checking Yes or No.
주어진 단어가 아는 단어이면 "예", 모르는 단어이면 "아니오" 에 동그라미(O) 하세요.

1. strange	예 (Yes)	아니오 (No)	21. basis	예	아니오
2. various	예	아니오	22. adult	예	아니오
3. sharp	예	아니오	23. familiar	예	아니오
4. normal	예	아니오	24. park	예	아니오
5. public	예	아니오	25. clear	예	아니오
6. salad	예	아니오	26. terrible	예	아니오
7. typical	예	아니오	27. wrong	예	아니오
8. vegetable	예	아니오	28. soft	예	아니오
9. sound	예	아니오	29. attack	예	아니오
10. brief	예	아니오	30. shy	예	아니오
11. annoyed	예	아니오	31. close	예	아니오
12. nervous	예	아니오	32. touch	예	아니오
13. previous	예	아니오	33. problem	예	아니오
14. mild	예	아니오	34. deep	예	아니오
15. easy	예	아니오	35. weather	예	아니오
16. news	예	아니오	36. glad	예	아니오
17. famous	예	아니오	37. sea	예	아니오
18. high	예	아니오	38. full	예	아니오
19. empty	예	아니오	39. money	예	아니오
20. fast	예	아니오	40. time	예	아니오

Appendix J Summary of Measures and Instruments

Table 9: Measures and Instruments.

Measures	Instruments	# of items
Word knowledge		
Availability		
Vocabulary size	Vocabulary Size Test (VST)	30
Paradigmatic association (synonym)	Paradigmatic Associates Test (PAT)	20
Syntagmatic association (collocation)	Syntagmatic Associates Test (SAT)	20
Accessibility		
Word recognition	Lexical Decision Test (LDT)	35
Paradigmatic association (synonym)	Computerized-paradigmatic Associates Test (CPAT)	35
Syntagmatic association (collocation)	Computerized-syntagmatic Associates Test (CSAT)	35
Working memory capacity		
Storage capacity	Visual Digit Span Task (DST)	24
Storage + processing capacity	Operation Word Span Task (OWST)	12
Grammatical knowledge	Grammaticality Judgment Test (GJT)	26
Reading comprehension	Gates-MacGinitie Reading Test-Level 4 (GMRT)	30
	Literal comprehension questions	15
	Inferential comprehension questions	15

REFERENCES

REFERENCES

Akbarian, I. (2010). The relationship between vocabulary size and depth for ESP/EAP learners. *System*, *38*(3), 391–401. https://doi.org/10.1016/j.system.2010.06.013

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

Arbuckle, J. L. (2011). *IBM SPSS Amos 20 user's guide*. Armonk, NY: IBM Corporation.

August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research & Practice*, *20*(1), 50–57.

Batty, A. (2012). Identifying dimensions of vocabulary knowledge in the word associates test. *Vocabulary Learning and Instruction*, *1*(1), 70–77.

Beck, I. L., & McKeown, M. G. (2007). Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *The Elementary School Journal*, *107*(3), 251–271.

Beck, I. L., Perfetti, C. A., & McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology*, *74*(4), 506–521. https://doi.org/http://dx.doi.org.proxy2.cl.msu.edu/10.1037/0022-0663.74.4.506

Bialystok, E. (1981). The role of conscious strategies in second language proficiency. *The Modern Language Journal*, *65*(1), 24–35. https://doi.org/10.2307/326204

Cain, K. (2005). Children's reading comprehension difficulties. In C. Wood & V. Connelly (Eds.), *Contemporary perspectives on reading and spelling* (pp. 59–75). New York, NY: Routledge.

Cain, K., & Oakhill, J. (2014). Reading comprehension and vocabulary: Is vocabulary more important for some aspects of comprehension? *L'Année Psychologique*, *114*, 647–662.

Cain, K., Oakhill, J., & Lemmon, K. (2004). Individual differences in the inference of word meanings from context: The influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology*, *96*(4), 671–681.

Carlo, M. S., August, D., McLaughlin, B., Snow, C. E., Dressler, C., Lippman, D. N., … White, C. E. (2004). Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classrooms. *Reading Research Quarterly*, *39*(2), 188–215.

Chapelle, C. A. (1994). Are C-tests valid measures for L2 vocabulary research? *Second Language Research*, *10*(2), 157–187. https://doi.org/10.1177/026765839401000203

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*(5), 769–786.

