Breadth and Depth of English Vocabulary Knowledge: Which Really Matters in the Academic Reading Performance of Chinese University Students?

Hsing-Fei Huang
Department of Integrated Studies in Education
McGill University, Montreal
January, 2006

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Arts in Second Language Education

© Hsing-Fei Huang, 2006
NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.
Abstract

This study explored the relationship between vocabulary size (i.e., breadth of knowledge), depth of vocabulary knowledge, and reading comprehension of Chinese-speaking ESL (English as a second language) university students in Canada. Both aspects of vocabulary knowledge, breadth and depth, continue to play roles in vocabulary research. Few studies, however, have focused on which aspect plays the predominant role in L2 reading. Using three language tests—the GRE (Graduate Record Examinations) for reading comprehension, Nation’s (1990) Vocabulary Levels Test, and Read’s (1998) Word Associates Test—and verbal reports, the general purpose of the study was to examine the relationship between vocabulary knowledge and reading comprehension, and the specific focus was to find out which aspect of vocabulary knowledge, breadth or depth, has greater impact on determining reading comprehension performance. The results demonstrate that (1) test scores on vocabulary size, depth of vocabulary knowledge, and reading comprehension are positively correlated, (2) vocabulary size is a stronger predictor of reading comprehension than depth of vocabulary knowledge, and (3) breadth and depth of vocabulary knowledge are closely interrelated and mutually facilitative. The findings suggest the importance of vocabulary size in reading comprehension for the population tested.
Résumé

La présente étude porte sur la relation entre la taille du lexique (étendue des connaissances), la profondeur de la connaissance lexicale et la compréhension écrite des sinophones en anglais langue seconde au Canada. Ces deux aspects de la connaissance lexicale, taille et profondeur du lexique, continuent d’intéresser la recherche thématique. Peu d’études, cependant, s’attachent à déterminer lequel des deux aspects joue le rôle prédominant dans la compréhension écrite en langue seconde. En utilisant trois tests linguistiques—le GRE (Graduate Record Examinations) pour la compréhension écrite, celui de Nation (1990) Vocabulary Levels Test et celui de Read (1998) Word Associates Test—ainsi que des rapports verbaux, cette étude vise à examiner en général la relation entre la connaissance lexicale et la compréhension écrite, et plus particulièrement à déterminer quel aspect de la connaissance lexicale, taille ou profondeur du lexique, influe d’avantage sur la compréhension écrite. Les résultats établissent que (1) l’analyse des tests de taille du lexique, de la profondeur de la connaissance lexicale et de la compréhension écrite présentent une corrélation positive, que (2) la taille du lexique est un meilleur indicateur de la compréhension écrite et que (3) la taille du lexique et la profondeur de la connaissance lexicale sont étroitement liées et mutuellement facilitantes. Les conclusions soulignent l’importance de la taille du lexique pour la compréhension écrite de la population étudiée.
Acknowledgements

I wish to express my gratitude to many people for their contributions in various ways to the completion of this thesis. I am thankful that many wonderful people have dedicated their time, energy, and knowledge to helping me achieve this goal.

I want to begin my sincere thanks to my supervisor, Dr. Carolyn Turner, for her constant guidance and encouragement throughout both my thesis work and graduate studies at McGill University. Without her support and confidence in me, this thesis could not have been started. Her patience, kindness, and thoughtfulness are very much appreciated. I am also deeply indebted to Dr. Tom Cobb (at Université du Québec à Montréal), my co-supervisor, for his insightful comments which have made a great difference to my thesis. His advice and expertise on second language reading and vocabulary have been extremely helpful throughout the entire course of my thesis writing. I feel fortunate to have both Dr. Carolyn Turner and Dr. Tom Cobb as my supervisors.

I would like to thank Dr. John Read (now at the University of Auckland, New Zealand) for providing his unpublished Word Associates Test for my study. A special thanks to Dr. David Qian (at the Hong Kong Polytechnic University) for his excellent research that inspired me to conduct this study. A heartfelt thanks to Dr. Janet Donin, of the Department of Educational and Counselling Psychology at McGill, for helping me out with her extensive knowledge on statistics. Also, I thank all the people who participated in this study; the investigation of this study could not be accomplished without their cooperation and participation.

My deepest appreciation goes to my good friend and colleague, Mark Maby, for his editorial help on my thesis; his sense of humor and useful suggestions have kept me moving along throughout this learning process. For all the support, I want to extend my sincere gratitude to my colleagues at McGill, my family and friends back home in Taiwan, and friends at MTGSA (McGill Taiwanese Graduate Students’ Association), especially to Catherine, Daniel, Eric, and Sidney for their continuous support, encouragement, and prayers. A big thanks to Diane Philip at McGill’s Education Library, for her patience, research skills, and knowledge of the English language.

I would like to dedicate this thesis to my mom, Li-Rong Lee, whose support has allowed me to have an unforgettable learning experience in Canada; I thank her for everything. Last, but not least, I thank God for giving me these supportive people and lending me strength to carry out the amazing journey of conducting this study.
Table of Contents

Page
Abstract ...................................................................................................................... i
Résumé ...................................................................................................................... ii
Acknowledgements ................................................................................................... iii

CHAPTER 1 INTRODUCTION ............................................................................. 1
Background ................................................................................................................ 2
Objective of the Study ............................................................................................... 3
Preview of the Chapters ............................................................................................. 5

CHAPTER 2 LITERATURE REVIEW ................................................................. 7
Overview ................................................................................................................... 7
The Nature of Vocabulary Knowledge ...................................................................... 7
What is a Word? .......................................................................................................... 7
What is Vocabulary Knowledge? .............................................................................. 10
Assessing Breadth and Depth of Vocabulary Knowledge and Their Relationship .... 14
Breadth of Vocabulary Knowledge ........................................................................ 14
Depth of Vocabulary Knowledge .......................................................................... 19
Relationship between Breadth and Depth of Vocabulary Knowledge ................. 23
The Role of Vocabulary Knowledge in Reading Comprehension ....................... 23
The Vocabulary-Reading Comprehension Chain in L1 Research ...................... 24
The Exposed Tip of the Iceberg in L2 Reading Comprehension: Vocabulary Size ... 26
A Relationship Needing Exploration: Vocabulary Depth and Reading
Comprehension ........................................................................................................ 27
Moving Forward ...................................................................................................... 28

CHAPTER 3 RESEARCH DESIGN AND METHODOLOGY .......................... 30
Overview .................................................................................................................. 30
Research Design ...................................................................................................... 30
Research Purpose ................................................................................................... 30
Research Question ................................................................................................ 31
Research hypotheses ............................................................................................... 31
Statement of inquiry ............................................................................................... 31
Method ..................................................................................................................... 32
CHAPTER 4 PRESENTATION AND DISCUSSION OF RESULTS .......... 47
Overview ......................................................................................................... 47
Scores on the Three Language Tests ............................................................... 47
Statistical Analyses and Results ..................................................................... 47
  Phase 1: Descriptive statistics and test reliability ......................................... 47
  Phase 2: Correlations .................................................................................... 48
  Phase 3: Multiple regression analyses .......................................................... 48
Examining the Proposed Hypotheses and Discussion .................................... 51
Hypothesis 1: Test scores on reading comprehension, vocabulary size, and depth of vocabulary knowledge will correlate positively with one another. 51
Hypothesis 2: Depth of vocabulary knowledge will be a more powerful predictor of reading comprehension performance than vocabulary size .... 53
Verbal Reports on the DVK ........................................................................... 55
Summary ......................................................................................................... 63

CHAPTER 5 CONCLUSIONS .............................................................................. 64
Overview ......................................................................................................... 64
Summary of the Findings ................................................................................... 64
  Main Findings in Response to the Research Question .................................... 64
  Other Related Findings .................................................................................. 66
Implications for L2 Instruction ........................................................................ 67
Limitations of the Study .................................................................................... 68
Directions for Future Research ......................................................................... 69
Final Remarks ................................................................................................... 71
REFERENCES ........................................................................................................... 72

APPENDICES

Appendix A .............................................................................................................. 79
Appendix B .............................................................................................................. 85
Appendix C .............................................................................................................. 90
Appendix D .............................................................................................................. 91
Appendix E .............................................................................................................. 92
Appendix F .............................................................................................................. 93
Appendix G .............................................................................................................. 94
List of Tables

<table>
<thead>
<tr>
<th>Tables</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1: The Participants’ Background Information by University, Educational Level, and Sex</td>
<td>33</td>
</tr>
<tr>
<td>Table 2: Distribution of the Participants’ Major Concentrations</td>
<td>33</td>
</tr>
<tr>
<td>Table 3: Ranges of the Participants’ Age</td>
<td>33</td>
</tr>
<tr>
<td>Table 4: The Participants’ Duration of Stay in Canada</td>
<td>34</td>
</tr>
<tr>
<td>Table 5: The Participants’ Period of Studying English in Their Native Country</td>
<td>34</td>
</tr>
<tr>
<td>Table 6: The Test Order Each Participant Undertook</td>
<td>44</td>
</tr>
<tr>
<td>Table 7: Descriptive Statistics and Reliability of the RC, VS, and DVK</td>
<td>48</td>
</tr>
<tr>
<td>Table 8: Pearson Correlation (2-tailed) between Scores on the RC, VS, and DVK</td>
<td>49</td>
</tr>
<tr>
<td>Table 9: Multiple Regression Analyses Using Score on the RC as Criterion Variable and Scores on the VS and DVK as Predictor Variables</td>
<td>50</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

Polonius: What do you read, my lord?
Hamlet: Words, words, words.

- Shakespeare, Hamlet, Act 2, Scene II

Reading is a complex process; reading in a foreign language is even more complex. Reading comprehension, in both first (L1) and second language (L2), is affected by many variables, the most researched being background knowledge, reading strategies, and vocabulary knowledge. This study investigates the role of vocabulary knowledge in L2 reading comprehension. It compares the relative importance of two aspects of vocabulary knowledge, the superficial, word-counting, or ‘breadth’ aspect, and the more complex, multidimensional, or ‘depth’ aspect, in the L2 (English) reading performance of Chinese university students.

Since the early 1980s, there has been a massive upsurge in vocabulary studies in second language acquisition (SLA) generally, and L2 reading specifically. Yet many of these vocabulary studies were based on an earlier understanding of the nature of L2 learners’ vocabulary knowledge (e.g., Cronbach, 1942). Two common assumptions were (1) that word knowledge is mainly or only about meaning, and (2) that learners either do or do not know what a word means. These assumptions were reflected in vocabulary testing (e.g., Nation's, 1990, Vocabulary Levels Test). The test-takers’ task is basically to match a word with its synonym or dictionary-type analytic definition. This kind of conventional vocabulary test simply equates word knowledge with knowledge of its meaning and form. Going hand in hand with a relatively simplistic notion of vocabulary knowledge was a strong emphasis on counting how many words learners might know, that is, measuring their vocabulary size (e.g., Goulden, Nation,
& Read, 1990). Of course, some researchers had all along emphasized the more complex nature of vocabulary knowledge (e.g., Richards, 1976), and the tendency since the late 1990's has been to incorporate more complexity into word knowledge models (e.g., Nation, 2001). This complexity includes learners’ knowledge of frequency, associations, collocations, and registers, the efficiency of their lexical access, and the distinction between receptive and productive knowledge. At present, both aspects of vocabulary knowledge, breadth and depth, continue to play roles in vocabulary research, but with few investigations into which aspect plays the predominant role for specific types of language use, such as L2 reading.

**Background**

Why was the role of vocabulary in L2 reading not paid much attention until the 1980s? Since the recent ‘re-discovery of vocabulary’ forms the main background to this study, it is worth pausing to fill in this background briefly. Viewing reading as a psycholinguistic process, Goodman (1968) and Smith (1978) argued that linking text context with readers’ knowledge about language and about the world in general can foster effective reading and that reading is a matter of bringing meaning to print. Thus, many ESL practitioners (e.g., Twaddell, 1973) have claimed that since skilled L1 readers can skip new words and still grasp the main ideas of a text, L2 readers should follow the same principle if they intend to reach a level as fluent and efficient as L1 readers’. These ESL practitioners therefore advocated an emphasis on building up reading strategies, rather than vocabulary knowledge. This is the ‘top-down approach’ to L2 reading that has been widely accepted in the last 30 years. Since the top-down approach sounds quite appealing, many researchers overlooked the fact that it is just part of the story. What if learners do not even know the meanings of many words in a
In L1 research (e.g., Thorndike, 1973), it has long been recognized that vocabulary knowledge is a distinct factor in determining reading comprehension. By contrast, due to the predominance of the top-down approach to L2 reading research in the 1970s and 80s, bottom-up processing in L2 reading attracted far less attention. Not until recently have researchers begun to realize that lower-level, lexical processing indeed plays an important role in L2 reading. Haynes and Baker (1993), for example, concluded that what differentiates Chinese college students from their American counterparts in reading an English text is their vocabulary knowledge, not their reading strategies. That is, the place of vocabulary among the potential sources of difficulty in L2 reading had been underestimated and this should no longer be the case.

Objective of the Study

The present study attempts to look closely at the relationship between vocabulary knowledge and reading comprehension. As a native speaker of Chinese, I wonder whether my own learning experience is typical of that of other Chinese L2 readers. As a result, Chinese students have been chosen as participants in this study. The general purpose of the study is to examine the relationship between vocabulary knowledge and reading comprehension of Chinese students, and the specific focus is to find out which aspect of vocabulary knowledge, breadth or depth, has greater impact on determining their reading comprehension performance.

Zimmerman (1997) noted that "Vocabulary is central to language and of critical importance to the typical language learner" (p. 5), and I could not agree more with this. When learning a new language, as an L2 learner myself, I find that vocabulary is the
most basic and influential element of the message compared with other elements such as syntax. For known words, I have found that lexical access (i.e., efficient retrieval or production of a word) is particularly important. When there are some unknown words in a text, even those L2 learners who have spent a great deal of time studying the L2, including myself, often have comprehension problems in reading. Many researchers have attempted to explain the source of reading difficulties for L2 learners and have demonstrated that insufficient vocabulary is the most serious handicap for L2 learners while reading (e.g., Yorio, 1971). From both my personal learning experience and the literature, vocabulary appears to play a more important role and constitute more of a problem than people usually recognize.

It is well-known that conventional Chinese education pays much attention to intensive word-by-word reading (Johnson & Ngor, 1996). When it comes to ESL (English as a Second Language) learning, the demands for translations of all types of reading materials then occur. Most Chinese students' attitude that knowing an English vocabulary item is restricted to knowing the meaning of a word could be traced back to their learning method, that is, list memorization. Many Chinese-speaking ESL learners, including myself, learned huge numbers of words from lists. We occasionally saw those vocabulary items in reading materials but had few chances to practice retrieving or producing them in context. In other words, we were, to some extent, encouraged to build up our vocabulary knowledge in a superficial way. In fact, most Chinese learners barely got into more dimensions of vocabulary knowledge other than the aspect of the word meaning.

Consequently, the role of vocabulary knowledge in L2 reading is a problem that needs attention. Seeking a better understanding of the nature of the learner’s vocabulary knowledge is a good starting point through such questions as the following: What is vocabulary knowledge? To what extent is vocabulary knowledge significant
in reading comprehension? Which aspect of vocabulary knowledge matters most in reading comprehension? Some of these questions have already received answers, as will be shown in the literature review of this thesis; others which have not received answers will form the focus of the present study.

The results of this study are expected to have instructional implications for Chinese academic ESL students in particular and possibly for academic L2 learners in general. If we knew which type of word knowledge was more important, we could design instruction to target it, and this could lead in turn to enhanced reading ability or even other language skills. Moreover, the results could provide the fields of vocabulary research and of vocabulary assessment with deeper insight into lexical learning and use. It is hoped that this study will also encourage researchers to pursue other perspectives on the breadth versus depth issue in further studies.

