UNDERSTANDING READING COMPREHENSION IN ENGLISH IMMERSION STUDENTS IN CHINA

by

Miao Li

A thesis submitted to the Faculty of Education
In conformity with the requirements for
the degree of Doctor of Philosophy

Queen’s University
Kingston, Ontario, Canada
(December, 2012)

Copyright ©Miao Li, 2012
Abstract

Compared to research on English as a First Language (EL1) reading comprehension, there is a dearth of studies investigating English as a Second Language (ESL) reading comprehension. This dissertation reports findings from two studies which examined the cognitive processes underlying English reading comprehension in Chinese ESL students. Two-hundred and forty six Grade 8 students in an English immersion program in a junior middle school in China were administered a battery of reading-related and reading comprehension tests. The first study investigated the cognitive predictors of English reading comprehension in these students. It tested whether the Simple View of Reading model (Gough & Tunmer, 1986) applied to Chinese ESL students and whether the effects of breadth and depth of vocabulary contributed to different levels of reading comprehension. The results showed that the Simple View of Reading model did not predict Chinese English-immersion students’ reading comprehension well. Vocabulary as well as inference and strategy added further variance to the general understanding of text. Vocabulary breadth and depth explained unique variance in reading comprehension, with vocabulary breadth contributing more to general understanding of text and vocabulary depth to deep processing of text.

The second study explored characteristics of reading comprehension difficulties among Chinese English-immersion students. Three groups of comprehenders matched on age, nonverbal intelligence, and word reading speed were identified from the 246 Grade 8 Chinese English-immersion students: 33 unexpected poor comprehenders, 28 expected average comprehenders, and 30 unexpected good comprehenders. The three groups differed in vocabulary and higher-level processes. Vocabulary breadth and depth distinguished between the unexpected poor comprehenders and the expected average comprehenders. Inference, strategy, listening comprehension, summary writing, and morphological awareness distinguished between the expected average comprehenders and the unexpected good comprehenders. The findings suggest
that vocabulary is the main source of reading comprehension difficulties in these unexpected poor comprehenders. The advantage of the unexpected good comprehenders group was primarily due to discourse comprehension and strategic processes, and is only possible with high language proficiency. Taken together, both vocabulary and higher-level skills are essential to successful reading comprehension in Chinese English-immersion students. The higher-level skills can be used more efficiently to help these students’ reading comprehension when they have acquired adequate vocabulary knowledge.
Acknowledgements

This dissertation would not be possible without the support and help from the following people in my life.

First of all, I would like to express my sincere appreciation to my supervisor Prof. John Kirby. In my past six years at Queen’s University, two years for a master’s program and four years for a PhD, John’s intelligence, humour, and wisdom left strong impression on me. Most important, he offered expertise in the Psychology of Reading and insightful guidance to me through the whole research process. It was him who taught me how to become a scientist. I was very fortunate to be his student and work with him. Because of him, my dream comes true!

My strong appreciation also goes out to my committee members: Profs. Liying Cheng and Lesly Wade-Woolley, whom I also have been working with for six years. Their understanding and beliefs in me encouraged me to go forward. They were very helpful in providing me with valuable advice when editing my work. Their assistance was always available for me whenever I need. Thank you very much, Liying and Lesly!

Crucial support to my research also comes from China. I am indebted to Dr. Wei Zhao, a professor in the Faculty of Education at Shannxi Normal University in China, and a director of China-Canada-USA English Immersion (CCUEI) collaborative Program. She arranged the English-immersion school for me to conduct my research in China and even assigned her graduate students to help me with my data collection. A special thank you to Mengmeng Liu, one of the graduate student assistants from Shannxi Normal University, who aided me through the whole data collection process. Thanks also go to Ms. Xianghong Zhou, who is in charge of the English-immersion program in the school and to teachers and administrators at the school where I collected my data. They were very friendly and supportive to my research. A big thank you goes to 248 participants in this research, whose names I cannot put forward. Their participation made this research succeed!
I would like to thank my mentors who have been influential in my academic journey, Prof. Hongan Li and my big sister Dr. Miao Sun for their continuous support for this degree after I obtained my master’s degree. Their words of encouragement meant a great deal to me. As role models, they guided me through my pursuit of the PhD degree.

I would also like to extend a thank you to educators, administrators, and my fellow graduate students at Queen’s University, who lent me generous help and support in both academic and non-academic areas. The process of my life would not be the same without their encouragement and inspiration.

My deep gratitude goes to my family members. A special thank you to my beloved husband, Dawei Wang, who always told me I could do anything I set my mind to. I am grateful for the many times he listened to my complaints and helped me through the hard times from his experience of being a doctoral student in the field of physics. He always has faith in me that I am able to achieve my goal. Dawei, thank you for believing in me. I made it!

My grandfather Jun Miao, who is 94 years old, cannot remember many of his grandchildren but still remembers me and that I have been pursuing my PhD degree. Now, I can proudly tell him that his granddaughter has received this degree and become the first one to get this degree among his third generation.