Cremer, M., & Schoonen, R. (2013). The role of accessibility of semantic word knowledge in monolingual and bilingual fifth-grade reading. *Applied Psycholinguistics*, *34*(6), 1195–1217.

Daller, H., Milton, J., & Treffers-Daller, J. (2007). Editors' introduction: Conventions, terminology and an overview of the book. In H. Daller, J. Milton, & J. Treffers-Daller (Eds.), *Modelling and assessing vocabulary knowledge* (pp. 1–32). Cambridge, UK: Cambridge University Press.

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, *22*(4), 499–533.

Droop, M., & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly*, *38*(1), 78–103. https://doi.org/http://dx.doi.org.proxy2.cl.msu.edu/10.1598/RRQ.38.1.4

Duke, N. K., & Carlisle, J. (2011). The development of comprehension. In M. L. Kamil, P. D. Pearson, E. B. Moje, & P. P. Afflerbach (Eds.), *Handbook of Reading Research, Vol. IV* (pp. 199–228). New York, NY: Routledge.

Eason, S. H., Goldberg, L. F., Young, K. M., Geist, M. C., & Cutting, L. E. (2012). Reader-text interactions: How differential text and question types influence cognitive skills needed for reading comprehension. *Journal of Educational Psychology*, *104*(3), 515–528. https://doi.org/10.1037/a0027182

Farvardin, M. T., & Koosha, M. (2011). The role of vocabulary knowledge in Iranian EFL students' reading comprehension performance: Breadth or depth? *Theory and Practice in Language Studies*, *1*(11), 1575–1580.

Good, R. H., Gruba, J., & Kaminski, R. A. (2002). Best practices in using dynamic indicators of basic early literacy skills (DIBELS) in an outcomes-driven model. Retrieved from http://psycnet.apa.org/psycinfo/2006-03715-045

Goulden, R., Nation, P., & Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics*, *11*(4), 341–363.

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York: Cambridge University Press.

Greidanus, T., Bogaards, P., van der Linden, E., Nienhuis, L., & de Wolf, T. (2004). The construction and validation of a deep word knowledge test for advanced learners of

French. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a Second Language: Selection, acquisition, and testing* (pp. 191–208). Amsterdam: John Benjamins Publishing.

Haastrup, K., & Henriksen, B. (2000). Vocabulary acquisition: Acquiring depth of knowledge through network building. *International Journal of Applied Linguistics*, *10*(2), 221–240.

Hamilton, S. T., Freed, E. M., & Long, D. L. (2013). Modeling reader and text interactions during narrative comprehension: A test of the Lexical Quality Hypothesis. *Discourse Processes*, *50*(2), 139–163. https://doi.org/10.1080/0163853X.2012.742001

Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, *14*(1), 25–38. https://doi.org/10.1017/S0272263100010457

Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition*, *21*(2), 303–317.

Henriksen, B. (2008). Declarative lexical knowledge. In *Vocabulary and writing in a first and second language* (pp. 22–66). Springer. Retrieved from http://link.springer.com/chapter/10.1057/9780230593404_2

Henriksen, B., Albrechtsen, D., & Haastrup, K. (2004). The relationship between vocabulary size and reading comprehension in the L2. In D. Albrechtsen, K. Haastrup, & B. Henriksen (Eds.), *Angles on the English-speaking world: Writing and vocabulary in foreign language acquisition* (Vol. 4, pp. 129–140).

Hirsh, D., Nation, P., & others. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language*, *8*, 689–689.

Horiba, Y. (2012). Word knowledge and its relation to text Comprehension: A comparative study of Chinese- and Korean-speaking L2 learners and L1 speakers of Japanese. *The Modern Language Journal*, *96*(1), 108–121. https://doi.org/10.1111/j.1540-4781.2012.01280.x

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

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

Huang, H.-F. (2006). *Breadth and depth of English vocabulary knowledge: Which really matters in the academic reading performance of Chinese university students?* (M.A.). McGill University (Canada), Canada. Retrieved from http://search.proquest.com.proxy1.cl.msu.edu/docview/304931830/abstract/24446A74D1084C0DPQ/1

Ikeno, O. (2006). L1 and L2 working memory: An investigation into the domain specificity and processing efficiency issues. *Bulletin of the Faculty of Education at Ehime University*, *53*, 113–121.