Preview of the Chapters

This thesis has five chapters. This chapter has introduced the objective, rationale, and background of the study. Following this introduction, Chapter 2 reviews the relevant literature, mainly from an assessment perspective. The definitions of key terms are also provided. It begins with the nature of vocabulary knowledge, then two dimensions of vocabulary knowledge are discussed, and related vocabulary measures are presented. The role that vocabulary knowledge plays in reading comprehension is reviewed in both L1 and L2 research. Chapter 3 explains the research design and methodology of the study. The research question, participants, instruments, and procedures for data collection are described. Chapter 4 presents the results of the data analyses and discusses the findings. Both quantitative and qualitative findings are displayed and interpreted. Finally, Chapter 5 summarizes the main and related
findings from the present study and puts forth concluding remarks.
CHAPTER 2
LITERATURE REVIEW

Not only every language but every lexeme of a language is an entire world in itself.
- Mel'čuk, 1981 (p.570)

Overview

This chapter provides a review of the relevant literature. It begins with the nature of vocabulary knowledge. Vocabulary knowledge is then discussed in terms of two dimensions, breadth and depth. The literature is reviewed mostly from an assessment perspective, in which several well-known vocabulary measures are introduced and discussed in terms of the two dimensions. Studies on the relationship between the two dimensions are also reviewed. Finally, the role of the research literature on vocabulary knowledge in reading comprehension is examined with regard to both L1 and L2 studies.

The Nature of Vocabulary Knowledge

What is a Word?

In vocabulary testing, a basic assumption is that we are assessing knowledge of words. As a result, it is essential to realize the nature of what we want to assess. The first question to ask, then, is ‘What is a word?’ Common-sense notions of a word, say orthographic units between white spaces or dictionary entries with their own lines, belie the fact that ‘word’ is not an easy concept to define. However, it is an issue of importance at both theoretical and practical levels, especially the latter. For instance, the definition of a ‘word’ becomes crucial when we attempt to estimate the size of a learner’s vocabulary with words as the basic unit of measurement.
When it comes to counting ‘words’, it is worth introducing a few key terms. In English, there is a variety of word forms. For example, we have the word *accept*, but then there are also *accepts, accepted, acceptable, acceptably*, and *acceptability*. Are they counted as six different words or different forms of the word *accept*? Before exploring this question, we need to point out that the word formation includes two types—inflections and derivations. Inflections are what is involved in adding inflectional endings to a base form (or root), without changing its meaning and its word class (Carter, 1998). That is, among the given example, *accept* is the base and *accepts* and *accepted* are its inflected forms. In contrast, the word forms which add a new element of meaning to the base or change its word class are regarded as derivations, such as *acceptable, acceptably* and *acceptability*. So, in the given example, they can be regarded as six different words, as they are displayed, or as one base form *accept* plus its two inflected forms (*accepts* and *accepted*) and three derived forms (*acceptable, acceptably*, and *acceptability*). For instance, Carroll, Davies and Richman (1971) counted words based on their displays due to the use of computer programs. That is, *accept, accepts, accepted, acceptable, acceptably, and acceptability* are counted as six different words in that way.

In many vocabulary studies, however, the base form and just its inflected forms are considered as a sort of super-word, or *lemma* (Read, 2000). When counting words in a text, the researcher’s first step is often to lemmatize the running words and then count just those. In the given example, *accept, accepts, and accepted* then belong to the same lemma. In other studies, researchers adopt *word family*, rather than lemma, as the counting benchmark. The base form, its inflected forms, and its derived forms are collectively known as a word family, having a common underlying meaning (Read, 2000). In this case, *accept, accepts, accepted, acceptable, acceptably, and acceptability* belong to the same word family and add up to one word. Goulden,
Nation and Read (1990), for instance, produced a list of base words, each representing a word family. So, is fly, flying, flies, and flyer four words or one word family? The answer is not either-or and what counts as a word usually depends upon the researcher's primary purposes.

In addition, in English a word often has multiple meanings applied in different contexts and many occurrences of a cluster of words begin with a word. **Polysemy** is 'the existence of several meanings in an individual word' (Carter, 1998, p.12) and those meanings may be close or distant. For example, the senses of line in straight line, power line, and railway line are relatively closer than the senses of raise in raise your hand and raise children. **Collocation** is a term used to represent 'a group of words which occur repeatedly in a language' (Carter, 1998, p.51). That is, certain words or certain types of words are expected before or after a word. These patterns of co-occurrence can be lexical or grammatical (Carter, 1998). For instance, we might expect 'make a decision' but not 'do a decision'; 'interest in something' but not 'interest at something.' The fact that one word may turn out to be many (i.e., in the sense of polysemy) and many may turn out to be one (i.e., in the sense of collocation) seems to make the concept of a word even more complicated. So, while assessing learners' knowledge of a word, is it simply 'a word' being assessed?

In view of the complexities in the 'What is a word' question, a word cannot be usefully defined, at least pedagogically, as simply any string of letters between white spaces. Presumably, then, nor should this definition be used, explicitly or implicitly, as the basis of a vocabulary test. When knowledge of the word liberal is assessed in a vocabulary test, for instance, and the learners answer it correctly, what does this correct answer mean? Do they simply know the word liberal? Is it valid to assume that they also know liberate, liberally, liberalism, and liberator? Do they know the word liberal through an associative network? Do they know liberal in the sense of
'open-minded', 'generous', or 'not strict'? Does their word knowledge about liberal result from the encounters of 'liberal views' and 'liberal parents'? Is liberal, the whole word family to which liberal belongs, or the associative network within which liberal lies, being assessed? Does vocabulary testing need to accommodate these complexities by taking more inclusive aspect of word knowledge into account?

Considering that testing more dimensions of vocabulary knowledge is at the expense of testing more words, is depth testing or size testing more useful? Before looking into this question, let us first know what vocabulary knowledge is all about. If 'What is a word' is a more complex question than it appears, 'What does it mean to know a word' is even more so.

What is Vocabulary Knowledge?

What does it mean to 'know a word'? As Labov (1973) mentioned, "Words have often been called slippery customers, and many scholars have been distressed by their tendency to shift their meanings and slide out from under any simple definition" (p. 341). Many L2 learners regard vocabulary learning as a matter of memorizing L2 word lists. Thus, they go to a bilingual dictionary whenever they encounter an unknown word. Very often, their vocabulary knowledge simply involves knowing rough L1 equivalents for L2 words, and yet with this knowledge they can do quite well on vocabulary tests which look only to see whether a learner can match or produce L2 words with their L1 equivalents. Nevertheless, as a matter of fact, vocabulary knowledge goes beyond this simplistic perspective.

Over the years, lexical researchers have built up various criteria for what is involved in knowing a word in both L1 and L2 research. In an earlier era, Cronbach (1942) classified vocabulary knowledge into five categories by focusing on word meaning (generalization, breadth of meaning, and precision of meaning) and use
(application and availability). Realizing the restriction in Cronbach's (1942) framework (i.e., that it relies solely on word meaning and little on word use), Richards (1976) integrated more aspects of lexical knowledge, such as morpho-syntactic properties, associations, frequency level, and register into Cronbach's framework. Richards (1976) presented a set of principles regarding lexical competence. His first principle is that native speakers' vocabulary continues to expand in adult life, as opposed to relatively little further development of their syntax. He then provided the other seven principles on what is meant by knowing a word:

2. Knowing a word means knowing a 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.

(Richards, 1976, p. 83)

Principle 2 suggests word knowledge includes knowledge of collocation. When encountering the word flower, people might expect the words such as beautiful, wild, and fragrant; that for meal people might expect delicious, wholesome, and substantial. Principle 3 concerns register restraints, as in petrol for the British but gas for Americans and Canadians. Principle 4 deals with the syntactic properties of words. For example, we cannot say *The committee named but The committee named a new
chairperson, because the transitive verb name requires a direct object. Knowing a word involves recognizing the base forms when they are combined with inflectional and derivational suffixes in Principle 5. Thus, if learners know the word adapt, it is most likely for them to know the words such as adapts, adapted, adaptable, adaptation, and adaptor. Principle 6 suggests words come to mind via associative networks. The association could be organized by coordinate classification (apple-banana), superordinate classification (apple-fruit), synonymy (happy-glad), or antonymy (good-bad). In Principle 7, knowing a word entails knowledge of semantic features, such as human/non-human, animate/inanimate, or positive/negative. For instance, we say John is laughing, rather than *The chair is laughing; Albert Einstein is famous but not notorious. Principle 8 states multiple meanings that an individual word often has. To elaborate this principle, the polysemous word off is given as an example:

Marie's off in Italy somewhere. (at a distance)
Take your coat off. (Something has been removed)
They're off. (= the race has begun)
The wedding is off. (cancelled)
The TV is off. (not functioning)
I need some time off. (away from work or duty)

The set of Richards' (1976) seven principles has frequently been taken as a comprehensive framework of vocabulary knowledge because it highlights the complexity of vocabulary knowledge, which involves a great deal more than simply knowing the meaning of a word.

Largely retaining Richards' (1976) definition of knowing a word, Nation (1990) added such important components as pronunciation and collocations to make the framework more inclusive. He proposed 16 questions on what are necessary to fully
know a word with the incorporation of the receptive-productive distinction. According to Nation’s (1990) classification system, he also indicated that the ability to use a word (i.e., production) requires further knowledge than the ability to understand it (i.e., reception) does. Later, Nation (2001) revised his earlier framework in order to improve the relatedness between various aspects of word knowledge and its practicability. Within his revised classification scheme, the refined 18 questions on knowing a word arose to reflect three different types of lexical knowledge—form (spoken form, written form, and word parts), meaning (form and meaning, concept and referents, and associations), and use (grammatical functions, collocations, and constraints on use).

The research reviewed above provides a clear state-of-the-art picture of what is involved in knowing a word. The implications for vocabulary testing, however, are far from clear. The next questions are ‘What construct do we want to measure with a vocabulary test?’ and ‘How do we operationalize the construct we want to measure?’ The main choice of constructs at present is breadth, based on the old word counting approach, and depth, which attempts to incorporate some of the new complexities about words and knowing words that was outlined above. Is there any real contest between the two? Although it is an ideal goal to assess the full range of vocabulary knowledge in a single test, in practice there are limits on how long a vocabulary test can be. Accordingly, various types of test are used to measure particular aspects of vocabulary knowledge, or to measure types of knowledge that can infer the existence of other types. The next section carefully unpacks the terms breadth and depth in the context of vocabulary knowledge and testing.
Assessing Breadth and Depth of Vocabulary Knowledge and Their Relationship

To conceptualize vocabulary knowledge in the present study, it is necessary to clarify the constructs breadth and depth. According to Anderson and Freebody (1981), breadth of knowledge refers to the quantity of words for which one knows at least some important aspects of meaning (i.e., how many words are known?), while depth involves understanding the quality of one’s lexical knowledge (i.e., how well are particular words understood?). Researchers have used numerous types of assessment with different formats to measure vocabulary knowledge (see Wesche & Paribakht, 1996, for a comprehensive discussion of various assessment types). In the discussion to follow, we will look more closely at how to measure these two dimensions of vocabulary knowledge and their relationship.

Breadth of Vocabulary Knowledge

Breadth of vocabulary knowledge has been defined as vocabulary size, which is what many lexical researchers have focused on measuring. Although the perspective of vocabulary size is easy to criticize, in fact it has been the central measure in a great deal of highly useful research, some highlights of which will be reviewed below.

According to Webster’s Third International Dictionary, there are around 54,000 word families in English (Nation & Waring, 1997). Nation and Waring (1997), however, reported the vocabulary size of an educated adult native English speaker to be only around 20,000 word families. Although there are clearly differences across individuals, this estimate is generally accepted. Other contemporary studies have also obtained similar results. However, Jamieson (as cited in Nation & Waring, 1997) found that many adult learners of English as a foreign language (EFL) have a vocabulary size of fewer than 5,000 word families, even after they have studied English for several years.
Also, simple size testing has shown that as well as fewer words, L2 learners tend mainly to know a different type of word from native speakers. Glancing at the variation in vocabulary size between native and non-native English speakers (i.e., 20,000 versus 5,000 word families), we know, simply, that there is a vocabulary size gap between them. But how different are their vocabulary profiles? It is important to first realize that not all English words are equally useful and that the way to measure usefulness is usually by word frequency, which considers how often the words occur in normal language use. For example, a small number of high-frequency words (e.g., *the, is, in*) occur repeatedly and therefore learners will know a large proportion of the running words in a text if they know these words. Besides these high-frequency words, most of the vocabulary that an educated adult native English speaker has (i.e., about 20,000 word families) is made up of relatively low frequency words (Nation, 1993) whereas L2 learners’ vocabulary mainly contains the high-frequency words.

Let us now turn to the question of how we measure vocabulary size. Vocabulary size measures are discussed here in two aspects: selection of test items and development of test formats. First, there are two sources of sampling for vocabulary size measures: dictionary-based and frequency-based. Dupuy (1974) randomly sampled words from *Webster's Third New International Dictionary* and then selected 123 words that met three criteria as basic words to be included in a vocabulary test. A basic word had to be a main entry, a single word form (i.e., not a derived form nor a compound), and could not be archaic, foreign, technical, or slang (Dupuy, 1974). Dupuy (1974) calculated that there are around 12,300 basic words in *Webster's Third New International Dictionary*, so the 123 vocabulary items represented 1% of the basic words in the dictionary. Therefore, the way to measure someone’s vocabulary size is to multiply the score on the test by 100. For instance, a person whose score is 50 would be approximated to have a vocabulary size of 5,000 words. In other words, a
person's vocabulary size is estimated by multiplying the number of sample words known by the ratio that those sample words bear to the total number of words in the dictionary. To demonstrate concretely the estimation of vocabulary size with the dictionary-sampling method, Nation (1990, p. 76) generated a simple formula as follows:

\[
\frac{N \text{ correct answers} \times \text{total N words in dictionary}}{\text{N items in test}} = \text{vocabulary size}
\]

This apparently straightforward procedure has not, however, been entirely successful. The procedure has yielded huge variation in estimates of vocabulary size (Lorge & Chall, 1963; Meara, 1996). As a result, most lexical research has tended to sample not from a dictionary but from a frequency distribution of words, the alternative basis for selecting test items for the assessment of vocabulary size. Anderson and Freebody (1981) claimed that “Frequency is a parameter which probably is very strongly related to probability that a word will be known” (p.101). The assumption is that words that occur more frequently will be learned relatively earlier, so that learners' knowledge of words at given frequency levels displays their overall vocabulary size. Insufficient knowledge of frequent words is often regarded as an indication of smaller vocabulary size, whereas knowing relatively infrequent words seems to reflect greater vocabulary size and proficiency. The frequency count approach of estimating vocabulary size involves the use of word frequency lists such as the Thorndike and Lorge (1944) and Kučera and Francis (1967) lists.

After knowing the two methods of selecting test items, the second aspect to consider is test formats. Several measures of vocabulary size were devised, based on researchers' different notions of knowing a word. Of these, two tests have been widely
used: the Eurocentres Vocabulary Size Test and the Vocabulary Levels Test, which are both frequency-sampling tests but with distinct test formats. The Eurocentres Vocabulary Size Test (EVST) was developed by Meara and his colleagues (Meara & Buxton, 1987; Meara & Jones, 1988), using a yes/no checklist format. According to Melka Teichroew (1982), the yes/no checklist format has been used with native speakers at least since 1890. Meara and his colleagues then employed the yes/no checklist format to work on estimating the vocabulary size of L2 learners. The EVST consists of a set of real words covering 10 different frequency bands drawn from the Thorndike and Lorge (1944) list, as well as carefully crafted plausible pseudo-words that adjust the possible overestimation of test-takers' lexical knowledge. This test makes very few demands on test-takers. The test-takers have to identify whether they know the meaning of a word by choosing 'yes' for a positive response and 'no' for negative. If test-takers claim to know some pseudo-words, their final score will be reduced. This test is available in both paper-and-pencil and computerized modes. The EVST has been viewed as a reliable and valid test (Meara, 1996; Meara & Jones, 1988) and some L2 researchers (e.g., Harley, Howard, & Roberge, 1996; Schmitt & Meara, 1997) have used this format to measure learners' vocabulary size. Most appealingly, this test is easy to construct and to score, and its simple format makes it possible to test a large number of words within a short time (around 10 minutes) (Meara, 1996; Read, 2000; Wesche & Paribakht 1996). The test has been used to place thousands of learners, reasonably successfully, in Eurocentre language courses.

Despite these promising features of the test, even Meara and his associates soon found that this test did not function well in some settings. For example, in Meara and Buxton's (1987) study, they found that the test works better with some L1 groups than with others. Speakers of Romance languages (e.g., French and Italian) had more difficulty rejecting non-words than Germanic speakers, due to the influence of the
cognate effect. In addition, since test-takers are simply required to self-report their word recognition without having word meanings verified, test-takers’ confidence in their lexical competence (i.e., willingness to select ‘yes’ or ‘no’) may play a role when they provide responses (Meara, 1996), and thus the validity of this test may be in doubt to some extent.