Finally, this dissertation is dedicated to my father Xiting Li and mother Yonghua Miao. For these years, they have loved me, believed in me, encouraged me, and supported me unconditionally. Without them, I would not be where I am today. They must be very proud of me for making this accomplishment in both their families.
# Table of Contents

Abstract.................................................................................................................................ii
Acknowledgements...............................................................................................................iv
List of Figures ......................................................................................................................... ix
List of Tables .......................................................................................................................... x

Chapter 1 General Introduction ......................................................................................... 1
  Purpose and Structure of the Dissertation................................................................. 2
  An Overview of Key Concepts in Reading Comprehension ................................... 4
    What is Reading Comprehension? .............................................................................. 4
    Measurement of Reading Comprehension ............................................................... 9
  Similarities and Differences in L1 and L2 Reading ..................................................... 12
    Similarities in L1 and L2 Reading ........................................................................... 12
    Differences in L1 and L2 Reading ........................................................................... 13
  English Immersion Programs in China ...................................................................... 15
  Rationale for the Dissertation ....................................................................................... 19
  Organization of the Dissertation ................................................................................... 20

Chapter 2 Predictors of Reading Comprehension in Chinese Students Learning English 21
  Predictors of Reading Comprehension in EL1 and ESL Children .............................. 25
    Decoding and Listening Comprehension ................................................................. 26
    Vocabulary .................................................................................................................... 28
    Inference and Strategy ............................................................................................... 31
    Reading Comprehension ......................................................................................... 33
  The Present Study ........................................................................................................... 34
  Method ............................................................................................................................. 36
    Participants .................................................................................................................. 36
    Measures ..................................................................................................................... 36
      Nonverbal Ability. ..................................................................................................... 36
      English Decoding .................................................................................................... 37
      English Listening Comprehension ......................................................................... 37
      Gates-MacGinitie vocabulary subtest. ................................................................. 37
      Word Definitions. .................................................................................................... 37
      Multiple-meaning Vocabulary. ............................................................................. 38
      Morphological Awareness. .................................................................................... 38
Chapter 3 Unexpected Poor Comprehenders among Adolescent ESL Students

Method

The Present Study

Reading Comprehension Difficulties in ESL Children

The Present Study

Results

Descriptive Statistics

Correlations

Exploratory Factor Analysis

Prediction of Reading Comprehension

Discussion

How Well Do Decoding and Linguistic Comprehension Predict English Reading Comprehension?

Should the SVR Model Include Inference and Strategy?

What are the Effects of Breadth and Depth of Vocabulary on English Reading Comprehension?

What are the Effects of Breadth and Depth of Vocabulary on English Word Reading?

Limitations

Conclusions and Implications

Chapter 3 Unexpected Poor Comprehenders among Adolescent ESL Students

Unexpected Poor Comprehenders in EL1 Children

Summary Writing

Chinese Reading Comprehension

English Reading Comprehension

Sentence Description Writing

Procedure

Results

Descriptive Statistics

Correlations

Exploratory Factor Analysis

Prediction of Reading Comprehension

Discussion

How Well Do Decoding and Linguistic Comprehension Predict English Reading Comprehension?

Should the SVR Model Include Inference and Strategy?

What are the Effects of Breadth and Depth of Vocabulary on English Reading Comprehension?

What are the Effects of Breadth and Depth of Vocabulary on English Word Reading?

Limitations

Conclusions and Implications

Chapter 3 Unexpected Poor Comprehenders among Adolescent ESL Students

Unexpected Poor Comprehenders in EL1 Children

Reading Comprehension Difficulties in ESL Children

The Present Study

Method

Participants

Measures

Nonverbal Ability

Word Reading

Vocabulary Breadth

Word Definitions

Multiple-meaning Vocabulary

Morphological Awareness

Listening Comprehension

Inference and Strategy

Inference and Strategy

Chinese Reading Comprehension

English Reading Comprehension

Summary Writing
Summary Writing………………………………………………………………………………….77
Reading Comprehension………………………………………………………………………..78
Procedure …………………………………………………………………………………………78
Results……………………………………………………………………………………………..79
Descriptive Statistics…………………………………………………………………………….79
Group Differences in Vocabulary and Higher Level Processes………………………81
Discussion…………………………………………………………………………………………82
Vocabulary Skills Distinguish Unexpected Poor Comprehenders from Expected Average
Comprehenders ……………………………………………………………………………………..82
Higher-level Skills Distinguish Expected Average Comprehenders from Unexpected Good
Comprehenders ……………………………………………………………………………………..84
Conclusions, Limitations, and Implications………………………………………………85
Chapter 4 General Discussion………………………………………………………………88
A Picture of English Reading Comprehension in Chinese Students…………………88
Critical Issues around ESL Reading Comprehension………………………………………90
   The Crucial Role of Vocabulary in ESL Reading Comprehension …………………..90
   Assessment of Vocabulary Depth……………………………………………………………92
   Assessment of Reading Comprehension…………………………………………………93
Higher-level Processes in ESL Reading Comprehension ………………………………94
Implications for Teaching and Future Research………………………………………..96
   Vocabulary Teaching………………………………………………………………………..96
   Teaching of Higher-level Skills……………………………………………………………..98
   ESL Reading Comprehension Model ……………………………………………………..99
Concluding Remarks………………………………………………………………………….100
References…………………………………………………………………………………………101
Appendix A Research Ethics Board Approval………………………………………….132
Appendix B Letter of Information ……………………………………………………………133
Appendix C Consent Form ……………………………………………………………………135
Appendix D English Word Definition Test and Scoring Guidelines…………………..137
Appendix E Multiple-meaning Vocabulary Test …………………………………………139
Appendix F Morphological Awareness Test and Scoring . .……………………………141
Appendix G Inference and Strategy Tests …………………………………………………143
Appendix H Summary Writing Test …………………………………………………………147
Appendix I Summary Writing Scoring……………………………………………………….148
List of Figures

Figure 1. The scatterplot for the distribution of unexpected poor comprehenders (UPCs), expected average comprehenders (EACs), and unexpected good comprehenders (UGCs), as well as those who were not selected for analyses (NSC), in the regression predicting reading comprehension from age, nonverbal intelligence, and word read efficiency.............................. 72
List of Tables