Jeon, E. H., & Yamashita, J. (2014). L2 Reading comprehension and its correlates: A meta-analysis. *Language Learning*, *64*(1), 160–212. https://doi.org/10.1111/lang.12034

Jeon, M. (2009). Globalization and native English speakers in English Programme in Korea (EPIK). *Language, Culture and Curriculum*, *22*(3), 231–243. https://doi.org/10.1080/07908310903388933

Jung, J. (2009). Second language reading and the role of grammar. *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics*, *9*(2), 29–48.

Kieffer, M. J., & Lesaux, N. K. (2012). Knowledge of words, knowledge about words: Dimensions of vocabulary in first and second language learners in sixth grade. *Reading and Writing*, *25*(2), 347–373. https://doi.org/http://dx.doi.org.proxy2.cl.msu.edu.proxy1.cl.msu.edu/10.1007/s11145-010-9272-9

Kim, Y.-S., Wagner, R. K., & Foster, E. (2011). Relations among oral reading fluency, silent reading fluency, and reading comprehension: A latent variable study of first-grade readers. *Scientific Studies of Reading*, *15*(4), 338–362. https://doi.org/10.1080/10888438.2010.493964

Kim, Y.-S., Wagner, R. K., & Lopez, D. (2012). Developmental relations between reading fluency and reading comprehension: A longitudinal study from Grade 1 to Grade 2. *Journal of Experimental Child Psychology*, *113*(1), 93–111. https://doi.org/10.1016/j.jecp.2012.03.002

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163–182. https://doi.org/http://dx.doi.org.proxy2.cl.msu.edu/10.1037/0033-295X.95.2.163

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford press.

Koda, K. (1996). L2 word recognition research: A critical review. *The Modern Language Journal*, *80*(4), 450–460. https://doi.org/10.2307/329725

Laufer, B. (1992a). How much lexis is necessary for reading comprehension? In *Vocabulary and applied linguistics* (pp. 126–132). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-1-349-12396-4_12

Laufer, B. (1992b). Reading in a foreign language: how does L2 lexical knowledge interact with the reader's general academic ability'. *Journal of Research in Reading*, *15*(2), 95–103.

Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, *22*(1), 15–30.

Leeser, M. J. (2007). Learner-based factors in L2 reading comprehension and processing grammatical form: Topic familiarity and working memory. *Language Learning*, *57*(2), 229–270. https://doi.org/10.1111/j.1467-9922.2007.00408.x

Li, M., & Kirby, J. R. (2015). The effects of vocabulary breadth and depth on English reading. *Applied Linguistics*, *36*(5), 611–634. https://doi.org/10.1093/applin/amu007

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2000). *Gates-MacGinitie Reading Tests* (4th ed.). Itasca, IL: Riverside Publishing Company.

Mancilla-Martinez, J., & Lesaux, N. K. (2010). Predictors of reading comprehension for struggling readers: The case of Spanish-speaking language minority learners. *Journal of Educational Psychology*, *102*(3), 701–711.

Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjaer, & J. Williams (Eds.), *Performance and competence in second language acquisition* (pp. 35–53). Cambridge, UK: Cambridge University Press.

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

Mehrpour, S., Razmjoo, S. A., & Kian, P. (2011). The relationship between depth and breadth of vocabulary knowledge and reading comprehension among Iranian EFL learners. *Journal of English Language Teaching and Learning*, *53*(222), 97–127.

Mezynski, K. (1983). Issues concerning the acquisition of knowledge: Effects of vocabulary training on reading comprehension. *Review of Educational Research*, *53*(2), 253–279.

Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: evidence from a longitudinal study. *Developmental Psychology*, *40*(5), 665.

Na, L., & Nation, I. S. P. (1985). Factors affecting guessing vocabulary in context. *RELC Journal*, *16*(1), 33–42.

Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly*, 304–330.

Nagy, W. E., & Scott, J. A. (2000). Vocabulary processes. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. III). Mahwah, NJ: Erlbaum.

Nassaji, H. (2003). Higher-level and lower-level text processing skills in advanced ESL reading comprehension. *The Modern Language Journal*, *87*(2), 261–276.

Nation, I. S. P. (1990). *Teaching and learning vocabulary.* New York, NY: Newbury House.