The Vocabulary Levels Test (VLT) was devised by Nation (1990) originally as a diagnostic vocabulary test for classroom teachers and then it was later used to estimate vocabulary size for L2 learners of general or academic English. It is a paper-and-pencil test with a word-short definition matching format. In response to the task, the VLT requires demonstration of knowledge on the meanings of the target words, which is more than what the EVST demands. The VLT was chosen as an instrument assessing breadth of vocabulary knowledge in the present study, and more details on it will be provided in the next chapter. Although making good estimates of individuals’ vocabulary size is not an easy task, the VLT has been accepted by a number of researchers (e.g., Laufer & Paribakht, 1998; Qian, 1998, 1999, 2002; Yu, 1996) as an appropriate measure of vocabulary size, in terms of reliability and validity. As a result, this test has become widely used both in vocabulary research and as a vocabulary assessment for L2 learners. Meara (1996) called it the “nearest thing we have to a standard test in vocabulary” (p. 38).

Among the various measures of vocabulary size, the EVST and VLT are simple but reliable, and both have achieved some success in providing both researchers and language teaching institutions with a measure of learners’ approximate vocabulary size (Qian, 1999). Nevertheless, tests with the frequency-sampling method, such as the EVST and VLT, will give an inaccurate estimation of learners’ knowledge of English words if the test-taker’s pattern of vocabulary acquisition does not fit that of the assumed frequency profile (Read, 2000). Within the frequency-sampling method,
the assumption is that words that occur more frequently in the language are more likely to be known by learners than less frequent words. Meara and Jones (as cited in Read, 2000) stated that most learners fit this pattern quite closely, but Meara (as cited in Read, 2000) found problems in interpreting the vocabulary size estimate of some groups whose frequency profile on English words varies from the assumed one, affected by their L1 backgrounds and educational experience.

In addition, most measures of vocabulary size, even the two well-known tests mentioned above, have been solely used to determine whether learners know a single (usually the most common) meaning of a word but have not been used to assess multiple meanings of words (Nation, 1990) or any of the other numerous dimensions of knowing words elaborated above. That is, the assessment of vocabulary breadth is a somewhat shallow, all-or-nothing measure. The bottom line of the vocabulary size measures is either that one word corresponds to one meaning or that it is a word recognition task (i.e., simply giving a yes/no response to the question ‘Do you know this word?’). In summary, despite their reasonable successes, there are practical problems with the two main vocabulary breadth tests, and in addition these tests make no attempt to incorporate the richer notions of word knowledge that have been disclosed in recent research. Accordingly, we now turn to considerations of how depth information can be incorporated in a vocabulary test.

*Depth of Vocabulary Knowledge*

Depth of vocabulary knowledge has been used to refer to the quality of vocabulary knowledge, that is, how well one knows a word. Vocabulary knowledge is often “a matter of degree” (Read, 1993, p. 357): even native speakers have only partial knowledge of words they know. In earlier years, several L1 lexical researchers produced scales to indicate the learner’s varying degrees of partial knowledge of the
meaning of words they know. Dale (1965, p. 898), for instance, outlined a four-stage developmental sequence on the continuum of knowing a word:

Stage 1: 'I never saw it before.'
Stage 2: 'I've heard of it, but I don't know what it means.'
Stage 3: 'I recognize it in context — it has something to do with . . .'
Stage 4: 'I know it.'

Nevertheless, the literature demonstrates that knowing a word should mean more than knowing its meaning in a particular context. Lexical researchers have elaborated many aspects of vocabulary knowledge, such as different meanings, underlying forms and derivations, and association network with other words (e.g., Nation, 1990; Richards, 1976). More often than not, L2 learners, even advanced ones, have incomplete word knowledge and do need depth in their vocabulary knowledge to perfect their comprehension and production. The question, then, is how much word knowledge, and of which specific aspect, is sufficient, for which purposes?

Unlike breadth of vocabulary knowledge, depth of vocabulary knowledge has not gained the attention it deserves. People have not investigated the multidimensionality of lexical knowledge until recently. The fact that there are so many components involved in depth of knowledge makes the measuring of depth a complex task. Read (1989) suggested that researchers conduct studies on L2 learners that involve moving beyond mere quantitative measures of learners' vocabulary size to a more qualitative investigation of how well selected words are understood. The most straightforward way to probe how well a learner knows about a set of target words is through an individual interview. Thus, Read (1989) described how he and his colleagues attempted to develop an interview protocol to elicit information from L2 learners, in order to best reflect the depth and quality of their vocabulary knowledge. Given the
lack of previous research of this kind on L2 learners, they outlined a rough framework of interview questions concerning learners’ word knowledge in pronunciation, meaning, collocations or associations, and underlying forms. With this unstructured procedure, however, it was difficult to attain consistency from one interview to the next. In addition, using such a method was time-consuming and the number of words that could be investigated in any degree of depth was limited.

Realizing the limitations of the interview method, Read (1993) proposed a less open test format measuring depth of vocabulary knowledge by means of word associations, known as the Word Associates Test (WAT). It is a simple response task and allows relatively broader coverage of a set of words. The task is to identify the related words, or associates, for each target word based on their three possible relationships: paradigmatic (synonyms), syntagmatic (collocates), and analytic (representing part of the word meaning). As reported by Read (1993), the test reliability (KR-20) of the WAT has reached .92. Read (1998) continued revising and improving the test, and it has been used to measure depth of vocabulary knowledge in a number of lexical studies (e.g., Nassaji, 2004; Qian, 1998, 1999, 2002). Also, it was chosen as the instrument for measuring depth of vocabulary knowledge in the present study. Further details on this test will be elaborated in the next chapter.

Another salient measure of depth was the Vocabulary Knowledge Scale (VKS) developed by Paribakht and Wesche (1993). The VKS was devised in the context of research on incidental vocabulary acquisition of ESL learners at the University of Ottawa in Canada. Similar to Dale’s (1965) four-point vocabulary development scale for L1 learners, the VKS involves five categories. However, other than indicating their degree of familiarity with target words, the VKS test-takers are asked to demonstrate their word knowledge by writing down the word meaning and composing a likely sentence with the target word if they further know the usage of the target word. The
sentence allows the researchers to collect and score information on the word’s grammar, collocation, and other dimensions in addition to the merely semantic (hence making it ‘deeper’ or more multi-dimensional than Dale’s 1965 test mentioned above). The VKS thus combines self-report with elicitation of verifiable responses. It is a generic instrument, making it possible to be applied with any set of words that researchers or testers intend to assess. Despite its high practicality, the VKS has limitations as a measure of depth of vocabulary knowledge. For example, multiple meanings of a word, an important aspect of lexical knowledge, are not necessarily measured by the VKS. Caution should be taken not to interpret learners’ well-composed sentences as evidence of word knowledge. McNeil (as cited in Read, 2000), who investigated vocabulary knowledge of Chinese-speaking ESL teachers, found that some of his participants were able to compose reasonable sentences to show the use of a target word even if they did not know what the target word meant. As McNeil mentioned, this may reflect Chinese students’ way of learning English by memorizing several multi-word lexical items without exactly comprehending the meaning of each individual word. Referring to the VKS, Read (2000) noted, “… the whole concept of a fixed scale may not be the most appropriate way to characterize the different forms of knowledge that people have of words” (p. 138).

Depth measures such as the WAT and VKS certainly probe deeper lexical knowledge than the EVST and VLT. Nevertheless, with depth measures, the number of words that can be assessed is fairly limited, and the process of responding and scoring usually takes quite an investment in time and labor. Breadth measures (i.e., estimates for vocabulary size), on the other hand, can cover a wide range of words from several frequencies or content areas within a relatively short time, and their simple content and format reduce the time for responding and scoring and also make it possible for them to be computerized. There is indeed a trade-off between how much
can be examined about each word and how many different words can be included.

**Relationship between Breadth and Depth of Vocabulary Knowledge**

Breadth and depth of vocabulary knowledge are obviously related (depth without some breadth is logically impossible), but how are they related? There are as yet few studies focusing on the relationship between breadth and depth of vocabulary knowledge. Empirical support for the fact that these two dimensions are related can be found in two studies. In Schmitt and Meara’s (1997) study on the English vocabulary knowledge of secondary and postsecondary Japanese students, they found that correlations between vocabulary size and word association (a depth measure) were relatively high (.61 for receptive knowledge and .62 for productive knowledge), as opposed to other intercorrelations in their study. Similarly, Nurweni and Read (1999), examining English vocabulary knowledge of Indonesian university students, reported that the overall correlation between the scores on the tests of breadth and depth of vocabulary knowledge was .62 and the relationship became even stronger ($r = .81$) when the sample group was narrowed down to high-proficiency students exclusively.

Based on those results, it could be inferred that the development of these two dimensions, namely breadth and depth of vocabulary knowledge, is associated to a large extent by the end of second language acquisition. But this does not mean the close association occurs at earlier stages in the process.

**The Role of Vocabulary Knowledge in Reading Comprehension**

After reviewing the research literature on assessment of breadth and depth of vocabulary knowledge and their relationship, we now come to the role of vocabulary knowledge in reading comprehension. Without specifically referring to breadth or
depth of vocabulary knowledge, a large number of studies that involve the relationship between overall vocabulary knowledge and reading comprehension have been well documented in L1 reading research.

**The Vocabulary-Reading Comprehension Chain in L1 Research**

Unanimous support for a strong correlation between overall lexical competence and reading comprehension can be found in pertinent L1 research (e.g., Davis, 1968; Klare, 1974; Thorndike, 1973). However, those studies are fairly descriptive rather than explanatory and there has been little agreement on the reason for the strong link between the two. Anderson and Freebody (1981) proposed three hypotheses to explain this relationship—the instrumentalist, knowledge, and aptitude hypotheses. The instrumentalist hypothesis claims that knowing the word is the essential and causative factor in text comprehension. The knowledge hypothesis suggests that background knowledge is important for text understanding and that a large vocabulary is merely a by-product of a large body of background knowledge. That is to say, it is the learning of concepts, not just of word meanings, that determines reading comprehension. The aptitude hypothesis suggests that vocabulary knowledge reflects verbal aptitude and that this verbal aptitude is the crucial factor in reading comprehension. Instead of accepting the idea of fixed capacity (i.e., the aptitude view), Mezynski (1983) added the fourth hypothesis, the access view, believing that reading comprehension may involve several ‘trainable’ skills such as accessing word meanings and applying meta-cognitive strategies, and that the amount of practice is the key to improving them. Some proponents of the access view (e.g., Perfetti & Lesgold, 1979) are interested in the automaticity of word knowledge and thus emphasize the importance of rapid lexical access in reading.

Which viewpoint is the most tenable, then? As Anderson and Freebody (1981)
noted, no serious scholar holds any one of these views to the exclusion of the others. In other words, none of them are considered mutually incompatible. Nevertheless, among these viewpoints, Anderson and Freebody (1981) regarded the instrumentalist perspective as embodying 'unquestioned tenets rather than hypotheses in need of verification' (p. 83). Several researchers (e.g., Becker, 1977; Yap, 1979) found evidence that vocabulary is a potent causal factor in reading comprehension. Yap (1979) concluded that 'casual links probably do exist between vocabulary and comprehension and that vocabulary is likely to be the predominant causal factor' (p. 58). The instrumentalist view has the most direct implications for instruction. Among studies investigating whether reading comprehension is facilitated by vocabulary training, some found positive effect (e.g., Kameenui, Carnine, & Freschi, 1982) and some reported no difference (e.g., Jenkins, Pany, & Schreck, 1978). However, those studies that fail to find support for the relationship between vocabulary knowledge and reading comprehension result from either the small number of participants with wide ranges of reading ability (Jenkins et al., 1978) or the inappropriate construction of assessment materials (Tuinman & Brady, 1974), rather than the flaw in the instrumentalist perspective per se. Consequently, the instrumentalist view appears to provide the best account of the role of vocabulary knowledge in reading. In a word, the relationship between overall vocabulary knowledge and reading comprehension in L1 research has been consistently proved to be associated and some studies even supported a causal relationship. The only thing is that the researchers did not clearly indicate what kind of vocabulary knowledge was involved.

When it comes to L2 studies, researchers did specify certain type of vocabulary knowledge and explored its role in reading comprehension; however, investigations on the impact of different aspects of vocabulary knowledge upon reading comprehension were unbalanced. More specifically, the majority focused on the relationship between
vocabulary size and reading comprehension but relatively few studies probed the relationship between vocabulary depth and reading comprehension.

*The Exposed Tip of the Iceberg in L2 Reading Comprehension: Vocabulary Size*

Examining the factors involved in the four hypotheses mentioned above, Nation (1993) used a triangular diagram with three variables—vocabulary size, skill in use, and knowledge of the world—to describe their interrelationship as L2 learners’ vocabulary develops. While discussing one of the relationships in his triangular diagram, Nation (1993) stated that learners’ skill in use (e.g., reading comprehension) depends on vocabulary size, in spite of Nagy’s warning against cautiously adopting the instrumentalist view that may “lead to a reductionist view of reading, i.e., if you know the vocabulary then that is about all you need to be able to read” (p. 117). It appeared that Nation’s position here supports the reductionist view of reading. Also, it seems that the dimension of depth of vocabulary knowledge has been ignored in L2 research ever since. As Meara (1996) remarked, “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).

There is substantial empirical evidence supporting the strong effect of vocabulary size on L2 proficiency, particularly in reading comprehension. In Liu and Nation’s (1985) study, they reported that it is necessary for an L2 learner to have a 95% coverage of a text (i.e., encountering one unknown word in approximately every 20 known words). Conducting research with 92 first-year university students whose L1 is Hebrew or Arabic, Laufer (1992) specifically indicated that this 95% lexical threshold for academic reading comprehension corresponds to knowing about 3,000 word families. That is, learners’ English reading will be impeded by having a vocabulary size below 3,000-word family level. By adopting Nation’s conversion formula, Laufer
translated 3,000 word families into around 5,000 lexical items (or individual word forms) (Laufer, 1992, 1997; Nation, 1993). In other words, L2 learners of English need a vocabulary of at least 3,000 word families, or 5,000 lexical items, to reach an adequate level of comprehension in general English academic texts.

Furthermore, high correlations between vocabulary size and reading comprehension were also revealed in Laufer’s (1992) study. She reported that the scores on reading comprehension correlated with both those on the VLT (Vocabulary Levels Test) \( r = .50 \) and those on the EVST (Eurocentres Vocabulary Size Test) \( r = .75 \). Knowing that this correlational evidence can not be interpreted as a casual relationship, Laufer (1992) concluded that vocabulary size is a good indicator of reading comprehension based on her results.

\*\*\*\*\*\*

A Relationship Needing Exploration: Vocabulary Depth and Reading Comprehension

Although studies on the relationship between vocabulary knowledge and reading comprehension in L2 gradually began to grow, the majority emphasized the role of vocabulary size in reading, as discussed in the preceding passage. In comparison, there have been few studies exploring the relationship between depth of vocabulary knowledge and reading comprehension. This is probably due to a relative shortage of depth measures for L2 learners (Read, 1988) and to the complexity of assessing depth of vocabulary knowledge (Qian, 1999). There has been only one study that I know of associating depth of vocabulary knowledge and reading comprehension. De Bot, Paribakht, and Wesche (1997) aimed to model L2 lexical processing in the context of reading through interviews and think-aloud protocols. Although de Bot et al. (1997) did not intend to probe the connection between depth of vocabulary knowledge and reading comprehension, their study found that various aspects of vocabulary knowledge, such as word morphology, word associations and other depth indicators
are closely linked to reading comprehension processes.

Therefore, to amend the imbalanced relationship between two dimensions of vocabulary knowledge (i.e., breadth and depth) and reading comprehension, more recently Qian (1998, 1999, 2002) investigated the interrelationships among vocabulary size, depth of vocabulary knowledge, and reading comprehension. Looking at the role of breadth and depth of vocabulary knowledge in 74 Chinese and Korean readers’ comprehension of general academic texts in English, Qian (1998, 1999) found that scores on vocabulary size, depth of vocabulary knowledge, and reading comprehension were highly correlated, and that depth of vocabulary knowledge made a unique contribution to the prediction of learners’ reading comprehension performance, over and above the prediction already provided by vocabulary size. He concluded that the dimension of vocabulary depth is just as important as that of vocabulary size in predicting reading performance. Later, Qian (2002) conducted a similar study with 217 participants from 19 different L1 backgrounds and obtained the same results, confirming the importance of the role of both depth and breadth of vocabulary knowledge in reading comprehension.