Table 1 Descriptive Statistics for All Measures ................................................................. 43
Table 2 Correlations between All Measures ................................................................. 45
Table 3 Principal Axis Factor Analysis with Direct Oblimin Rotation of Vocabulary Measures .46
Table 4 Summary of Hierarchical Regression Analyses Predicting Different Levels of Two Reading Comprehension Measures from the Simple View of Reading Model (N = 246) ........... 48
Table 5 Summary of Hierarchical Regression Analyses Predicting Different Levels of Two English Reading Comprehension Measures (N = 246) ........................................................................ 50
Table 6 Summary of Hierarchical Regression Analyses Predicting English Word Reading from Breadth and Depth of Vocabulary .............................................................................. 51
Table 7 Means, Standard Deviations, and Analysis of Variance Results for Unexpected poor comprehenders, Expected average comprehenders, and Unexpected good comprehenders on Age, Nonverbal Intelligence, Word Reading Efficiency, and Reading Comprehension ...................... 74
Table 8 Means, Standard Deviations, and Analysis of Variance Results for Unexpected poor comprehenders, Expected average comprehenders, and Unexpected good comprehenders on Vocabulary Measures and Higher Level Processes .............................................................................. 80
Chapter 1

General Introduction

Reading is regarded as a necessary skill for academic, economic, and social success. It has two major elements: word reading and reading comprehension. Word reading is the ability to recognize individual words. Reading comprehension is the ability to draw meaning from the text. Comparatively speaking, reading comprehension is more complex than word reading as it involves a broader range of cognitive processes and thus it is viewed as the “essence of reading” (Durkin, 1993). It is not only essential to academic learning in all subjects but also to lifelong learning.

Reading comprehension in learners whose English is their first language (EL1) has been extensively researched. The National Reading Panel (2000) reported that lower-level skills (e.g., phonological awareness, decoding, fluency, and vocabulary knowledge) are important to the ability to read English successfully. At the same time, it has become clear that higher-level reading skills (e.g., inferencing skills, comprehension monitoring, integration of text information) are also essential for success in English reading comprehension (Oakhill, Cain, & Bryant, 2003; Snow, 2002; van den Broek et al., 2005). Compared to research on EL1 reading comprehension, less is known about English as a Second Language (ESL) or English as a Foreign Language (EFL) reading comprehension.

In China, there are a great number of students learning English. Mandatory English instruction in schools begins in Grade 3, and continues until post-secondary schools. Furthermore, English immersion programs have started in several major cities in China in order to bring in an innovative English teaching method which focuses on teaching some non-language subjects in
English besides English Language Arts. Traditionally, students learning English in China are considered as learning English as a Foreign Language (EFL) because their English input is much less than that of ESL students taught by English native speakers in English-speaking countries. However, as English continues to spread, more and more English input has been brought to English learning in Chinese students, and the Chinese students in English immersion programs receive even more English input than other students. Therefore, I will refer to the Chinese English-immersion students as ESL students in my dissertation. This is also consistent with the literature, which mainly refers to ESL. Also, ESL and EFL are basically the same: both refer to the use or study of English by speakers with different native languages (Reid, 1995).

ESL reading ability is in great demand as many people need to read in English as an L2 at high levels of proficiency to achieve personal, occupational, and professional goals. A better understanding of learning to read in English as an L2 can provide us with a picture of ESL reading processes and help ESL learners read more accurately and efficiently.

**Purpose and Structure of the Dissertation**

This dissertation aims to understand the cognitive processes underlying English reading comprehension in Chinese ESL students who are in English immersion programs. It is guided by EL1 reading theories because EL1 reading comprehension research has made impressive progress in understanding the English reading process. Therefore it makes good sense for ESL researchers to learn from EL1 reading research about the nature of English reading comprehension. Also, reading comprehension theories in ESL are not yet well developed. Most previous ESL reading studies focused on top-down inferential and schema-driven explanations for reading comprehension rather than decoding and lower-level skills (e.g., Barnett, 1988; Bernhardt, 1991; Carrell, 1984; Coady, 1979; Goodman, 1967). They were limited by the theoretical frameworks
that were then in use. Their approaches are insufficient for an integrative understanding of the development of ESL reading comprehension. Therefore, I intend to test whether some comprehensive EL1 reading models can be extended to ESL children reading English. This dissertation uses a multiple manuscript approach, consisting of two separate but inter-related studies, each providing a different perspective on of the cognitive processes of English reading comprehension for Chinese English-immersion students.

The first study examines the cognitive predictors influencing English reading comprehension in Chinese English-immersion students. Reading research has identified powerful predictors of reading comprehension in EL1 students. According to the Simple View of Reading (SVR) model, reading comprehension can be predicted by decoding and listening comprehension (Gough & Tunmer, 1986). How well these two predictors account for reading comprehension in ESL students needs investigating: ESL students learning to read in L2 use a two-language processing system (L1 and L2), which is different from EL1 students who depend on only one language processing system. Research on the SVR model has suggested that these two predictors are not adequate to explain the complexities of reading comprehension in L1 and that other factors such as inference and strategy need to be considered (Cain, Oakhill, & Lemmon, 2004; Kirby & Savage, 2008). Furthermore, studies have shown that vocabulary has proved to be a powerful predictor of reading comprehension in ESL students (August, Carlo, Dressler, & Snow, 2005). However, the nature of vocabulary is not well understood. There are different dimensions of vocabulary, e.g., breadth and depth, but usually only one dimension (breadth) is researched in ESL reading comprehension. Therefore, the first study explores whether the SVR model applies to Chinese English-immersion students and how different dimensions of vocabulary (i.e., breadth and depth of vocabulary) contribute to reading comprehension.
Research has examined ESL students with deficits in decoding skills, which in turn, lead to difficulties in reading comprehension (Lesaux, Geva, Koda, Siegel, & Shanahan, 2008). However, there are some students who have adequate decoding skills but who still have trouble comprehending a text (e.g., Cain & Oakhill, 1999). This group of students has been neglected in ESL reading research. Thus, the second study investigates the sources of difficulties for those Chinese English-immersion students who experience difficulties in English reading comprehension in spite of adequate decoding skills.