Nation, I. S. P. (1993). The bilingual lexicon. In R. Schreuder & B. Weltens (Eds.), *Vocabulary size, growth, and use* (pp. 115–134). Amsterdam: John Benjamins.

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

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review / La Revue Canadienne Des Langues Vivantes*, *63*(1), 59–81. https://doi.org/10.1353/cml.2006.0049

Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, *31*(7), 9–13.

Nation, K., & Norbury, C. F. (2005). Why reading comprehension fails: Insights from developmental disorders. *Topics in Language Disorders*, *25*(1), 21–32.

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: broader language skills contribute to the development of reading. *Journal of Research in Reading*, *27*(4), 342–356. https://doi.org/10.1111/j.1467-9817.2004.00238.x

Oakhill, J., Cain, K., & McCarthy, D. (2015). Inference processing in children: The contributions of depth and breath of vocabulary knowledge. In E. J. O'Brien, A. E. Cook, & R. F. Lorch (Eds.), *Inferences during reading* (pp. 140–159). Cambridge, UK: Cambridge University Press.

Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology*, *98*(3), 554–566.

Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 174–200). Cambridge, UK: Cambridge University Press.

Perception Research Systems. (2007). *Paradigm Stimulus Presentation*. Retrieved from http://www.paradigmexperiments.com

Perfetti, C. A. (1985). *Reading ability.* Oxford University Press. Retrieved from http://doi.apa.org/psycinfo/1985-97290-000

Perfetti, C. A. (1999). Comprehending written language: A blueprint of the reader. In C. M. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 167–208). Oxford, UK: Oxford University Press.

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*(4), 357–383.

Perfetti, C. A., & Hart, L. (2001). The lexical basis of comprehension skill. In D. S. Gorfein (Ed.), *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity.* (pp. 67–86). Washington, DC: American Psychological Association.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (Vol. 11, p. 189-213). Amsterdam: John Benjamins.

Price, K. W., Meisinger, E. B., Louwerse, M. M., & D'Mello, S. (2016). The contributions of oral and silent reading fluency to reading comprehension. *Reading Psychology*, *37*(2), 167–201. https://doi.org/http://dx.doi.org.proxy2.cl.msu.edu/10.1080/02702711.2015.1025118

Proctor, C., Carlo, M., August, D., & Snow, C. (2005). Native Spanish-speaking children reading in English: Toward a model of comprehension. *Journal of Educational Psychology*, *97*(2), 246–256. https://doi.org/10.1037/0022-0663.97.2.246

Qian, D. D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review/ La Revue Canadienne Des Langues Vivantes*, *56*(2), 282–308. https://doi.org/10.3138/cmlr.56.2.282

Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*, *52*(3), 513–536. https://doi.org/10.1111/1467-9922.00193

Qian, D. D., & Schedl, M. (2004). Evaluation of an in-depth vocabulary knowledge measure for assessing reading performance. *Language Testing*, *21*(1), 28–52.

RAND Reading Study Group. (2002). *Reading for understanding: toward a research and development program in reading comprehension*. Santa Monica, CA: RAND.

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC nonword database. *The Quarterly Journal of Experimental Psychology: Section A*, *55*(4), 1339–1362.

Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing*, *10*(3), 355–371. https://doi.org/10.1177/026553229301000308

Read, J. (1998). Validating a test to measure depth of vocabulary knowledge. In A. Kunnan (Ed.), *Validation in language assessment* (pp. 41–60). Mahwah, NJ: Erlbaum.

Read, J. (2000). *Assessing vocabulary*. Cambridge, UK: Cambridge University Press. Retrieved from http://ebooks.cambridge.org/ref/id/CBO9780511732942

Read, J. (2004). Plumbing the depths: How should the construct of vocabulary knowledge be defined? In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition, and testing* (p. 209227). Amsterdam, Netherlands: John Benjamins.

Richards, J. C. (1976). The role of vocabulary teaching. *TESOl Quarterly*, 77–89.

Richardson, J. T. E. (2007). Measures of short-term memory: A historical review. *Cortex*, *43*(5), 635–650. https://doi.org/10.1016/S0010-9452(08)70493-3

Schmitt, N. (2010a). Key issues in teaching and learning vocabulary. In M. R. C. Beltrán, C. A. Contesse, & M. del M. T. López (Eds.), *Insights into non-native vocabulary teaching and learning* (pp. 28–40). Bristol, UK: Multilingual Matters.