Moving Forward

As an L2 learner of English whose L1 is Chinese, I personally experience the important role that vocabulary plays in English proficiency, especially reading. Even at an advanced level, most L2 learners are still aware of their limited vocabulary knowledge. Unlike native speakers, L2 learners’ process of vocabulary acquisition is more conscious and demanding. Typically, Chinese students have vocabulary breadth but do not receive enough vocabulary depth, due to their method of learning vocabulary, that is, list memorization. Realizing the need for vocabulary research exclusively on Chinese speakers and inspired by Qian’s (1998, 1999, 2002) excellent
study, I have thus carried out a study investigating the role of breadth and depth of vocabulary knowledge in the academic reading performance of Chinese-speaking ESL students. The present study partially replicates Qian’s (1999) earlier study. Minor changes to Qian’s instrumentation and supplementation of his data collection with qualitative information are described in the next chapter, which deals with research design and methodology.
CHAPTER 3
RESEARCH DESIGN AND METHODOLOGY

Overview

This chapter first outlines the research design for the present study by stating the purpose and the research question of the study. A mixed-methods—quantitative and qualitative—approach is applied to explore the research question. The methods for carrying out this study are then presented by introducing the selected participants, instruments, and procedures for data collection. A pilot experiment is reported in which I attempt to establish the appropriateness of the materials within the chosen instruments and within a feasible test time. At the end of this chapter, the data analysis is explained as preparation for presenting results in the next chapter.

Research Design

Research Purpose

The general purpose of this study is to investigate Chinese university students' English vocabulary knowledge in academic settings and its impact on their reading performance. The vocabulary knowledge here includes two aspects: breadth and depth. In the present study, the relationship between vocabulary size (i.e., breadth of knowledge), depth of vocabulary knowledge, and reading comprehension is explored through language tests and self-reports to be described later in this chapter. The main focus of this study is to compare the relative importance of the two aspects of vocabulary knowledge, breadth and depth, in reading comprehension performance for these learners.
Research Question

The research question under examination is: What is the relationship between reading comprehension, vocabulary size, and depth of vocabulary knowledge for Chinese-speaking ESL university students? To address this question, a mixed-methods approach is used. The quantitative part is hypothesis-driven and the qualitative part focuses on self-reports of the learner’s vocabulary knowledge (i.e., data-driven) through verbal protocol.

Research hypotheses (Quantitative part)

Based on the findings of pertinent studies reviewed in the previous chapter, the following hypotheses are proposed:

1) Test scores on reading comprehension, vocabulary size, and depth of vocabulary knowledge will correlate positively with one another.

2) Depth of vocabulary knowledge will be a more powerful predictor of reading comprehension performance than vocabulary size.

Statement of inquiry (Qualitative part)

According to Qian’s (1998, 1999, 2002) study, depth of vocabulary knowledge seemed to be a very important factor in reading comprehension. As a result, the depth of participants’ vocabulary knowledge and the differences in vocabulary knowledge across participants will be further explored through verbal reports. The verbal reports will also give an insight into the relationship between the participants’ vocabulary breadth (i.e., vocabulary size) and depth. The verbal reports serve as a source of data for the researcher to infer cognitive processes and to obtain in-depth information.
Method

Participants

Participants in the present study were 24 university students. They were attending either McGill University or Concordia University, the two English-speaking universities in Montreal, Quebec, Canada. Of the participants, 12 were male and 12 female. They were recruited from different academic fields on a voluntary basis. The age range of the participants was from 22 to 37, with a mean age of 26.33 (see Tables 1 through 3 for details).

The target language tested in this study was English. The L1 for all participants was Mandarin Chinese, a language that is not cognate with English and is very different from English phonologically and orthographically. Therefore, it was expected that guessing unknown English words through the assistance of L1 in language tests of this study would be almost impossible.

The 24 native Chinese speakers were from either Taiwan (7) or China (17), two countries where English is spoken as a foreign language. The participants' duration of stay in Canada was controlled for those who had been resident in Canada for less than three years. This was done to restrict the study to participants who had primarily learned English in EFL (English as a Foreign Language) settings. Also, Chinese students were not accepted if they had attended high school or university in Canada (or any English-speaking country) either in English or French before attending McGill University or Concordia University. Information on the participants’ duration of stay in Canada and period of studying English in their native country is summarized in Tables 4 and 5.

International students whose native language is not English are generally required to take a standardized language proficiency test, usually either TOEFL (Test of English as a Foreign Language) or IELTS (International English Language
Table 1
_The Participants’ Background Information by University, Educational Level, and Sex (n=24)_

<table>
<thead>
<tr>
<th>University</th>
<th>McGill=15</th>
<th>Concordia=9</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ed. level</strong></td>
<td><strong>Undergraduate</strong></td>
<td><strong>Cont. Ed.</strong></td>
</tr>
<tr>
<td>Undergraduate</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Sex</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Note. <sup>a</sup>Cont. Ed. = Continuing Education
<sup>b</sup>Of 13 graduate students at McGill University, 6 were master’s students and 7 were doctoral students.
<sup>c</sup>The 9 graduate students at Concordia University were all master’s students.

Table 2
_Distribution of the Participants’ Major Concentrations (n=24)_

<table>
<thead>
<tr>
<th>Major concentrations</th>
<th>No. of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agricultural &amp; Environmental sciences</td>
<td>1</td>
</tr>
<tr>
<td>Education</td>
<td>5</td>
</tr>
<tr>
<td>Engineering</td>
<td>10</td>
</tr>
<tr>
<td>Management studies</td>
<td>1</td>
</tr>
<tr>
<td>Medicine</td>
<td>2</td>
</tr>
<tr>
<td>Music</td>
<td>1</td>
</tr>
<tr>
<td>Science</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 3
_Ranges of the Participants’ Age (n=24)_

<table>
<thead>
<tr>
<th>Age range</th>
<th>No. of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>21-25</td>
<td>13</td>
</tr>
<tr>
<td>26-30</td>
<td>10</td>
</tr>
<tr>
<td>Over 30</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 4

*The Participants’ Duration of Stay in Canada (n=24)*

<table>
<thead>
<tr>
<th>Duration of stay (months)</th>
<th>No. of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-5</td>
<td>11</td>
</tr>
<tr>
<td>6-10</td>
<td>0</td>
</tr>
<tr>
<td>11-15</td>
<td>2</td>
</tr>
<tr>
<td>16-20</td>
<td>4</td>
</tr>
<tr>
<td>21-25</td>
<td>4</td>
</tr>
<tr>
<td>26-30</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 5

*The Participants’ Period of Studying English in Their Native Country (n=24)*

<table>
<thead>
<tr>
<th>Period of studying English (years)</th>
<th>No. of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-5</td>
<td>1</td>
</tr>
<tr>
<td>6-10</td>
<td>7</td>
</tr>
<tr>
<td>11-15</td>
<td>14</td>
</tr>
<tr>
<td>16-20</td>
<td>2</td>
</tr>
</tbody>
</table>

Testing System), and to obtain the required score before being admitted to a course of study. In the present study, all the participants met this requirement, except for one who was admitted to Concordia University after taking CELDT (Concordia English Language Diagnostic Test) as an alternative language test. Among the 23 participants who took the standardized language proficiency test, 21 took the TOEFL and their scores ranged from 207 (543 for the paper-based version) to 283 (657 for the paper-based version); 19 out of these 21 participants obtained scores above 213(550 for the paper-based version). For the other two participants who took the IELTS, one scored 6 and the other 7. According to their performance on the TOEFL or IELTS, the participants were regarded as advanced English learners.

*Instruments*

The instruments used in the present study include three language tests, verbal
The three language tests were administered by paper and pencil, the verbal reports by audio recording, and the questionnaire by e-mail.

**Reading comprehension test**

All but three of the participants in the present study (two who had taken the IELTS and one the CELDT) had taken the reading comprehension section of the TOEFL and had scored at an acceptable level. However, Qian (1999) has pointed out that a more sophisticated test for reading comprehension than the reading comprehension section of the TOEFL is needed to discriminate between test-takers for research as opposed to placement purposes. In light of this consideration, the reading comprehension test (hereafter RC) used for this study, instead of the TOEFL reading comprehension section, was the GRE (Graduate Record Examinations) reading comprehension subtest of the verbal section developed by ETS (Educational Testing Service). The RC test items were taken from the GRE General Test, in which the topics are not related to any specific field of study. In each edition of the GRE General Test, there are three or more reading comprehension passages, each providing the basis for answering two or more multiple-choice questions. The passages are across different subject areas: the humanities, the social sciences, and the natural sciences (ETS, 2003). The RC test items for the present study are drawn from the official GRE General Test administered in 1989 (ETS, 1990, Form GR90-14, pp. 208-210, 239-241), which contains two relatively short reading comprehension passages, each with four multiple-choice questions, and two relatively long passages, each with seven multiple-choice questions. There are 22 multiple-choice questions in total. Each multiple-choice question is based on the content of its corresponding passage. On the basis of what is stated or implied in the passage, the participants were asked to choose
one best answer to each question from the five options provided after reading the passage.

According to ETS (2003), the purpose of the GRE’s reading comprehension questions is to measure test-takers’ ability to read with understanding, insight, and discrimination. Sufficient context is provided for test-takers to analyze the written passage from different perspectives, including in both explicit and implicit modes. As for the reading comprehension questions, they focus on (1) the main idea or primary purpose of the passage, (2) information explicitly stated in the passage, (3) information or ideas implied or suggested by the author, (4) possible application of the author’s ideas to other situations, (5) the author’s logic, reasoning, or persuasive techniques, and (6) the tone of the passage or the author’s attitude as it is revealed in the language used (ETS, 2002). In other words, the reading comprehension questions are intended to assess test-takers’ comprehension of the texts as a whole rather than the meanings of specific words. It was expected that this RC would discriminate enough to pinpoint test-takers’ general reading comprehension ability.

*Vocabulary size test*

The English vocabulary size test (hereafter VS) chosen for this study was the Vocabulary Levels Test (Nation, 1990). It is in five sections, representing five different word frequency levels—the 2,000-word level, the 3,000-word level, the 5,000-word level, the University Word List (UWL) level, and the 10,000-word level. According to Nation (1990), the 2,000- and 3,000-word level sections include only high-frequency words in English; the 5,000-word level is a boundary level between the high-frequency and low-frequency word levels; and the 10,000-word level section includes relatively low-frequency words. For example, the words *accident, jump,* and *slow* are from 2,000-word level, *brilliant, charity,* and *prevail* from 3,000-word level,
blend, fragrant, and trumpet from 5,000-word level, and froth, morose, and squint from 10,000-word level. Words in these sections were taken from Thorndike and Lorge’s (1944) list, with reference to the word-frequency data from the General Service List (West, 1953) and Kučera and Francis (1967). Without specifically reflecting the frequency distribution of words in general English, however, the UWL level section contains medium-frequency academic vocabulary that is from the 3,000-to 5,000-word level with respect to the language at large but much more frequent within and across academic domains. The words coordinate, hypothesis, and ultimate, for instance, are from the UWL level. Words in the UWL level section were sampled from Campion and Elley’s (1971) academic vocabulary list for non-native English speakers attending universities in English-speaking countries. In the VS, the words in each section were randomly selected (but proper nouns and compound words were excluded) so that they would represent all the rest of the words at that level (Nation, 1990). That is, its results can indicate the proportion of the total number of words at each level the test-takers know.

As for the test format, the VS consists of matched words and short definitions. Unlike the conventional practice, the definitions, rather than the words, are the test items. In other words, test-takers are required to match the definition to the words. They write the corresponding number of the correct word next to its definition. At each level, there are 36 words and 18 definitions, in groups of six and three respectively, as in the following example from the 2,000-word level:

1. original
2. private complete
3. royal first
4. slow not public
5. sorry
6. total
At each level, test-takers need to check 36 words against 18 definitions to make the correct matches. Such a test structure may reduce the chances of guessing correctly to one in six as against one in four in a standard multiple-choice test. In addition, the words used in the definitions are always more common than the matched words. To avoid providing any grammatical clue for the definitions, all the words in each group belong to the same word class. Moreover, the words and definitions that are related in meaning, except for the correct matches, do not cluster together. There are 18 test items at each level and thus a total of 90, which can be processed by most test-takers in less than 50 minutes (Nation, 1990). Brief test instructions with an example are provided in advance of the actual test. (see Appendix A for the complete test).

*Depth of vocabulary knowledge test & Verbal report*

The depth of vocabulary knowledge test (hereafter DVK) in this study was the Word Associates Test (WAT), developed by Read (1993, 1998). This test was devised to measure test-takers' depth of receptive English vocabulary knowledge in terms of three elements: synonymy, polysemy, and collocation. That is, the WAT is intended to assess learners' depth of vocabulary knowledge through word associations, which are the semantic and collocational relationships that a word has with other words in the language. As mentioned in Chapter 2, this test replaced the time-consuming production approach that Read (1989) had previously experimented with. The test used in the present study was version 4.0 of the WAT (J. Read, personal communication, October, 2004).

Each test item comprises one stimulus word, which is always an adjective, and eight words located in two different columns, each containing four words. Among these eight words, four are associates (i.e., related words) to the stimulus word and the other four are distractors. Most adjectival stimulus words were select from Barnard's
Second and Third Thousand Word Lists (Nation, 1986), which involve more high frequency academic vocabulary than the UWL does, but those adjectives that have only one basic meaning or restricted usage, however, were excluded. An example taken from the actual WAT is presented below:

<table>
<thead>
<tr>
<th>beautiful</th>
</tr>
</thead>
<tbody>
<tr>
<td>enjoyable</td>
</tr>
<tr>
<td>expensive</td>
</tr>
<tr>
<td>free</td>
</tr>
<tr>
<td>loud</td>
</tr>
<tr>
<td>education</td>
</tr>
<tr>
<td>face</td>
</tr>
<tr>
<td>music</td>
</tr>
<tr>
<td>weather</td>
</tr>
</tbody>
</table>

The words in the left column are adjective forms, among which the associates are either synonyms of the stimulus word (i.e., testing knowledge of synonymy) or they represent one aspect of its various meanings (i.e., testing polysemy). In the right column are four nouns, among which the associates can collocate with the stimulus word (i.e., testing collocation). The DVK obviously measures more dimensions of word knowledge than the VS does: the VS examines only meaning while the DVK does that (i.e., with synonymy) plus two other components (i.e., polysemy and collocation). These components represent important aspects of depth of vocabulary knowledge. However, fewer words can be tested in the DVK than in the VS, given that depth testing requires more time on each test item. The DVK in this study thus contains 40 items as against the VS’s 90. Before the actual test, test instructions along with an example question were prepared to guide the test-takers in how to respond to the test items, because the structure of this test is not self-explanatory.

One modification in Read’s (1998) test instructions was made for the present study. In such comprehension tasks with selected responses as the DVK, the problem of guessing seems inevitable. Although it is good to provide test-takers with several options to select, their answers do not necessarily indicate whether they understand the test items. In other words, without understanding the test items, test-takers may be
able to apply whatever strategies or techniques to guess or guess blindly. Read revised his old test structure designed earlier in order to reduce the effect of guessing, and reached a more refined test structure (Read, 1998), this being the one which was used in the present study. In Read’s (1998) study, however, he pointed out that the problem of guessing is still a factor in test-takers’ performance and that a further investigation of the test-taking process remains to be done. Therefore, to minimize the possible influence of guessing in the DVK, for the present study, a modification of the instructions of the original test by Read has been attempted. In Read’s (1998) test, the test-takers were instructed that there are always four correct answers in each test item, whereas in the present study the test-takers were not informed of the exact number of correct answers. By taking this measure, it was hoped that the test-takers’ responses would clearly reflect what they know about those test items and were not merely the result of filling in the required number of choices. Although there are four correct answers in the example given to the participants, they were told that the example was just a demonstration of responding and that they should follow the instructed principle of circling associates that they found related to the stimulus word (see Appendix B for the complete test).