The remainder of this general introduction chapter briefly reviews the literature regarding the key concepts of reading comprehension used to guide my dissertation. Then the similarities and differences between L1 and L2 reading are addressed because the relationship between these two should influence our understanding of L2 reading comprehension. The context of this dissertation is then clearly presented to help the readers understand the setting better. In addition, a rationale is provided to explain why it is important to conduct this dissertation, followed by the organization of the dissertation.

**An Overview of Key Concepts in Reading Comprehension**

**What is Reading Comprehension?**

Reading comprehension is the ability to understand information from texts and interpret this information appropriately (Grabe & Stoller, 2011). However, this definition is insufficient to understand the true nature of reading comprehension. It does not really tell us how we comprehend a text. In fact, reading comprehension is a highly complex cognitive process which involves the coordination of many different processing skills.

Reading comprehension can be seen in distinct levels of understanding of text. Kintsch’s (1998) *Construction-Integration* (CI) model outlined three distinct levels of memory
representations that are formed during reading comprehension. When reading a text, the first level consists of a surface representation of the text in which a sequence of propositions (the smallest units of meaning within individual sentences and clauses) is constructed directly from words and phrases in the text. At the second level, the propositions generated are connected with previous and subsequent propositions in a network, representing the local meaning relationships (microstructure), and with more important concepts in the network, representing the more global relationships in the text (macrostructure). Together the microstructure and macrostructure form the textbase. The textbase involves deeper processing than the surface representation because it links more information in the text and indicates relative importance. These first two levels are the construction phase of the CI model. The final level is an integration process in which the constructed textbase is integrated with the reader’s prior knowledge to form a coherent mental representation of the text (situation model). The situation model is a more elaborated level; it involves deeper processing than the textbase because it incorporates information from long-term memory (prior knowledge) to make sense of the text and because it is more related to meaning.

It is important to note that Kintsch’s CI model assumes efficient word reading, which is the lowest level of surface processing. This assumption may apply to the word reading of skilled adults, but is less likely to be applicable in younger, or less skilled, or ESL readers. Therefore, in a population where proficiency in this skill cannot be assumed, it is important that its influence be considered.

Information can be processed at different levels, such as words, phrases, ideas/propositions, main ideas, or themes (Kirby, 2007). Comprehension at the phrase or proposition level is literal comprehension (e.g., textbase), which is a relatively shallow understanding of what the text states; further processing of ideas or propositions to main ideas
and themes is generally inferential comprehension (e.g., situation model), which is a deep understanding of what the text states (Kintsch, 1998). There are two types of processing: bottom-up and top-down. In bottom-up processing, several words are formed into one phrase, several phrases are formed into one idea or proposition, etc. In top-down processing, higher-level information helps the reader identify lower-level information to build a framework for comprehension. Efficient functioning of lower-level information allows higher-level processing to function more efficiently (Perfetti, 1985), but higher-level knowledge and goals are required to guide the lower-level processing. Both bottom-up and top-down processing occur interactively in reading comprehension and both are important.

It should be noted that knowledge is represented in hierarchical levels of representation, in which higher levels represent general or abstract ideas while lower levels represent specific details (Cohen, 2000). Craik and Lockhart (1972) proposed levels of processing, which suggests a continuum of processing about the qualitative nature of processing from shallow to deep, leading to more durable memories (Craik, 2002). The levels of processing proposal concerns the degree to which each level of processing has been done during encoding, with shallow processing representing visual word processing and depth of processing reflecting enriched meaning and leading to better memory. When readers comprehend a text, they progress from relatively shallow levels to deeper levels and deeper processing leads to a good quality understanding and learning from the text (Kirby, Cain, & White, 2012).

The lower-level reading skills are those which can help word reading, such as phonological awareness, decoding, fluency, and vocabulary knowledge (National Reading Panel, 2000). The higher-level reading comprehension skills are those which can aid the construction of the meaning-based representation of the text, such as inferencing skills, strategies, comprehension
monitoring, and integration of text information (Oakhill et al., 2003; Snow, 2002; van den Broek et al., 2005). It should be noted that the information to be integrated from lower levels to higher levels is held in working memory (Kintsch, 1998). Working memory refers to the memory used to store and process information simultaneously (Baddeley, 2000). It serves as a buffer for the most recently read propositions in a text, enabling their integration to establish coherence (Baddeley, 2007). However, working memory is limited. Inaccurate and slow word reading is proposed to affect comprehension by using too much memory capacity with little remaining for reading comprehension processes (Perfetti, 1985).

The higher-level information is stored in the form of schemas. Anderson and Pearson (1984) proposed schema theory which accounted for the configuration of information in memory and the interaction between the text and the readers’ background knowledge. In schema theory, prior knowledge structures are stored in memory, and reading comprehension is a process of mapping the information from the text onto these prior knowledge structures (Kendeou, Rapp, & van den Broek, 2003). Higher levels of processing require prior knowledge to get to the deeper implications of the text, and both E11 and ESL reading studies have shown that prior knowledge can be activated both automatically and strategically, which facilitates the understanding of text (Carrell, 1988; Kendeou et al., 2003; Pritchard, 1990; Stahl, Hare, Sinatra, & Gregory, 1991).