Schmitt, N. (2010b). *Researching vocabulary: A vocabulary research manual*. Basingstoke, England: Palgrave Macmillan.

Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, 1–39. https://doi.org/10.1111/lang.12077

Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, *95*(1), 26–43. https://doi.org/10.1111/j.1540-4781.2011.01146.x

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. https://doi.org/10.1177/026553220101800103

Schoonen, R., & Verhallen, M. (2008). The assessment of deep word knowledge in young first and second language learners. *Language Testing*, *25*(2), 211–236. https://doi.org/10.1177/0265532207086782

Segalowitz, N. (2003). Automaticity and second languages. In C. J. Doughty & M. H. Long (Eds.), *The Handbook of Second Language Acquisition* (pp. 382–408). Blackwell Publishing Ltd. Retrieved from http://onlinelibrary.wiley.com.proxy1.cl.msu.edu/doi/10.1002/9780470756492.ch13/summary

Shin, H. W. (2015). Psychometric properties of word association test with regard to adolescent EFL learners. *Vocabulary Learning and Instruction*, 9.

Shiotsu, T., & Weir, C. J. (2007). The relative significance of syntactic knowledge and vocabulary breadth in the prediction of reading comprehension test performance. *Language Testing*, *24*(1), 99–128. https://doi.org/10.1177/0265532207071513

Silverman, R. D., Speece, D. L., Harring, J. R., & Ritchey, K. D. (2013). Fluency has a role in the Simple View of Reading. *Scientific Studies of Reading*, *17*(2), 108–133.

Snow, C. E. (2010). Academic Language and the Challenge of Reading for Learning About Science. *Science*, *328*(5977), 450–452. https://doi.org/10.1126/science.1182597

Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. National Academies Press.

Tannenbaum, K. R., Torgesen, J. K., & Wagner, R. K. (2006). Relationships between word knowledge and reading comprehension in third-grade children. *Scientific Studies of Reading*, *10*(4), 381–398. https://doi.org/10.1207/s1532799xssr1004_3

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*(2), 127–154.

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York, NY: Academic Press.

van Gelderen, A., Schoonen, R., de Glopper, K., Hulstijn, J., Simis, A., Snellings, P., & Stevenson, M. (2004). Linguistic knowledge, processing speed, and metacognitive knowledge in first- and second-language reading comprehension: A componential analysis. *Journal of Educational Psychology*, *96*(1), 19–30. https://doi.org/http://dx.doi.org.proxy2.cl.msu.edu/10.1037/0022-0663.96.1.19

van Gelderen, A., Schoonen, R., Stoel, R. D., de Glopper, K., & Hulstijn, J. (2007). Development of adolescent reading comprehension in language 1 and language 2: A longitudinal analysis of constituent components. *Journal of Educational Psychology*, *99*(3), 477–491. https://doi.org/http://dx.doi.org.proxy2.cl.msu.edu/10.1037/0022-0663.99.3.477

Verhoeven, L. (2000). Components in early second language reading and spelling. *Scientific Studies of Reading*, *4*(4), 313–330. https://doi.org/10.1207/S1532799XSSR0404_4

Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics*, *22*(2), 217–234.

Walter, C. (2004). Transfer of reading comprehension skills to L2 is linked to mental representations of text and to L2 working memory. *Applied Linguistics*, *25*(3), 315–339.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale - Third Edition (WAIS®-III)* (3rd ed.). San Antonio, TX: Psychological Corporation.

Wesche, M. B., & Paribakht, T. S. (1996). Assessing second language vocabulary knowledge: depth versus breadth. *Canadian Modern Language Review*, *53*, 13–40.

Wright, T. S., & Cervetti, G. N. (2016). A systematic review of the research on vocabulary instruction that impacts text comprehension. *Reading Research Quarterly*, 1–24. https://doi.org/10.1002/rrq.163

Zhang, D. (2012). Vocabulary and grammar knowledge in second language reading comprehension: A structural equation modeling study. *The Modern Language Journal*, *96*(4), 558–575. https://doi.org/10.1111/j.1540-4781.2012.01398.x

Zhang, D., & Yang, X. (2016). Chinese L2 learners' depth of vocabulary knowledge and its role in reading comprehension. *Foreign Language Annals, 49*(4), 699–715.