In addition, to monitor the possible guessing factor, if any, and to further explore the relationship between the participants’ breadth and depth of vocabulary knowledge, four participants from the Chinese group were asked to give verbal reports on the DVK once they had finished working on all the test items. In this case, a delayed retrospective verbal protocol (i.e., test-takers’ reflections on the whole process after completion of the test) was undertaken in order not to intervene in the test-takers’ performances during the tests. That is, the four participants were asked to recall and report orally on their test-taking processes. For the purpose of this study, the term ‘verbal report’ is used throughout the discussion of the study hereafter. To take
advantage of their fresh memories for why and how they had chosen their answers, the four participants produced verbal reports right after taking the DVK. If needed, the researcher reminded them to keep talking by giving prompts such as ‘Why did you do that?’ or ‘How did you guess?’ In other words, the participants’ test-taking strategies were elicited whenever applicable. The participants’ retrospection and the researcher’s probing were conducted in the participants’ L1 (i.e., Chinese). This appeared to help the participants to verbalize their thoughts. It was expected that verbalization in the L1 would generate more information. During the protocol, the participants’ reports were digitally recorded for later analysis.

Background questionnaire

The questionnaire was designed to obtain the participants’ background information, such as educational level, duration of stay in Canada, and number of years studying English. It consisted of 10 items. The information was used to better understand the participants in the study and enabled the researcher to describe their characteristics. Since the questionnaire was not used as a measure of language ability, it was presented in Chinese, the participants’ native language, so that they could read the questions faster and without misunderstanding. The participants were presented with a questionnaire written in the orthographic form most familiar to their country: for those from Taiwan, one written with traditional Chinese characters, and for those from China, one written with simplified Chinese characters (see Appendix C for the complete questionnaire and Appendix D for its English version).

Scoring for the Three Language Tests

There were 22 multiple-choice questions in the RC and 90 word-definition matching items in the VS. The responses were coded as either correct or incorrect.
One point was given for each correct answer, so the maximum score was 22 for the RC and 90 for the VS. As for the DVK, it consisted of 40 test items and each had eight options. Different from Read's (1993, 1998) scoring method of valuing only four correct responses out of eight options, all responses, correct or incorrect, were taken into account in the present study. That is, the response for each option (i.e., circled or not) was assessed separately and each correct response was given one point. Even though an option was not circled, it was possible to obtain one point if the response of not circling was correct. However, test-takers who gave no response for a complete test item did not earn a point. The maximum score was eight points for each test item and 320 points for the entire test.

The Pilot Experiment

Before being administered to the 24 participants in the present study, the three language tests—RC, VS and DVK—and the background questionnaire were piloted with two native English speakers and four native Chinese speakers assumed to be comparable to the participants in the present study. Since the RC in this study was taken from the GRE, a standardized test, all the test items had already been through numerous pretests for reliability and validity for a general academic population. Also, both the VS and DVK have satisfactory test reliability, and thus had been used by many lexical researchers. For these reasons, the primary purpose for this pilot experiment was to confirm the appropriateness of the test materials, and to determine the time that test-takers would need to complete the tests.

The result of the pilot test was that the test instructions and items on the three language tests seemed to be well-articulated, without any ambiguity, according to the feedback from the two native English speakers. The questions in the background questionnaire (in Chinese) were also considered comprehensible and unambiguous to
the four native Chinese speakers. In addition, based on the test-taking time of the four
native Chinese speakers, it was decided that the maximum time allotment would be 80
minutes for the RC, 30 minutes for the VS, and 25 minutes for the DVK.

*Procedures for Data Collection*

All required research ethical procedures were followed (see Appendix E). Before
partaking in this study, the participants were asked to sign a consent form (see
Appendices F and G). All the participants (i.e., 24 students) signed the consent form
(in Chinese) and therefore all participated in the present study. The three tests, the RC,
VS and DVK, were administered to each participant in a single testing session. To
eliminate the possibility of an order effect, the tests were administered in a
counterbalanced order. That is, each participant took the three tests in one of six
different orders. The orders are listed as follows:

1. RC-VS-DVK
2. RC-DVK-VS
3. VS-RC-DVK
4. VS-DVK-RC
5. DVK-RC-VS
6. DVK-VS-RC

Test-takers usually do less well on their later tests when some of them may feel
fatigued. It was felt that giving the three tests in a fixed order would negatively bias
the results of the later tests. By varying the order, however, the effect of the fatigue
factor is randomized and hence reduced to a minimum. Each participant was assigned
a random identification code for the purpose of confidentiality and the distribution of
the test order was randomly given. The test order each participant took is summarized
in Table 6.
Table 6

*The Test Order Each Participant Undertook*

<table>
<thead>
<tr>
<th>Test Order</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) RC- VS- DVK</td>
<td>#2, #7, #9, #15</td>
</tr>
<tr>
<td>2) RC- DVK- VS</td>
<td>#1, #3, #14, #18</td>
</tr>
<tr>
<td>3) VS- RC- DVK</td>
<td>#6, #10, #17, #22</td>
</tr>
<tr>
<td>4) VS- DVK- RC</td>
<td>#4, #8, #21, #24</td>
</tr>
<tr>
<td>5) DVK- RC- VS</td>
<td>#13, #19, #20, #23</td>
</tr>
<tr>
<td>6) DVK- VS- RC</td>
<td>#5, #11, #12, #16</td>
</tr>
</tbody>
</table>

Since the participants were volunteers and they all had different schedules, it was difficult to arrange a single testing time for everyone. As a result, four testing sessions were held with groups of different sizes, the largest being ten and the smallest three. The tests were given on the same day or on successive days. However, all the testing sessions were conducted within the same timeframe and the time allotted for each test was strictly controlled. In addition, irrespective of test order, instructions for all three tests were given to the participants before the first test. All the test-takers received the same instructions and explanations, in both oral and written forms (The oral instruction was given in Chinese to ensure that the participants did not misunderstand what they were supposed to do).

Besides the three language tests, four participants from the Chinese-speaking sample were asked to provide extra verbal reports on the DVK in order to monitor the possible guessing factor and to give insight into the participants' vocabulary knowledge. Among the four participants, two were from the RC-VS-DVK order (participants no. 7 and no. 9) and two were from the VS-RC-DVK order (participants no. 6 and no. 10), ensuring that giving verbal reports did not bother them in the middle of their tests. Before the actual provision of verbal reports, these four participants were given individual instructions concerning the retrospective protocol which
included practice items to help familiarize them with the procedure. In this case, there was no time limit for them to produce verbal reports.

After finishing the three tests, the participants were given the background questionnaire. Considering that most of the participants had already spent a great deal of time on the three tests and that they were all volunteers, the questionnaire was sent via e-mail after the participants had finished the three tests. Moreover, they were asked to send their completed questionnaires back to the researcher within five days. All of the participants completed the three tests and the questionnaire in January, 2005.

Data Analysis

The main purpose of the data analysis in the present study was multifold. To test the two research hypotheses proposed in the beginning of this chapter, the quantitative data (i.e., scores of the three language tests) was used (1) to reveal the relationship between reading comprehension, vocabulary size, and depth of vocabulary knowledge, and (2) to determine the more powerful predictor of reading comprehension from scores on vocabulary size and depth of vocabulary knowledge. As for the qualitative data (i.e., verbal reports on the DVK), the results not only checked over the possible guessing factor but also gave insight into the participants' vocabulary knowledge. The relationship between breadth and depth of their vocabulary knowledge was further examined through the verbal reports.

There were two parts of data analysis in the study. First, I used SPSS (Statistical Package for Social Sciences), version 11.0 for Windows, to run statistical analyses of the three language tests. Two-tailed Pearson correlation and multiple regression analysis were the main statistical techniques chosen to meet the research purpose of the study. To explore the intercorrelations among the RC, VS, and DVK, two-tailed
Pearson correlations were calculated. When determining the more powerful predictor of reading comprehension from the variables vocabulary size and vocabulary depth, correlational analysis however seemed not sufficient because correlation only allows the examination of two variables at a time. For this reason, a multiple regression analysis was conducted to provide a more realistic picture of the phenomenon by simultaneously examining a number of variables. Multiple regression analysis has advantages over simple correlational techniques when there are more than two variables involved. In the multiple regression analysis, scores on the VS and DVK were used as the predictors (or independent variables) and score on the RC as the criterion (or dependent variable), in order to determine the stronger predictor of reading comprehension. The results would indicate which predictors are statistically significant in their contribution to the variance in the criterion and how much variance each contributed.

Second, verbal reports on the DVK, the qualitative data, were analyzed without the aid of computer software. Since the verbal reports were in the form of digital recording, I, as a researcher, first transcribed all the recording data in order to focus directly on the participants' test-taking behaviors. From the participants' verbal reports, the guessing factor was monitored. Also, their vocabulary knowledge was further explored through the verbal reports by looking at the commonalities or variations of their behaviors when taking the DVK. During the analysis of the verbal report data, I read and reread the transcription, segmented all meaning units, and identified which ones were relevant to the present study. Categories were then developed from the pertinent data and coded (Bogdan & Biklen, 2003; Tesch, 1990). The patterns that emerged from the verbal reports constituted the results.

In the next chapter, the analyses and results of the study will be presented in turn, and then the findings will be interpreted and discussed.
CHAPTER 4
PRESENTATION AND DISCUSSION OF RESULTS

Overview

This chapter presents and discusses the results of the study. Based on the background questionnaire, all of the 24 Chinese-speaking participants met the criteria set for the study (as described in Chapter 3) and therefore none were excluded. As described in Chapter 3, a mixed-methods approach was used in the study. The results are therefore presented and discussed in two parts: scores on the three language tests and verbal reports on the DVK.

Scores on the Three Language Tests

Statistical Analyses and Results

To test the two proposed hypotheses of the study (as described in Chapter 3), correlation and multiple regression analyses were used. The procedure included three phases.

Phase 1: Descriptive statistics and test reliability

The purpose for this phase of the analysis was to obtain descriptive statistics and reliability coefficients for the participants' scores on the three tests, namely RC, VS, and DVK, for the sample (i.e., Chinese-speaking ESL participants). The results of the descriptive and reliability analyses are shown in Table 7, which lists the score range, mean, standard deviation, and reliability (Cronbach Alpha) of the three tests. Due to the fact of unequal total scores of the three tests, the percentages for the score ranges
Table 7

Descriptive Statistics and Reliability of the RC, VS, and DVK (n=24)

<table>
<thead>
<tr>
<th>Test</th>
<th>Maximum Score</th>
<th>Score Range</th>
<th>Mean Score</th>
<th>Standard Deviation</th>
<th>Reliability (Cronbach Alpha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC</td>
<td>22</td>
<td>5 (23%) – 21 (95%)</td>
<td>12.29 (56%)</td>
<td>4.19</td>
<td>.77</td>
</tr>
<tr>
<td>VS</td>
<td>90</td>
<td>38 (42%) – 90 (100%)</td>
<td>72.33 (80%)</td>
<td>13.05</td>
<td>.94</td>
</tr>
<tr>
<td>DVK</td>
<td>320</td>
<td>216 (68%) – 276 (86%)</td>
<td>249.63 (78%)</td>
<td>15.72</td>
<td>.79</td>
</tr>
</tbody>
</table>

and mean scores are inserted alongside in parentheses to make the three tests more comparable. As shown in Table 7, the reliabilities of the RC, VS, and DVK—.77, .94, and .79 respectively—are all at satisfactory levels.

Phase 2: Correlations

The purpose for this phase of the analysis was to determine the correlations between the scores on the RC, VS, and DVK. A two-tailed Pearson correlation analysis was conducted and the results are displayed in Table 8. The intercorrelations among the three tests are all both positive and statistically significant. The correlation between the RC and VS ($r = .71$) is higher than that between the RC and DVK ($r = .67$), and the correlation between the VS and DVK is the highest ($r = .81$).

Phase 3: Multiple regression analyses

Multiple regression analyses were conducted in this phase to predict reading comprehension performance from, on the one hand, vocabulary size, and on the other, depth of vocabulary knowledge. To determine the more powerful predictor of reading comprehension, scores on the VS and DVK were taken as the predictor (or independent) variables and score on the RC as the criterion (or dependent) variable.
Table 8

*Pearson Correlations (2-tailed) between Scores on the RC, VS, and DVK (n=24)*

<table>
<thead>
<tr>
<th>Test</th>
<th>RC</th>
<th>VS</th>
<th>DVK</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC</td>
<td>—</td>
<td>.71**</td>
<td>.67**</td>
</tr>
<tr>
<td>VS</td>
<td>.71**</td>
<td>—</td>
<td>.81**</td>
</tr>
<tr>
<td>DVK</td>
<td>.67**</td>
<td>.81**</td>
<td>—</td>
</tr>
</tbody>
</table>

**p < .01

Through a multiple regression analysis using SPSS 11.0, I chose the ‘enter’ method, a forced entry option, to enter predictor variables into the regression equation. With this method, SPSS processes all specified variables at one time and reveals the status of all the entered variables, regardless of their significance levels. Table 9 summarizes the results of the multiple regression analysis. The column labeled $R^2$ indicates the proportion of the total variance in the criterion variable (i.e., the RC in this case) accounted for by the predictor variables (i.e., the VS and DVK in this case), and $R^2$ is an estimate of the population value. Adjusted $R^2$ value, however, is developed for the model to better fit the population (i.e., compensating for the optimistic bias of $R^2$ value) and is of concern for models with more than one predictor (SPSS Inc., 1998). Reporting both the $R^2$ and the adjusted $R^2$ is important when there are numbers of predictors and a small sample size (Green, Salkind, & Akey, 1997). $R^2$ change, the difference between an $R^2$ value for the preceding predictor and an $R^2$ value for the predictor being entered, is the highlight of the regression analysis, indicating the magnitude of the contribution of a variable at the point where it is entered into the regression equation.

The previous phrase indicated that the predictor variable VS has a stronger correlation with the criterion variable RC ($r = .71, p< .01$) than the predictor variable DVK ($r= .67, p< .01$). Therefore, the predictor variable VS was chosen to be entered into the regression equation first. The first section of Table 9 (labeled A) presents the
Table 9
*Multiple Regression Analyses Using Score on the RC as Criterion Variable and Scores on the VS and DVK as Predictor Variables (n=24)

<table>
<thead>
<tr>
<th>Step</th>
<th>Predictor variable(s) involved</th>
<th>R²</th>
<th>Adjusted R²</th>
<th>R² Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>A)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>VS</td>
<td>.500*</td>
<td>.477*</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>VS, DVK</td>
<td>.524</td>
<td>.479</td>
<td>.024</td>
</tr>
<tr>
<td>B)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>DVK</td>
<td>.443*</td>
<td>.418*</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>DVK, VS</td>
<td>.524</td>
<td>.479</td>
<td>.081</td>
</tr>
</tbody>
</table>

*p < .05

results where VS was entered first into the equation, followed by DVK. As Table 9 shows, when VS was entered into the equation first, the R² value at this step was .500 and the adjusted R² value .477. VS alone accounted for 50.0% (R² = .500) of the variance in the criterion variable RC. Also, VS, as a predictor, explained a significant amount of the RC variability (F (1, 22) = 22.00, *p* < .05). As VS remained in the equation, DVK was then added at the second step. At this point, the R² value changed to .524 and the adjusted R² value changed to .479. VS and DVK jointly accounted for 52.4% (R² = .524) of the variance in RC but the DVK measure did not predict significantly over and above the VS measure (R² change = .024, F (1, 21) = 1.08, *p* = .311). The entry of DVK at the second step contributed only an additional 2.4% (R² change = .024) of the variance in RC. In other words, DVK did not explain a significant proportion of RC variance after controlling for the effect of VS. The finding that the DVK measure made small but independent further contribution is similar to Qian’s (1998, 1999, 2002) finding, except that the further contribution the DVK measure made in Qian’s study is statistically significant but in this study it is not.
To further investigate the unique contribution made by the DVK measure, another regression analysis was carried out by reversing the order of entry for the predictor variables into the equation. The second section of Table 9 (labeled B) displays the results where VS was entered after DVK. When DVK was entered into the equation at the first step, the $R^2$ value was $0.443$ ($F(1, 22) = 17.51, p< .05$), indicating that DVK alone explained $44.3\%$ of the variance in RC. Next, when VS was added to the equation at the second step, the $R^2$ value increased by $0.081$ ($R^2$ change = $0.081$) to $0.524$. That is, VS explained an additional $8.1\%$ of the variance in RC above the $44.3\%$ variance already accounted for by DVK. However, the VS measure did not predict significantly over and above the DVK measure ($F(1, 21) = 3.59, p = .072$).