Inference is one of the most widely accepted schema notions; inferences are made on the basis of the readers’ prior knowledge and input (Anderson & Pearson, 1984). During reading comprehension, readers routinely generate possible inferences to connect information for establishing coherence in understanding the text. Inferences to establish global coherence are also generated when local coherence cannot achieve the goal (Graesser, Singer, & Tenenbaum, 1994). There are two main types of inferences. One type consists of coherence inferences which connect
different pieces of information from within the text, and the other type consists of elaborative inferences which connect information from the text with prior knowledge (Barnes, Denis, & Haefele-Kalvaitis, 1996; Cain, Oakhill, Barnes, & Bryant, 2001; McNamara & Magliano, 2009). Improving readers’ inference-making skills leads to substantial reading benefits in both EL1 and ESL learners (e.g., Cain & Oakhill, 1999; Hammadou, 1991; Horiba, 1996; van den Broek, Lorch, Linderholm, & Gustafson, 2001). A large number of empirical studies have indicated that readers must engage in extensive inferential processing to achieve successful reading comprehension (Cain et al., 2001; van den Broek, Risden, & Husebye-Hartman, 1995; van den Broek, 1994).

Strategy, another important factor responsible for successful reading, refers to methods for approaching a task and they are conscious rather than automatic processes. Reading comprehension strategies can aid readers in comprehension when additional processing is required. For example, when readers find information in the text which does not make sense, they can generate inferences to connect previous and subsequent ideas to establish coherence, and they can also use prior knowledge to help them understand. When readers meet unfamiliar or unknown words during a text, they can guess their meanings using contextual cues. A large number of empirical studies have indicated that improving readers’ reading strategies leads to substantial reading comprehension benefits (Cain et al., 2004; Rapp, van den Broek, McMaster, Kendeou, & Espin, 2007). The reading comprehension strategies, such as locating information, inferring the meanings of new words from context, making inferences, finding main ideas, and activating background knowledge can help readers gain a better understanding of text (Gascoigne, 2008). For beginning and less-skilled readers, strategic processing is more crucial because they need to apply reading strategies to reading comprehension to compensate for their lack of lower-level
reading skills. Less-skilled readers tend to use a surface strategy (focusing on the details in the text) rather than a deep strategy (focusing on main ideas) when remembering a text (e.g., Cain, 1999).

It should be noticed that Kintsch’s CI model assumes that readers have efficient word reading skills (Kintsch, 1998). However, this is not true for beginners or second language learners. For them, reading comprehension problems may arise from inefficiencies in English basic skills. These skills include phonological awareness, decoding (e.g., Gough & Tunmer, 1986; White, 2005), fluency (e.g., LaBerge & Samuels, 1974), and vocabulary knowledge (e.g., Beck, Perfetti, & McKeown, 1982; Perfetti, 1985). If some words cannot be recognized, the higher levels can compensate to some extent through top-down processing (Stanovich, 1986). However, if there are too many unknown words, comprehension becomes impossible. In this way, lower levels can act as bottleneck (LaBerge & Samuels, 1974). This is especially relevant in second language reading because bottom-up processes may not be sufficiently automated in L2 readers and basic skills need to be strongly established before higher-level skills exert effects in L2 readers (Zhang, Li, & Kirby, 2012).

**Measurement of Reading Comprehension**

The measurement of reading comprehension is a crucial component in understanding what reading comprehension is, how reading comprehension studies are conducted, and how reading comprehension can best be taught. Reliable and valid measures of reading comprehension can inform teachers’ instruction in reading. Reading comprehension can be assessed by different types of measures (Paris & Stahl, 2005), e.g., multiple-choice questions after silently-read short passages, cloze tests that require the student to read passages silently and provide missing words, short-answer constructed response tests, and long constructed responses tasks such as text
recalling and summary writing after reading a passage. There is growing recognition that different measures assess different reading skills (Cutting & Scarborough, 2006; Keenan, Betjemann, & Olson, 2008).

Some measures may be more demanding of bottom-up skills than others, e.g., cloze tests vs. question-and-answer tests (passage reading with short-answer questions) (Cutting & Scarborough, 2006). In their study with a sample of 97 first through tenth graders, Cutting and Scarborough (2006) included three measures of global reading comprehension commonly used by researchers: the Gates-MacGinitie Reading Test-Revised which involves silent passage reading with passages present during multiple-choice questions (GM; MacGinitie, MacGinitie, Maria, & Dreyer, 2000), the Gray Oral Reading Test-Third Edition which requires oral passage reading with passages removed during multiple-choice questions (GORT-3; Widerhold & Bryant, 1992), and the Weschler Individual Achievement Test which consists of silent passage reading with passages present during open-ended questions (WIAT; Wechsler, 1992). Cutting and Scarborough (2006) reported that nearly twice as much unique variance was accounted for by word reading measures (composite of word reading accuracy and decoding) on the WIAT (11.9%) than on GM (6.1%) or GORT (7.5%). The unique variance explained by oral language proficiency was similar for the WIAT and GORT-3 (each 9%) but higher for the GM (15%). Similarly, in Nation and Snowling’s (1997) sample of 7- through 10-year-olds, decoding explained more variance in reading comprehension scores when cloze tests versus question-and-answer tests were used (79% versus 53%).

Some reading comprehension measures may tend to tap higher-level skills, e.g., text recalling and summary writing tasks. For example in summary writing, students have to select what is most important in the text, in some cases they need to construct main ideas to integrate
text elements, and then write the summary in a coherent form, all of which are thought to require deeper or higher-level processing (Kirby & Woodhouse, 1994). Task conditions have been shown to influence summary writing; one task condition variable is the absence or presence of the text during summarization. Research has demonstrated that text absent summarization (i.e., instructing subjects to read a text and then summarize it without being able to refer back to the text) encourages the reader to process the text more actively, thus facilitating deeper processing, while text present summarization encourages lower level strategies (such as verbatim copying) that lead to lower level processing (Hidi & Anderson, 1986; Kirby & Pedwell, 1991; Stein & Kirby, 1992). Kirby and Woodhouse (1994) found that students who reported adopting a deep approach to learning performed better when they were required to summarize text without the text being present; however, students who reported a surface approach to learning were adversely affected by the absence of text. Later recall of the text by students in text absent summarization group was significantly greater than that by students in the text present summarization group (Stein & Kirby, 1992). Therefore, text absent summarization should be helpful in the higher-level aspects of reading comprehension as it encourages and requires deeper processing. Text summarization can be evaluated at different levels, for instance details, main ideas, and themes (Kirby & Pedwell, 1991; Stein & Kirby, 1992). The main idea and thematic levels should represent deeper levels of processing than details, just as Kintsch’s (1998) situation model represents deeper processing than the textbase and surface models.