Examining the Proposed Hypotheses and Discussion

*Hypothesis 1: Test scores on reading comprehension, vocabulary size, and depth of vocabulary knowledge will correlate positively with one another*

In light of the results of the Pearson correlation analysis, scores on the three language tests did positively correlate with one another, which supports Hypothesis 1. Besides that, the intercorrelations were moderate to high. The correlation between the RC and VS ($r = .71$) was higher than that between the RC and DVK ($r = .67$). This indicates that the score on vocabulary size was more strongly associated with the test-taker's reading comprehension performance than the score on their depth of vocabulary knowledge. This further suggests that vocabulary size appeared to be in a stronger relationship with reading comprehension than vocabulary depth. It should be remembered, however, that correlation does not prove causation. Although Qian (1998) claimed that the strong association between vocabulary knowledge and reading comprehension in his study had provided empirical support for the instrumentalist
hypothesis (Anderson & Freebody 1981, see the previous review in Chapter 2), I would rather only confirm the crucial role of vocabulary knowledge in reading comprehension by this strong association because the causal relationship between the two was not examined in the present study.

Moreover, among the intercorrelations of the three tests, it is intriguing that the correlation was the highest between the scores on the VS and DVK ($r = .81$). It reveals that the two variables, vocabulary size and depth of vocabulary knowledge, are themselves highly correlated. The strong relationship between breadth (i.e., vocabulary size) and depth of vocabulary knowledge is found to be consistent with the results of previous studies (Nurweni & Read, 1999; Schmitt & Meara, 1997). The high correlation shown in the study ($r = .81$) leads us to suppose that the development of breadth and depth of vocabulary knowledge is closely interrelated and may even be interdependent. This appears plausible, for one would not normally have vocabulary size knowledge without also acquiring some depth knowledge. The particularly robust link between vocabulary breadth and depth in the study may result from the high language proficiency of the participants. Since they are mostly advanced English learners whose vocabulary knowledge is generally richer, the breadth and depth of their vocabulary knowledge could interrelate with each other to a large extent. It may also be explained by the overlapping construct of the two measures. Despite the fact that the DVK explores more and deeper aspects of vocabulary knowledge, the synonymy and polysemy that the DVK attempts to measure is actually the basic word meaning that the VS requires, and the knowledge of collocation is more or less affected by knowledge of individual word meaning.
Hypothesis 2: Depth of vocabulary knowledge will be a more powerful predictor of reading comprehension performance than vocabulary size

In the multiple regression analysis, the results suggested that both VS and DVK contributed significantly to the prediction of RC. When comparing the unique contributions they made, however, the analysis yielded results that the VS measure alone accounted significantly for 50.0% of the variance in RC, while only 44.3% of the variance in RC was explained by the DVK measure. In other words, it turned out that vocabulary size is a more powerful predictor of reading comprehension performance than depth of vocabulary knowledge. This is a finding that runs against the original hypothesis. It was initially hypothesized that depth of vocabulary knowledge would be a more powerful predictor of reading comprehension performance than vocabulary size in Hypothesis 2 of the present study; thus, Hypothesis 2 is rejected. Nevertheless, due to the stronger correlation with reading comprehension that vocabulary size shows in this study, this result does not seem surprising at the moment. Although the finding in the present study appears contrary to that of Qian’s (1998, 1999, 2002) study, the patterns shown in Qian’s study and the present study are in fact more similar than different. The one that had a stronger relationship with reading comprehension (as shown in the correlations) would be the more powerful predictor of reading comprehension. In this study, the results suggest the salient role of vocabulary size in reading comprehension.

To better understand this finding, a closer comparison with Qian’s (1998, 1999, 2002) related finding may be useful. In contrast to Qian’s results, what is striking in the present study is that depth of vocabulary knowledge did not make a noticeable independent contribution to the prediction of reading comprehension beyond the prediction afforded by vocabulary size, and neither did vocabulary size. Although both vocabulary size and depth of vocabulary knowledge explained the variance in reading
comprehension significantly (i.e., \( p < .05 \) in both cases), neither of them was able to predict a significant amount of the variability in reading comprehension after controlling for the effect of the other (i.e., \( p > .05 \) in both cases). In other words, vocabulary size and depth of vocabulary knowledge are both good predictors of reading comprehension in their own right, but neither of them is sufficient to account for the variance in reading comprehension when they co-exist. The finding that both vocabulary size and depth of vocabulary knowledge were competing for the existing supply of the variance in reading comprehension, but that neither one made it to the significance level, is confirmed by the close interrelationship and interdependence between vocabulary size and depth of vocabulary knowledge (as found in testing Hypothesis 1). Theoretically, breadth and depth of vocabulary knowledge may be discussed separately; but, apparently, in practice, they are actually inseparable and interrelated. In addition, in terms of the importance of vocabulary to reading generally, the results showed that scores on the breadth and depth of vocabulary knowledge measures jointly accounted for 52.4% of the variance in reading comprehension scores. Although another 47.6% of the variance remained unexplained and unidentified in the present study, the major share of the variance in reading comprehension was already explained by vocabulary knowledge. Again, there is no denying that the role of vocabulary knowledge in reading comprehension is essential, but which aspect of vocabulary knowledge, breadth or depth, plays a more crucial role in reading comprehension, remains in debate.

Given the results in the present study, however, it may be no longer meaningful for this population (i.e., relatively more advanced English learners) to argue for which aspect of vocabulary knowledge is more important in that they are both useful predictors of reading comprehension. Instead, the results further suggested that combining the two dimensions of vocabulary knowledge in vocabulary assessment is
more beneficial than keeping them apart. Nevertheless, the result might depend on language proficiency levels of students. That is, the result of the present study could be due to the strong relation of vocabulary breadth (i.e., vocabulary size) and depth that the participants had. These Chinese-speaking ESL students were advanced learners in an English-speaking context and might be different from those in Chinese universities.

Verbal Reports on the DVK

In Qian's (1998, 1999, 2002) study, the findings demonstrated that depth of vocabulary knowledge is a crucial factor in reading comprehension. To further explore the Chinese-speaking students' vocabulary depth, verbal reports on the DVK were collected from four participants. Through the verbal reports, the relationship between breadth and depth of their vocabulary knowledge was investigated. The four participants were two master's and two doctoral students. Based on their TOEFL and/or GRE scores reported in the background questionnaire and their performance on the three language tests in the present study, they were two high proficiency (hereafter HP) English learners (participants no. 9 and no. 10) and two relatively low proficiency (hereafter LP) learners (participants no. 6 and no. 7). With the delayed retrospective verbal protocol, the four participants verbalized their test-taking process, including identifying strategies being employed. Their audio-taped verbal reports ranged from 47 to 60 minutes. The verbal reports were transcribed and then analyzed. The findings are discussed below.

I identified four common strategies used by the four Chinese participants. First, they were more inclined to circle associates they were certain of, or at least relatively more certain of. They would not guess randomly in the first place unless no or few
responses were evoked. Second, a process of elimination occasionally occurred. They first eliminated those words least associated with the stimulus word, leaving one or two possible associates. This process of elimination might also leave one or few unknown or unfamiliar associates (or distractors), since the participants had already eliminated those words they knew to be unrelated. Thirdly, the positive-negative dichotomy for the adjectival stimulus word was a method of choosing responses, especially when the participants had a vague impression of the stimulus word. That is, instead of knowing the precise meaning of the stimulus word, they claimed they knew whether the word was positive or negative, and were therefore able to choose words of the same trait as their responses. An excerpt from the verbal reports (participant no. 10) in item 19 is presented below as an example.

19  favourable

<table>
<thead>
<tr>
<th>helpful</th>
<th>legal</th>
<th>possible</th>
<th>positive</th>
<th>habit</th>
<th>response</th>
<th>teacher</th>
<th>weather</th>
</tr>
</thead>
</table>

Participant (P): favourable ... ummm ....  
Researcher (R): Do you know this word?  
P: Well, [pause] ... it's hard to explain favorable in words. Favorable ... mm ... could be related to 'favorite' ... should be something good ... something people like ... so it's a positive adjective. Then, I chose helpful and positive because they are also positive adjectives in a sense.

Fourthly, the application of grammatical rules about the stimulus word was also a strategy. These grammatical rules were either taught through their previous vocabulary instruction or internalized from reading. The following are excerpts from the verbal reports (participant no. 7) in items 3 and 15.

3  calm

<table>
<thead>
<tr>
<th>open</th>
<th>quiet</th>
<th>smooth</th>
<th>tired</th>
<th>cloth</th>
<th>day</th>
<th>light</th>
<th>person</th>
</tr>
</thead>
</table>

...
P: In the right-hand side of the column, I only chose *person*.
R: How so?
P: Mmm ... it sounds just right. Oh yes, I was told that this adjective *calm* is used to modify a person, not an object when I first learned this word *calm*. So I eliminated *cloth, day* and *light* and chose *person* only.

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>helpful</td>
<td>interested</td>
<td>missing</td>
<td>strange</td>
</tr>
<tr>
<td>accident</td>
<td>child</td>
<td>computer</td>
<td>steel</td>
</tr>
</tbody>
</table>

P: I chose *curious child*. If *curious* goes with the other three—*accident, computer,* and *steel*, those expressions will just be weird. *Curious* is a word for ‘humankind.’
R: Oh? How do you learn this?
P: Through reading! I kind of generate rules from English texts I’ve read. You know, a result of cumulative learning experience. Now I seldom look up words in a dictionary.

Furthermore, the HP and LP learners’ strategies appeared to vary on the DVK, which is grouped into two parts—meaning and collocation. In the ‘meaning’ section (i.e., the left-hand side of the column in the example given), when the participants saw the stimulus word, its meaning first came out. The LP learners immediately translated the meaning of each word into Chinese and spoke them aloud, trying to make some matches between the stimulus word and the associates (or distractors) whereas the HP learners did not behave so. The four participants chose the semantically-related associates based on what they already knew (i.e., from their mental lexicon) or what they derived from the primary meaning, or core meaning, of the stimulus word. The main differences between the HP and LP learners lie in the vocabulary knowledge they already have. The HP learners knew almost all of the stimulus words and also their associates and distractors; other than that, they mostly knew more than one meaning of the stimulus word before thinking of word associations, if needed. The LP learners, on the other hand, reported that they did not know the meanings of a few
words on the DVK. In other words, the HP learners seemed to have a relatively larger vocabulary size and a deeper knowledge of word meanings (in terms of synonymy and polysemy) than the LP learners. Generally speaking, since the way to present words on the DVK was in isolation (i.e., without any context provided), they found it difficult to recall other meanings of the stimulus word if they did not originally know them.

In the ‘collocation’ section (i.e., the right-hand side of the column in the example given), if the participants immediately and correctly picked their answers, it was mainly because they had seen or heard those expressions before, especially in an English text or from a native English speaker. They were usually common usages between the two languages (e.g., beautiful face) or widely-used English expressions (e.g., general ideas). Also, the participants examined the collocational possibility by putting words together to see whether the combination made sense to them. I observed that the LP learners tended to use logic in a manner typical to Chinese people; that is, they translated the stimulus word and the associates (or distractors) into Chinese and then checked if the existence of the combination was acceptable in Chinese. Sometimes they were fortunate to guess correctly because there are also the equivalent statements used in Chinese (e.g., conservative estimate and sensitive topic). However, the HP learners seldom did this. Instead, they considered the collocational relationship more in an English fashion. One of them said, ‘I put them together, read them silently in my heart, and see if they sound like English. If they sound clumsy, I will not choose those answers.’ There may be evidence that the way the HP learners dealt with English made their English more native-like. Since Chinese is the participants’ native language, all responses were more or less influenced by Chinese, consciously or unconsciously, while taking the DVK. And yet, the LP learners relied more on translation from Chinese than the HP learners did. Moreover, sometimes when the participants saw the
options in the collocational column, those collocations made the participants retrieve
the fixed expressions they learned before (e.g., bare feet, acute angle, and crude oil).
To some extent, these associates gave the participants a cue to recall other meanings
or usages of the stimulus word.

The verbal reports provided further insight into the word association test taken by
Chinese participants. First of all, the participants' confidence in giving responses
appeared to be a factor. There seemed to be a relationship between language
proficiency and confidence as they gave the verbal reports. The HP learners were
confident in their English ability and tended to give definite responses, using absolute
words like 'obviously related' and 'impossibly associated'. The LP learners, however,
could not really explain, at times, why they selected the answers they did. Very often,
before making their responses, they expressed doubts, uttering sentences such as 'I'm
not sure that...'. Although the HP learners occasionally could not give direct reasons
for their answers, they were willing to explain their techniques or explain how they
made their association network function. They also appeared more focused on and
engaged in the test than the LP learners.

Secondly, the English-Chinese equivalent concept was fairly strong for these
Chinese participants, especially for the LP learners. Their initial reaction to seeing an
English word was to transform it into the translation equivalent in Chinese. It
appeared that their 'English-Chinese equivalent formula' was not very flexible. This
may have resulted from the way most Chinese learners learned vocabulary when they
first studied English; that is, to memorize vocabulary lists. Frequently only one
Chinese equivalent, using concise Chinese characters, is provided, and these concise
Chinese characters are stored in memory as the standard translation equivalent. It is
unlikely for them to use another Chinese expression with the same meaning to replace
the standard one. They, however, overlooked the fact that there exists no precise
equivalent between the two languages and that many English words are polysemous. From the verbal reports I obtained, the participants in general translated those English words into exactly the same Chinese characters. As a native Chinese speaker, I was not surprised at all, because it seems that we learned the same translations for English vocabulary. Yet, the LP learners often knew only one basic meaning of the stimulus word. As reflected in the DVK, they immediately picked up one equivalent associate and quickly denied the other three options, because the corresponding Chinese translations of the other three options were not as close to the standard Chinese equivalent of the stimulus word learned by heart.

Thirdly, with respect to the effect of guessing, both HP and LP learners did guess at responses according to stimulus words that they did not know or only partially knew. The HP learners were willing to guess through whatever strategies or techniques they had. And they were usually quite resourceful—and successful—at seeking the correct responses even though they reported not knowing the stimulus word very well. The LP learners, on the other hand, guessed blindly, without a concrete reason, and most of their responses were incorrect.

Finally, there was behavior specific to each group of learners. One of the LP learners mispronounced vocabulary items on the DVK from time to time while producing his verbal report. Moreover, based on the verbal reports of LP learners, they sometimes became confused about the meanings of similar-sounding words (e.g., _crude-rude_ ) or familiar-looking words (e.g., _flight-fight_ ), and consequently made incorrect responses. The HP learners, on the other hand, came up with phrases that include the stimulus word even before scanning all the options listed beside the stimulus word. Frequently, these phrases involved the associates in the collocational column. This might have resulted because the HP learners have deeper word knowledge of the target words and because their lexical network is more active. Also,
one of the HP learners applied the strategy of prefix analysis he learned when he
prepared for the GRE. Moreover, it was very interesting that both HP learners became
confused about the stimulus word *fertile*. It was not the case that they did not know its
meaning. As a matter of fact, they did know but had never well distinguished *fertile*
from *sterile*, because usually they were given sufficient context to infer its meaning.
One of them even asserted that she would not doubt the meaning of the word *fertile* if
she read it in a text. That also implied the importance of the context for the HP learner.
The following two excerpts illustrate the HP learners’ strategies just mentioned.

23  chronic

<table>
<thead>
<tr>
<th>continuing</th>
<th>local</th>
<th>serious</th>
<th>unplanned</th>
<th>accident</th>
<th>examination</th>
<th>illness</th>
<th>shortage</th>
</tr>
</thead>
</table>

Participant no. 10: [Look at the stimulus word only and point to it with her pen]

*Chronic* ... ahh ... *chronic illness*! That’s how I learned this vocabulary in the very beginning.

Participant no. 9: *Chronic*... a vocabulary item of déjà vu. Mmm ... aha, look at the prefix of this word! *Chron-* is a prefix for something related to ‘time’, so I chose *continuing* in the left-hand side of the column.

28  fertile

<table>
<thead>
<tr>
<th>dark</th>
<th>growing</th>
<th>private</th>
<th>special</th>
<th>business</th>
<th>egg</th>
<th>mind</th>
<th>soil</th>
</tr>
</thead>
</table>

Participant no. 10: *Fertile*. I know this word, but I often confuse it with another word. It means either ‘productive’ or ‘barren.’ Mmm ... I wondered which one it is. But it became so clear when I scanned the options provided in the left-hand side of the column. *Growing* is the one. Yes, *fertile* means ‘productive’. By the way, too bad *fertile* was presented here out of context. If I encounter it in a passage, I can surely tell its meaning from the context right away, without the above process of suspecting.
Compared with the VS, the DVK is more complex, in that test-takers require not only a basic word meaning but also multiple meanings and collocations. According to the Chinese participants' verbal reports, they found the collocation part particularly challenging. They often wondered whether some collocations were acceptable or not, because they are not native English speakers and did not grow up in an English-speaking environment. However, the participants stated that their recent residence in Canada did help them understand authentic English expressions that they had encountered in their daily life, such as natural foods and sensitive skin. Besides this, I found that learning words through their associated phrases seemed to offer a good source of making correct responses.

Generally speaking, the participants did not have very good mastery of deeper vocabulary knowledge, according to results of the verbal reports. Concerning word knowledge, they often emphasized simply the aspect of word meaning. Moreover, many Chinese participants in this study employed similar strategies to deal with isolated words presented on the DVK. Sometimes, the reason they selected the responses to the stimulus word was even identical. Nevertheless, from the results, it was learned that larger vocabulary size makes more words available for associations. That is, the more vocabulary items the learners knew, the wider their lexical networks. It seemed that breadth and depth of their vocabulary knowledge are interdependent and mutually facilitative. Those learners with wider lexical networks were more successful in taking such tests measuring depth of vocabulary knowledge, such as the DVK. In other words, it appears with these participants that having bigger vocabularies plays a role in performance on a vocabulary depth test.
Summary

Both vocabulary size and depth of vocabulary knowledge correlated positively with reading comprehension performance. That is, the wider and the deeper one’s vocabulary knowledge is, the better reading comprehension he or she has. Although vocabulary size appeared to be a stronger predictor of reading comprehension than depth of vocabulary knowledge in this study, it was shown that both dimensions (i.e., breadth and depth of vocabulary knowledge) play very important roles in reading comprehension and that combining vocabulary breadth and depth together has the greatest effect on reading comprehension. Through the statistical analysis of scores on the language tests, a remarkably high correlation between vocabulary breadth and depth was found. In addition, the interdependent and inseparable relationship between vocabulary breadth and depth was observed in the participants’ verbal reports on the DVK and their DVK scores. In other words, the quantitative and qualitative results were consistent. With the mixed-methods approach, the accordance of the results strengthens the reliability of the findings in the study. The next chapter summarizes the main and related findings from the study and puts forth concluding remarks.
CHAPTER 5
CONCLUSIONS

Overview

I summarize the main findings in response to the research question and also mention a few related findings from the study in this closing chapter. Deriving from these findings, there are some pedagogical implications for L2 learners in general and Chinese students in particular. In addition, limitations of the study and directions for future research are put forth.

Summary of the Findings

Main Findings in Response to the Research Question

Before summarizing the main findings in this study, let me restate the research question, the two proposed hypotheses on test scores, and statement of inquiry on the verbal reports. The research question under investigation is:

What is the relationship between reading comprehension, vocabulary size, and depth of vocabulary knowledge for Chinese-speaking ESL university students?

and the hypotheses on test scores are:

1) Test scores on reading comprehension, vocabulary size, and depth of vocabulary knowledge will correlate positively with one another and

2) Depth of vocabulary knowledge will be a more powerful predictor of reading comprehension performance than vocabulary size.

The verbal reports aim to reveal the participants' vocabulary knowledge in detail and to further examine the relationship between vocabulary breadth and depth of the participants. To respond to the research question, several major findings are displayed
by the results with different sources of data.

On the one hand, from the results of the quantitative data (i.e., test scores), the first hypothesis is retained but the second one is rejected. First, test scores on reading comprehension, vocabulary size and depth of vocabulary knowledge did positively correlate with one another; in addition, correlations between them were fairly strong. The association between vocabulary size and depth of vocabulary knowledge was the most remarkable, as vocabulary size appeared to have a stronger relationship with reading comprehension than depth of vocabulary knowledge did. This indicates that interrelationships among reading comprehension, vocabulary size, and depth of vocabulary knowledge do indeed exist. Due to the notably high correlation between vocabulary size and depth, moreover, there appears a strong interaction between the development of breadth (i.e., vocabulary size) and depth of vocabulary and they are most likely interdependent to a certain extent. Second, by comparison, vocabulary size predicted reading comprehension performance more powerfully than depth of vocabulary knowledge did. That is, the learners’ vocabulary breadth seemed to play a critical role in reading comprehension.

On the other hand, the results of the qualitative data (i.e., verbal reports) support those of the quantitative data. It was observed that knowing more vocabulary items does make a difference. Knowing the meaning of a word appears to be the cornerstone of having deeper vocabulary knowledge: the greater the participants’ vocabulary size is, the better they perform on the DVK, which is an in-depth vocabulary test. It leads us to suppose that knowing more vocabulary items makes it possible to broaden their lexical network. Further, the broader lexical network results in a larger vocabulary size which delivers deeper vocabulary knowledge. This deeper vocabulary knowledge then promotes the occurrence of more vocabulary items. That is, the two dimensions of vocabulary knowledge (i.e., breadth and depth) are interdependent and mutually
facilitative. The complementary relationship between these two dimensions is in accord with the results found in the high correlation between breadth and depth of vocabulary knowledge, as shown in the quantitative analysis.

Other Related Findings

First of all, it was learned from the statistical analyses of the test scores that both vocabulary size and depth of vocabulary knowledge explained the variance in reading comprehension significantly. Although neither of them was able to predict a significant amount of the variability in reading comprehension after controlling for the effect of the other, both vocabulary measures (i.e., breadth and depth) made independent further contribution to reading comprehension. The results further suggest that both are useful predictors of reading comprehension performance and that in a predictive sense, a combination of the two correlates better with reading comprehension than either one alone. This principle may be applied to teaching and learning new vocabulary.

Secondly, specific characteristics concerning these Chinese learners’ vocabulary depth were revealed in their verbal reports on each vocabulary item. Very often, they made word associations simply based on arbitrary reasons. The learners mostly indicated that their rationale was generated from their reading experience. They also said that they received direct vocabulary instruction only when a word illustrated a grammatical rule, which was at times incomplete. However, learning along with a chunk of words (e.g., phrases or idioms) showed its effectiveness in the word association task. Generally speaking, without special vocabulary instruction, these Chinese students’ lexical knowledge seems insufficient or inaccurate, which explains their overall performance on the DVK.

Thirdly, the Chinese learners’ emphasis on the word meaning, or a translation
equivalent, was largely observed from their verbal reports. The majority tended to first translate every English vocabulary item they saw. Also, the notion that one word corresponds to one meaning frequently emerged in their verbal report on the DVK, which is a test of depth of vocabulary knowledge. In general, they did not show mastery of deeper lexical knowledge. That is, to some extent, these Chinese participants’ vocabulary knowledge is shallow.

**Implications for L2 Instruction**

Based on the findings of this study, a relationship between vocabulary knowledge and reading comprehension does exist. Increasing vocabulary knowledge is considered essential and beneficial to the learners’ reading comprehension. However, it does not necessarily indicate that words are going to be learned naturally and correctly through reading. According to the verbal reports of the study, lexical knowledge generated from the learners’ personal reading experience is not guaranteed to be accurate and complete. Vocabulary needs to be taught. Once the necessity of vocabulary instruction is accepted, the only real issue is the best manner in which to deliver it. The results of the study suggest the top priority of widening learners’ vocabulary size (i.e., breadth of vocabulary knowledge). Nevertheless, the results also reveal that vocabulary depth seems to tag along with vocabulary size for these Chinese participants. Therefore, building both vocabulary size and depth needs to be specifically encouraged.

Vocabulary needs not only to be taught but also to be taught thoroughly. From this study, we learned that the word list memorization method of learning English vocabulary which misleads Chinese learners as to the notion of one-word-one-meaning correspondence do not work well in establishing the Chinese
students' vocabulary knowledge effectively. It may help expand learners' vocabulary size but not vocabulary depth at all. This one-to-one translation equivalent type of approach might not be the favored one for L2 learners in general and possibly should be reconsidered in L2 vocabulary learning. Vocabulary knowledge has not only breadth but also depth, so in-depth vocabulary practices and activities are recommended. Depth characteristics such as polysemy and collocation in the present study appear to be appropriate elements to integrate into L2 vocabulary instruction. In other words, language teachers need to reinforce vocabulary instruction while they train learners for reading comprehension.

Limitations of the Study

There were limitations of the study. First, the number and language proficiency level of participants may have affected the results of the study. There were only 24 participants. The increase of the number of participants in future research may strengthen the reliability of the results from the present study. Also, participants in the study were advanced ESL learners who studied at English-speaking universities in Canada and read substantially. Investigating other Chinese-speaking ESL learners from different language proficiency levels and in non-English-speaking countries should be considered.

Moreover, the vocabulary tests used in the study measure solely the participants' receptive knowledge. Such comprehension tasks as the VS and DVK in the study usually obtain high test reliability and are easy to be administered and scored because they are objective in the sense that the test items can be scored without involving human raters. However, the absence of the test-takers' productive vocabulary knowledge made this study incomplete to some degree. As Read (1989) noted, the
standard view among researchers on L1 vocabulary acquisition is that receptive vocabulary surpasses productive vocabulary. That is, people can recognize and comprehend a greater number of words than they can actually use. The same phenomenon occurs in L2 vocabulary acquisition to an even greater extent (Nation & Waring, 1997; Schmitt & Meara, 1999). Since production needs greater knowledge than comprehension (Nation, 1990), assessing productive vocabulary knowledge should be included.

**Directions for Future Research**

Over the years, the role of vocabulary knowledge has been explored in L2 studies. Also, an association between vocabulary knowledge and reading comprehension has consistently been found in both L1 and L2 studies. Unsurprisingly, a relationship between vocabulary knowledge and reading comprehension was also found to exist in the present study; however, whether this relationship involves causality was not investigated. For future research regarding the examination of this causal relationship, a study involving the comparison between two different treatments—reading only and reading plus vocabulary enhancement activities—within the same group of learners may be useful.

The present study shows that these Chinese learners did not have very good mastery of deeper word knowledge, as shown by their verbal reports. It indicates that we need to learn more about different types of vocabulary knowledge and how to effectively promote their learning. To learn more about the depth of the learners' lexical knowledge, we need in-depth vocabulary measures. Test design for research or assessment purposes is an area that needs attention. Although the majority of vocabulary tests concern vocabulary size and some of them have been shown to have
an affirmative degree of reliability and validity, there are to date few test formats assessing depth of vocabulary knowledge in an efficient way. That is, there is a gap between research and assessment in vocabulary studies. To take a step further, developing well-designed depth measures is much needed.

In the present study, the Chinese participants were advanced ESL learners in Canada. It is unknown whether the finding of no effect for vocabulary depth over size as a stronger predictor of reading comprehension will remain if the Chinese participants are those who study English in China or Taiwan. It is possible that a strong relationship between vocabulary depth and reading comprehension may be found, in conjunction with a rather weak one between vocabulary size and reading comprehension in beginner or intermediate ESL participants who are not studying in an English-speaking setting. It will be interesting to conduct another study to see whether the results from the present study can be applied to Chinese ESL learners in general or only to advanced ESL learners.

Furthermore, although L2 researchers and language professionals recognize vocabulary as an area of increasing importance and there have been several studies on how L2 learners acquire their vocabulary through reading, as well as on the impact of vocabulary training upon reading or other language skills, unfortunately most of these studies were administered within a short period of individuals' or groups' vocabulary learning, such as one testing session or one school semester. Since learning vocabulary is an on-going process, one-off studies do not necessarily provide an appropriate view of how learners acquire vocabulary knowledge. Conducting longitudinal studies could be a good direction to pursue and it is much more likely to unfold a comprehensive story of what it means to learn a word.
Final Remarks

With a mixed-methods approach, the findings of this empirical study have demonstrated close relationships among vocabulary size, depth of vocabulary knowledge, and reading comprehension. The quantitative data confirmed the important role of vocabulary size (i.e., breadth of knowledge) while the qualitative data suggested that gaining more vocabulary knowledge (i.e., vocabulary depth) might be beneficial. Yet they are not conflicting. One of the most noteworthy findings in the study is that the development of breadth and depth of vocabulary knowledge is complementary, not mutually exclusive. That is to say, while the importance of vocabulary size is highlighted for L2 instruction, strengthening learners' vocabulary depth should not be ignored. The results indicate that combining vocabulary breadth and depth appears to be the most beneficial and that the importance of vocabulary knowledge in reading comprehension cannot be overemphasized. The results of the study shed new light on the relationship between vocabulary knowledge and reading comprehension in the advanced Chinese-speaking ESL learner and also provide direction for future L2 vocabulary research, assessment, and instruction. It is hoped that more studies pursuing other perspectives on the issue of vocabulary breadth versus depth will continue to occur in the future.
References


Educational Testing Service. (1990). *Practice to take the GRE general test-No.8.*
Princeton, NJ: Educational Testing Service (for the Graduate Record Examinations Board).


instructional procedures for teaching word meanings on comprehension and recall. 


Schmitt, N., & Meara, P. (1997). Researching vocabulary through a word knowledge


University of Toronto, Toronto, Ontario, Canada.

Appendix A

Nation’s Vocabulary Levels Test (1990)

Name: ______________________

A Vocabulary Levels Test

This is a vocabulary test. You must choose the right word to go with each meaning. Write the number of that word next to its meaning. Here is an example.

1. business
2. clock ______ part of a house
3. horse ______ animal with four legs
4. pencil ______ something used for writing
5. shoe
6. wall

You answer it the following way.

1. business
2. clock ______ 6____ part of a house
3. horse ______ 3____ animal with four legs
4. pencil ______ 4____ something used for writing
5. shoe
6. wall

Some words are in the test to make it more difficult. You do not have to find a meaning for those words. In the example above, these words are business, clock, shoe.