Reading comprehension is clearly a complex construct which involves many different components. There are different levels of processing in reading and each level is necessary to achieve successful reading comprehension (Kirby & Woodhouse, 1994). As said earlier, reading comprehension is processed from a surface representation of the text (propositions), to a textbase
that distinguishes main ideas from details, and then to a situation model that integrates text information with prior knowledge (Kintsch, 1998). A number of factors at both lower and higher levels contribute to reading comprehension. The contributions of factors to reading comprehension might differ depending on which reading comprehension measure or which level of the reading comprehension measure is used. Multiple measures should be used to ensure that all relevant aspects of reading comprehension are covered.

**Similarities and Differences in L1 and L2 Reading**

There is some debate among researchers about the extent to which the English reading development of ESL learners is different from that of EL1 learners. It is certainly clear that there are unique features and processes of ESL reading, including the complex influences from L1 skills (e.g., Anderson & Han, 2009; Dressler & Kamil, 2006; Kieffer, 2011). It is also clear that English reading development draws on many of the same reading skills in both ESL and EL1 learners (August & Shanahan, 2006; Grabe & Stoller, 2011).

**Similarities in L1 and L2 Reading**

Research has shown that, at the beginning stage of learning English, L2 readers follow a similar pattern to L1 readers in learning English (Bialystok, Craik, Green, & Gollan, 2009; Geva & Yaghoub Zadeh, 2006). Many studies have pointed to the similarities between L1 and L2 reading processing (e.g., Dropp & Verhoeven, 2003; Genesee, Geva, Dressler, & Kamil, 2008; Lesaux, Rupp, & Siegel, 2007). The factors that predict reading comprehension in EL1 children are essentially identical to the requisites for reading acquisition among ESL students. Phonological awareness, decoding, fluency, and vocabulary have been shown to be significant predictors of EL1 reading and ESL reading (Geva & Siegel, 2000; Geva, Wade-Woolley, & Shany, 1997; Geva & Wang, 2001; Chen, Fen, Nguyen, Hong, & Wang, 2010; Wang, Yang, &
Cheng, 2009). When L2 students are exposed to L2 reading sufficiently, they appear to develop word reading skills at a level equivalent to those attained by L1 students (e.g., Chiappe, Siegel, & Wade-Woolley, 2002; Geva, Yaghoub-Zadeh, & Schuster, 2000; Lesaux & Siegel, 2003). However, L2 students usually perform more poorly on measures of vocabulary and reading comprehension than L1 students at the same grade level (Carlisle, Beeman, Davis, & Spharim, 1999; Lesaux et al., 2008).

Research has also shown that, at very advanced levels of learning L2, L1 and L2 reading abilities tend to be similar (Grabe & Stoller, 2011). For example, Wang and Koda (2007) examined advanced Chinese and Korean ESL adults on English word reading processes and found that both groups performed in ways similar to native English readers. Both groups recognized high-frequency and regular words faster and more accurately than low-frequency and irregular words, which is similar to the recognition demonstrated by EL1 learners when performing this task. Similar to EL1 readers, ESL readers tend to employ strategies to aid them in comprehension (Anderson, 2005). For ESL readers, strategic processing may be more crucial because it may compensate for their lower word reading skills.

**Differences in L1 and L2 Reading**

However, there exist considerable cognitive processing differences between L1 and L2 reading development. When L1 students enter Grade 1, they usually already know 5000 words orally, but this does not apply to L2 students (Nagy & Herman, 1987). Therefore, reading processing will be different for L2 and L1 students when they try to match the words they read to the words they know. Evidence from think-aloud protocols has shown that students tend to pay more attention to bottom-up reading strategies such as focusing on individual words or sentences when they read L2 texts because of their low proficiency in L2 (Grabe, 2009). However, this does
not happen often in L1 reading comprehension, in which students use more top-down reading strategies such as inference and prediction and they associate prior knowledge to the text to help them understand the text better (Grabe, 2009).

As L2 students come from different L1s, the linguistic differences between L1 and L2 may influence L2 reading processing. L2 students learn to read in L2 using a two-language processing system (L1 and L2) rather than just their L2 system only because their L1 never completely turns off (Akamatsu, 2002; Grabe & Stoller, 2011; Wang & Koda, 2007). For example, Wang and Koda (2007) found that advanced Chinese and Korean ESL students recognized English words in a way similar to that of English native speakers. However, Korean students read words faster and more accurately than did Chinese students, although they were matched on their English language proficiency. The authors attributed this result to the L1 influence on the two groups. They argued that Korean, unlike Chinese, is an alphabetic language, and this facilitates Korean students’ ability to recognize English words. Therefore, L2 students tend to draw on L1 processing skills when they read L2.

In general, when learners learn to read in L2 at a very young age, they perform the same way as L1 learners do in learning to read in L1 (Bialystok et al., 2009). The reading processes are also similar when learners have a very high proficiency in L2 (Grabe & Stoller, 2011). L1 and L2 reading differences can be found among those who begin the L2 learning at a later stage. Their L2 reading development is influenced by the prior or concurrent development of L1 skills (Wang & Koda, 2007). However, it does not mean that L2 readers are different from L1 readers in terms of learning to read. Rather, L2 reading research suggests that L2 readers combine both L1 and L2 literacy knowledge in learning to read in L2. More exploration of L2 reading processes, the interaction between L1 and L2 reading, and the development of L2 input is thus needed.
English Immersion Programs in China

For readers to better understand the focus of the present study, it is necessary to describe the context within which it took place. This dissertation explored the cognitive processes of English reading comprehension of middle school English-immersion students in China. The purpose of immersion education is to allow children who speak the language of the majority at home to achieve proficiency in another language (Swain, 1996). Immersion is distinct from traditional approaches to bilingual education because L2 is not only explicitly taught but is also the medium of curriculum instruction (Genesee, 1985, 1987). English-immersion programs in China borrowed the model of French-immersion programs in Canada. French immersion programs started in Canada in the 1960s and have prospered for over four decades. French immersion programs in Canada vary in form. They may start once the child is in kindergarten or Grade 1 (early immersion) or after a number of years of schooling in the mother tongue (late immersion). They also vary in the quantity of teaching time in the second language (total vs. partial immersion) (Comeau, Cormier, Grandmaison, & Lacroix, 1999). Research in French immersion programs has shown that immersion is an effective means of facilitating preschool and primary school students’ language proficiency and literacy without any detrimental effects to their first language (Cummins & Carson, 1997; Genesee, 1995; Lapkin, Hart & Turnbull, 2003; Swain, 1996; Turnbull, Lapkin, & Hart, 2001). Furthermore, Genesee (2004) demonstrated that French immersion students continue to perform as well as their peers in non-immersion programs in all English reading skills after Grade 6. However the purpose of this dissertation is not to compare French immersion programs to English immersion programs or to evaluate English immersion programs but rather to understand the cognitive processes of reading comprehension.
in English immersion students in China; therefore, literature about French immersion programs is not reviewed here.

English immersion programs in China the China-Canada-USA English Immersion (CCUEI) Project, starting from kindergarten to Grade 6 in 1998 (for a review, see Cheng, 2012; Qiang, Huang, Siegel, & Trube, 2011). Xi’an was the first city that initiated the English-immersion programs and then other major cities followed. The programs were modelled after French immersion programs in Canada and were viewed as an educational transfer (Qiang & Kang, 2011). However, these English-immersion programs differ from French-immersion programs in several ways. They do not have as much second language (L2) input as is typical in immersion programs in Canada. For example, a minimum of 50% instruction is given through French (L2) in French immersion programs in Canada (Genesee & Jared, 2008). The curriculum design in English-immersion programs is that 30-40% of the curriculum is taught in English, including English language arts, social studies, and science. The other 60-70% of the curriculum is taught in Chinese, including key subjects such as Chinese and mathematics (Cheng, Li, Kirby, Wade-Woolley, & Qiang, 2011). Other subjects, such as physical education, art, and music are taught in Chinese or English varying from immersion program to immersion program according to teacher skills. With regards to the teacher resources in the English immersion programs in China, all of the English teachers are native Chinese speakers. Most of them have a certificate or diploma of teaching English but few of them have a Bachelor’s degree (Song & Cheng, 2011). Their English proficiency levels are not considered equal to those of native English speakers (Knell et al., 2007). In addition, these immersion teachers have not been trained systematically on subject content in English, which is considered to be a constraint in immersion education in China (Feng & Wang, 2007; Hu, 2007)
Despite the teachers’ lack of competence, empirical studies in Chinese English-immersion programs from Grades 1 to 6 have shown that: (a) immersion students performed significantly better than non-immersion students on English after considering L1 vocabulary or reading ability (Cheng et al., 2010; Knell et al., 2007), and there was no significant difference between immersion and non-immersion students on Chinese and mathematics in Grades 2 and 4 but immersion students outperformed non-immersion students in Grade 6 (Cheng et al., 2010), (b) English phonological awareness was a significant predictor of English reading achievement and listening comprehension for the immersion students in the early grades (Knell et al., 2007; Li, Cheng, & Kirby, 2012; Li, Kirby, Cheng, Wade-Woolley, & Qiang, 2012), and there was little evidence of cross-linguistic transfer from Chinese phonological awareness and naming speed to English reading (Li et al., 2012; Li, Kirby, & Georgiou, 2011), and (c) immersion students demonstrated more advanced English communication skills than non-immersion students with classmates and teachers (Liang, 2011; Pei, Du, & Wang, 2009).

However, these findings should be interpreted with caution because relevant variables such as intelligence and SES were not controlled in these studies. Future research should consider the relevant variables when examining the academic performance of immersion students and non-immersion students. Despite the overall positive results from English immersion programs, some researchers have addressed challenges that immersion programs face, including limited time and limited subject areas permitted to be used for implementation of the program, large class sizes, a lack of qualified teachers, and limited supportive resources (Huang, Trube, & Yu, 2011). For example, Hu’s (2008) review of Chinese-English bilingual education in China questioned the effectiveness of immersion programs in China due to the limited resources for English-medium
instruction and the quality of such instruction. He also argued that lack of a sociolinguistic context for using English after school was another major concern for immersion programs.