Try to do every part of the test.
### The 2,000-word level

<table>
<thead>
<tr>
<th>1. original</th>
<th>2. private</th>
<th>3. royal</th>
<th>4. slow</th>
<th>5. sorry</th>
<th>6. total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>_________</td>
<td>complete</td>
<td>_______</td>
<td>not public</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1. apply</th>
<th>2. elect</th>
<th>3. jump</th>
<th>4. manufacture</th>
<th>5. melt</th>
<th>6. threaten</th>
</tr>
</thead>
<tbody>
<tr>
<td>_________</td>
<td>_________</td>
<td>_________</td>
<td>_________</td>
<td>_________</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1. blame</th>
<th>2. hide</th>
<th>3. hit</th>
<th>4. invite</th>
<th>5. pour</th>
<th>6. spoil</th>
</tr>
</thead>
<tbody>
<tr>
<td>_________</td>
<td>_________</td>
<td>_________</td>
<td>_________</td>
<td>_________</td>
<td>_________</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1. accident</th>
<th>2. choice</th>
<th>3. debt</th>
<th>4. fortune</th>
<th>5. pride</th>
<th>6. roar</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>_________</td>
<td>_________</td>
<td>_________</td>
<td>_________</td>
<td>_________</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1. apply</th>
<th>2. elect</th>
<th>3. jump</th>
<th>4. manufacture</th>
<th>5. melt</th>
<th>6. threaten</th>
</tr>
</thead>
<tbody>
<tr>
<td>_________</td>
<td>_________</td>
<td>_________</td>
<td>_________</td>
<td>_________</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1. blame</th>
<th>2. hide</th>
<th>3. hit</th>
<th>4. invite</th>
<th>5. pour</th>
<th>6. spoil</th>
</tr>
</thead>
<tbody>
<tr>
<td>_________</td>
<td>_________</td>
<td>_________</td>
<td>_________</td>
<td>_________</td>
<td>_________</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1. accident</th>
<th>2. choice</th>
<th>3. debt</th>
<th>4. fortune</th>
<th>5. pride</th>
<th>6. roar</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>_________</td>
<td>_________</td>
<td>_________</td>
<td>_________</td>
<td>_________</td>
</tr>
</tbody>
</table>
The 3,000-word level

<table>
<thead>
<tr>
<th>1. administration</th>
<th>1. marble</th>
</tr>
</thead>
</table>
| 2. angel  
managing business and affairs | 2. palm  
inner surface of your hand |
| 3. front  
spirit who serves God | 3. ridge  
excited feeling |
| 4. herd  
group of animals | 4. scheme  
plan |
| 5. mate | 5. statue |
| 6. pond | 6. thrill |

<table>
<thead>
<tr>
<th>1. bench</th>
<th>1. discharge</th>
</tr>
</thead>
</table>
| 2. charity  
part of country | 2. encounter  
use pictures or examples to show the |
| 3. fort  
help to the poor | 3. illustrate  
meaning |
| 4. jar  
long seat | 4. knit  
meet |
| 5. mirror | 5. prevail  
thrown up into air |
| 6. province | 6. toss |

<table>
<thead>
<tr>
<th>1. coach</th>
<th>1. annual</th>
</tr>
</thead>
</table>
| 2. darling  
a thin, flat piece cut from something | 2. blank  
happening once a year |
| 3. echo  
person who is loved very much | 3. brilliant  
certain |
| 4. interior  
sound reflected back to you | 4. concealed  
wild |
| 5. opera | 5. definite |
| 6. slice | 6. savage |
The 5,000-word level

1. alcohol
2. apron __________ cloth worn in front to protect your
3. lure __________ clothes
4. mess __________ stage of development
5. phase __________ state of untidiness or dirtiness
6. plank

1. bruise
2. exile __________ agreement using property as security
3. ledge __________ for a debt
4. mortgage __________ narrow shelf
5. shovel __________ dark place on your body caused by
6. switch __________ hitting

1. circus
2. jungle __________ speech given by a priest in a church
3. nomination __________ seat without a back or arms
4. sermon __________ musical instrument
5. stool
6. trumpet

1. blend
2. devise __________ hold tightly in your arms
3. embroider __________ plan or invent
4. hug __________ mix
5. imply
6. paste

1. apparatus
2. compliment __________ set of instrument or machinery
3. revenue __________ money received by the government
4. scrap __________ expression of admiration
5. tile
6. ward

1. desolate
2. fragrant __________ good for your health
3. gloomy __________ sweet-smelling
4. profound __________ dark or sad
5. radical
6. wholesome
The University Word List level

1. affluence
2. axis
3. episode
4. innovation
5. precision
6. tissue

1. deficiency
2. magnitude
3. oscillation
4. prestige
5. sanction
6. specification

1. anonymous
2. indigenous
3. maternal
4. minimum
5. nutrient
6. modification

1. elementary
2. negative
3. static
4. random
5. reluctant
6. ultimate

1. coincidence
2. coordinate
3. expel
4. frustrate
5. supplement
6. transfer
The 10,000-word level

1. acquiesce
2. contaminate  _______ work at something without serious intentions
3. crease  _______ intentions
4. dabble  _______ accept without protest
5. rape  _______ make a fold on cloth or paper
6. squint

1. anterior
2. concave  _______ small and weak
3. interminable  _______ easily changing
4. puny  _______ endless
5. volatile
6. wicker

1. blaspheme
2. endorse  _______ give care and food to
3. nurture  _______ speak badly about God
4. overhaul  _______ slip or slide
5. skid
6. straggle

1. dregs
2. flurry  _______ worst and most useless parts of anything
3. hostage
4. jumble  _______ natural liquid present in the mouth
5. saliva  _______ confused mixture
6. truce

1. auxiliary
2. candid  _______ full of self-important
3. dubious  _______ helping, adding support
4. morose  _______ bad-tempered
5. pompous
6. temporal

1. auspices
2. casualty  _______ being away from other people
3. froth  _______ someone killed or injured
4. haunch  _______ noisy and happy celebration
5. revelry
6. seclusion
Appendix B

Read’s Word Associates Test (unpublished), Version 4.0, with Modified Test Instructions

Name: __________________________

WORD ASSOCIATES TEST

This is a test of how well you know the meaning of adjectives that are commonly used in English. Each item looks like this:

A sudden

<table>
<thead>
<tr>
<th>beautiful</th>
<th>quick</th>
<th>surprising</th>
<th>change</th>
<th>doctor</th>
<th>noise</th>
<th>school</th>
</tr>
</thead>
<tbody>
<tr>
<td>thirsty</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There are eight words in the two boxes (left & right boxes). 

The words here on the left side may help to explain the meaning of "sudden".
The words here on the right side are nouns that may come after "sudden" in a phrase or a sentence.

"Sudden" means "happening quickly and unexpectedly", so the correct answers on the left side are "quick" and "surprising".
We do not normally say "a sudden doctor" or "a sudden school", but we often say "a sudden change" and "a sudden noise", so "change" and "noise" are the correct answers on this side.

From the two boxes, select words that you think are relevant to the stimulus word (i.e., ‘sudden’ in this example), according to the criteria mentioned above. Put a circle around each word that you choose.

Please circle the answers like this:

A sudden

<table>
<thead>
<tr>
<th>beautiful</th>
<th>quick</th>
<th>surprising</th>
<th>change</th>
<th>doctor</th>
<th>noise</th>
<th>school</th>
</tr>
</thead>
<tbody>
<tr>
<td>thirsty</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: In this example, there are two correct answers on the left and two on the right, but this is just an example. Do NOT assume there is a consistent number of correct answers on the left or on the right. Just remember: try to find as many related words as you can.
1. beautiful
   - enjoyable, expensive, free, loud
   - education, face, music, weather

2. bright
   - clever, famous, happy, shining
   - colour, hand, poem, taste

3. calm
   - open, quiet, smooth, tired
   - cloth, day, light, person

4. natural
   - expected, helpful, real, short
   - foods, neighbours, parents, songs

5. fresh
   - another, cool, easy, raw
   - cotton, heat, language, water

6. general
   - closed, different, usual, whole
   - country, idea, reader, street

7. bare
   - empty, heavy, uncovered, useful
   - cupboard, feet, school, tool

8. acute
   - hidden, often, rich, sharp
   - angle, hearing, illness, stones

9. common
   - complete, light, ordinary, shared
   - boundary, circle, name, party

10. complex
    - angry, difficult, necessary, sudden
    - argument, passengers, patterns, problem
11 broad

full moving quiet wide night river shoulders smile

12 conscious

awake healthy knowing face decision effort student laughing

13 convenient

easy fresh near suitable experience sound time vegetable

14 dense

crowded hot noisy thick forest handle smoke weather

15 curious

helpful interested missing strange accident child computer steel

16 distinct

clear famous separate true advantage meanings news parents

17 dull

cloudy loud nice secret colour knife place rock

18 direct

honest main straight wide fence flight heat river

19 favourable

helpful legal possible positive habit response teacher weather

20 secure

confident enjoyable fixed safe game job meal visitor
<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>tight</td>
<td>close</td>
<td>rough</td>
<td>uncomfortable</td>
<td>wet</td>
<td>bend</td>
<td>pants</td>
<td>surface</td>
</tr>
<tr>
<td>22</td>
<td>violent</td>
<td>expected</td>
<td>smelly</td>
<td>strong</td>
<td>unlucky</td>
<td>anger</td>
<td>death</td>
<td>rubbish</td>
</tr>
<tr>
<td>23</td>
<td>chronic</td>
<td>continuing</td>
<td>local</td>
<td>serious</td>
<td>unplanned</td>
<td>accident</td>
<td>examination</td>
<td>illness</td>
</tr>
<tr>
<td>24</td>
<td>compact</td>
<td>effective</td>
<td>small</td>
<td>solid</td>
<td>useful</td>
<td>group</td>
<td>kitchen</td>
<td>medicine</td>
</tr>
<tr>
<td>25</td>
<td>crude</td>
<td>clever</td>
<td>fair</td>
<td>rough</td>
<td>valuable</td>
<td>behaviour</td>
<td>drawing</td>
<td>oil</td>
</tr>
<tr>
<td>26</td>
<td>domestic</td>
<td>home</td>
<td>national</td>
<td>regular</td>
<td>smooth</td>
<td>animal</td>
<td>movement</td>
<td>policy</td>
</tr>
<tr>
<td>27</td>
<td>profound</td>
<td>bright</td>
<td>deep</td>
<td>exact</td>
<td>great</td>
<td>effect</td>
<td>machine</td>
<td>taste</td>
</tr>
<tr>
<td>28</td>
<td>fertile</td>
<td>dark</td>
<td>growing</td>
<td>private</td>
<td>special</td>
<td>business</td>
<td>egg</td>
<td>mind</td>
</tr>
<tr>
<td>29</td>
<td>formal</td>
<td>fast</td>
<td>loud</td>
<td>organised</td>
<td>serious</td>
<td>bomb</td>
<td>education</td>
<td>growth</td>
</tr>
<tr>
<td>30</td>
<td>independent</td>
<td>changed</td>
<td>equal</td>
<td>important</td>
<td>separate</td>
<td>child</td>
<td>country</td>
<td>ideas</td>
</tr>
</tbody>
</table>
### 31 original
- careful
- closed
- first
- proud
- condition
- mind
- plan
- sister

### 32 sensitive
- feeling
- interesting
- sharp
- thick
- clothes
- instrument
- skin
- topic

### 33 professional
- paid
- public
- regular
- religious
- advice
- manner
- musician
- transport

### 34 critical
- clear
- dangerous
- important
- rough
- festival
- illness
- time
- water

### 35 synthetic
- artificial
- electronic
- expensive
- simple
- drug
- meal
- radio
- sound

### 36 liberal
- free
- moderate
- plenty
- valuable
- crops
- furniture
- parents
- transport

### 37 dramatic
- exciting
- official
- surprising
- worried
- adventure
- change
- patient
- salary

### 38 conservative
- hopeful
- safe
- together
- traditional
- clothes
- estimate
- meeting
- signal

### 39 coherent
- clear
- normal
- recent
- together
- crime
- health
- speech
- theory

### 40 ample
- heavy
- large
- plentiful
- windy
- amount
- climate
- feelings
- time
Appendix C

Background Questionnaire [Chinese version]
(Note: This version is of traditional Chinese characters. For the version of simplified Chinese characters, please contact the author)

問卷調查

此為一針對英文為外國語學習者的背景資料調查。請完成下列所有問題以協助此研究計劃。謝謝！

個人背景資料

1. 姓名：__________________________
2. 性別： □男 □女
3. 國籍：__________________________
4. 年齡：__________________________
5. 教育程度：

   □□□□□ 大學部，之 □第一 □第二 □第三 □第四年

   □□□□□ 研究生，之 □碩士 □博士 學位

6. 學校：_____________________________________________________________________

   學院/系別：_________________________________________________________________

7. 停留在加拿大已滿： __________年 __________月

8. 學習英語已經：

   □ 1~5年 □ 6~10年 □ 11~15年 □ 16~20年 □ 21年以上

9. 您曾考過 TOEFL/IETLS/GRE 嗎？(假如您有，請寫下您的分數及年份)

   □ 有，我的 TOEFL 分數是 ________ 在 __________年。

   □ 有，我的 IELTS 分數是 ________ 在 __________年。

   □ 有，我 GRE 語文部分的分數是 ________ 在 __________年。

   □ 沒有，我從沒考過上述之任何測驗。

10. 聯絡資料—

    電話：（_________）________________________________________________________

    電子信箱：________________ ______@_____________________________
Appendix D

Background Questionnaire [English version]

QUESTIONNAIRE

This is a background questionnaire for learners of English as a foreign language. Please complete the following questions to help this research project. Thank you!

Background Information

1. Name: ____________________________  2. Sex: □ Male  □ Female

3. Name of home country: ______________  4. Age: ______________________

5. Educational Level:

□ an undergraduate student, at the □ first  □ second  □ third  □ fourth  year

□ a graduate student, pursuing a □ Master’s □ Ph. D. degree

6. Name of university: __________________________

Faculty / Department / Program: __________________________

7. Length of stay in Canada: __________ year(s) __________ month(s)

8. I have been studying English for

□ 1–5  □ 6–10  □ 11–15  □ 16–20  □ more than 21 years.

9. Have you ever taken TOEFL and/or GRE? (If you have, please write down your score(s) and when you obtained it if you remember.)

□ Yes. My TOEFL score was ___________ in Year __________; and/or

□ Yes. My IELTS score was ___________ in Year __________; and/or

□ Yes. My GRE score (Verbal Section) was ___________ in Year __________.

□ No. I have not taken any tests mentioned above.

10. Contact information:

Telephone number: (_______) __________________________

Email address: ___________________________ @ _______________________
Appendix F

Informed Consent Form [Chinese version]
(Note: This version is of traditional Chinese characters. For the version of simplified Chinese characters, please contact the author)

參與研究同意書

此為一聲明：我同意參與 McGill 大學第二語言教育學系研究生黃杏妃所策劃名為 “英語字彙知識的廣度及深度：何者在母語為中文的大學生之學術閱讀能力中扮演重要角色?” 的研究計劃。

目的：當學習一語言時，字彙常被視為最基本的要素。因此，本研究論文目的在於探索母語為中文的大學生之英語字彙知識所扮演的角色。本研究即策劃回答下列問題：對於母語為中文的大學生，英語之閱讀理解力、字彙量、字彙知識的深度及三者間的關係為何？

程序：參與者之身分將會保持匿名。當將來研究者提及其資料時，參與者將以代碼的形式來被識別。所有在研究期間所收集的資料將是保密且僅用於此研究目的。

參與情況：參與者將需完成三項語言測驗（其一為閱讀理解力測驗；另二項為字彙測驗）及一後續問卷調查。在語言測驗中，將呈現若干英語文章或字彙，而參與者則需回答在測驗中的問題來展現其理解程度。此外，將有四位參與者對於其一字彙測驗給予額外的口頭報告。問卷調查則是以自己的觀點或經驗來回答字彙學習之相關問題。時間是唯一的不便之處。本研究內容並無涉及任何危險，而可能對於母語為中文的大學生的英語字彙知識角色之了解有所貢獻。

請詳細閱讀以下:
1. 我了解此研究論文的目的並且知曉此研究計劃所之風險、利益及不便之處。
2. 我了解我能在任何時候退出此研究計劃並無任何懲處及偏見。
3. 我了解此研究將不會影響我的成績或我工作的表現。
4. 我了解在此研究計劃期間收集之資料將被保持為機密。
5. 我了解此資料預期的用處（尤其是在出版方面）及其資料結果傳播的方式。

我已細查上述之聲明，且了解在此協議提及之參與情況。我同意並自願參與此研究計劃。

姓名 (請書寫工整) __________________________________________

簽名 __________________________________________ 日期 _____________
Appendix G

Informed Consent Form [English version]

INFORMED CONSENT FORM TO PARTICIPATE IN RESEARCH

This is to state that I agree to participate in the research project entitled:

*Breadth and depth of English vocabulary knowledge: Which really matters in the academic reading performance of Chinese university students?*

as conducted by Hsing-fei Huang, MA student in the Second Language Education program at McGill University.

• **Purpose:** Vocabulary is usually regarded as the most basic unit when learning a language. This study attempts to explore the role of vocabulary knowledge of Chinese university students and is devised to answer the following question: What are the relationships between reading comprehension, vocabulary size and depth of vocabulary knowledge for Chinese university students?

• **Procedures:** Participants’ identities will be kept anonymous and participants will be identified by number when the researcher refers to them. All the data collected during this study will be confidential and will be used for research purposes only.

• **Conditions of participation:** Participants will be asked to complete three language tests—one reading comprehension test and two vocabulary tests—and a follow-up questionnaire. They will be presented with a few English passages or vocabulary items and will be asked to show how well they know them by answering questions in the three tests. In addition, four of the participants will be asked to produce extra verbal reports on one of the vocabulary tests. The questionnaire concerns what relates to vocabulary knowledge in a self-report fashion. Time is the only inconvenience. There are no risks, and the benefits involve the possibility of contributing to our understanding of the role of vocabulary knowledge of Chinese university students.

Please carefully read the following:

1. I understand the purpose of this study and know about the risks, benefits and inconveniences that this research project entails.
2. I understand that I am free to withdraw at anytime from the study without any penalty or prejudice.
3. I understand that this research will not affect my grades or evaluation of my work.
4. I understand how confidentiality will be maintained during this research project.
5. I understand the anticipated uses of data, especially with respect to publication, communication and dissemination of results.

I have studied the above and understand my participation in this agreement. I freely consent and voluntarily agree to participate in this study.

Name (please print) ________________________

Signature __________________________ Date __________________