After the English-immersion programs in elementary schools had run for six years, the immersion programs were extended to junior middle schools. In Xi’an, the capital city of Shannxi province in China, three junior middle schools started English-immersion programs in 2004. Later, another four junior middle schools implemented English-immersion programs in Xi’an. The immersion programs in junior middle schools are similar to those in elementary schools but the amount of English language instruction is reduced. Thirty percent of the curriculum is taught in English (L2) including English Language Arts and either Social Studies or Science, which varies from school to school. Most of the teachers have obtained Bachelor’s degrees in English. These middle school English immersion teachers reported feeling a lack of professional training although they have undertaken training in various forms provided by the CCUEI project team; these include workshops led by English professors, inter-school lesson observations, and teacher discussions (Kong, Hoare, & Chi, 2011). It is noted that the English-immersion programs in junior middle schools have not borrowed the entire model of French immersion programs in Canada but have just adopted the concept underlying immersion programs to ensure that some non-language subjects are taught in the second language, which is more like “content-based language teaching (CBLT)” (Brinton, Snow, & Wesche, 1989). English immersion programs in China have been referred to as “CBLT” or “Chinese-English bilingual education” by different educators and researchers (Gu, 2006; Hoare, 2010; Hu, 2008). In this dissertation, I will still refer to them as “English immersion programs” because that is how they are often described in China.
Rationale for the Dissertation

A great deal of research has been carried out on reading comprehension in L1 contexts (especially in English as an L1). However, comparatively few studies have explored reading comprehension in an ESL context. This is perhaps due to the greater complexity of ESL contexts. In an ESL context, many students have already had experience learning to reading in their L1s. This L1 linguistic knowledge may either support L2 learning through positive linguistic transfer or interfere with L2 learning through negative transfer; this makes research in L2 contexts more complex. In addition, diverse L2 learning environments exist among L2 students. Some L2 students are in countries in which the native language is English and they are more exposed to English than those who learn English in their home country, where English is spoken as an L2 or foreign language. Therefore, it is difficult to generalize results from reading research in L2 because there are great differences in L2 exposure and proficiency levels in L2 learners from different backgrounds.

Many studies of ESL students’ reading ability have been conducted with Spanish-speaking ESL students (e.g., Gottardo & Mueller, 2009; Lindsey, Manis, Bailey, 2003; Proctor, Carlo, August, & Snow, 2005). However, there is a large population learning English as an L2 in Asia. The number of users of English in Asia has surpassed that of English native speakers (Cheng & Curtis, 2010; McKay, 2002). According to Crystal (2009), there are 200-500 million Chinese people studying English just in China. Thus, it is crucial to understand English reading development of Chinese students. The findings of such research could have important implications for ESL instructional and assessment practices.

Although English-immersion programs were implemented in Chinese middle schools almost 10 years ago, all the existing studies are limited to elementary school students and no
empirical research has been done on English literacy development of immersion students in middle schools. Reading research in English (L1) has demonstrated that both basic and higher-order reading skills are crucial to successful reading comprehension (National Reading Panel, 2000; Oakhill et al., 2003; van den Broek et al., 2005). According to the report from the National Literacy Panel on Language-Minority Children and Youth (August & Shanahan, 2006), many of the skills that predict reading comprehension among L1 learners play similar roles in predicting reading comprehension among L2 learners but these studies were all conducted with elementary school learners. Thus, there is a need to investigate whether the factors involved in L1 reading comprehension play similar roles in L2 reading comprehension in higher grades.

**Organization of the Dissertation**

The remainder of this dissertation is organized as follows. In Chapters 2 and 3, two empirical studies of ESL reading comprehension are described. Each includes individual literature review, methodology, results, and discussion sections. General discussion of the two studies with educational implications and conclusion are presented in Chapter 4.
Chapter 2

Predictors of Reading Comprehension in Chinese Students Learning English

Miao Li and John R. Kirby
Queen’s University, Canada

Abstract

This study examined the cognitive predictors of English reading comprehension in Chinese students learning English. Two hundred and forty-six Grade 8 students in China were administered measures of decoding, listening comprehension, vocabulary breadth, vocabulary depth, inference and strategies, and reading comprehension. Two questions are addressed: (a) How well does the Simple View of Reading model apply to Chinese ESL students? (b) Do two dimensions of vocabulary knowledge (i.e., breadth and depth) make different contributions to two lower and higher levels of reading comprehension? The results showed that decoding and linguistic comprehension, the two factors in the SVR model, did not predict ESL reading comprehension well. Inference and strategies added further variance to general understanding of text. Vocabulary breadth and depth explained unique variance in reading comprehension, with vocabulary breadth contributing more to general understanding of text and vocabulary depth to deep processing of text. Discussion focuses on how theories of reading comprehension need to be adapted to address second language reading.
Reading comprehension is the ability to extract meaning from text. It is regarded as a fundamental skill necessary for educational achievement. The National Reading Panel (2000) reported that lower-level skills (e.g., phonological awareness, decoding, fluency, and vocabulary knowledge) are important for successful English reading. It is also clear that higher-level skills (e.g., inferencing skills, reading strategies, comprehension monitoring, integration of text information) are essential to successful English reading (Oakhill, Cain, & Bryant, 2003; Pressley, 2002; Snow, 2002; van den Broek et al., 2005). In English as a first language (EL1), a great deal of research has been conducted to explore the cognitive factors which are related to reading comprehension. For example, the Simple View of Reading (SVR) model has been proposed as a framework within which to address the complexities of reading comprehension in EL1 children (e.g., Gough & Tunmer, 1986; Kendeou, van den Broek, White, & Lynch, 2009; Kirby & Savage, 2008). In this model, decoding and linguistic comprehension are the two main predictors of reading comprehension. Decoding is the ability to read isolated words quickly, accurately, and silently, while linguistic comprehension is the process by which words, sentences, and discourses are interpreted (Gough & Tunmer, 1986). When they proposed this SVR model, Gough and Tunmer clearly asserted that linguistic comprehension can be estimated by listening comprehension tasks, in which participants answer questions after listening to texts. Recently, studies have shown more complex relations among decoding, listening comprehension, oral vocabulary, and reading comprehension than were implied by the original simple view of reading model (e.g., Braze, Tabor, Shankweiler, & Mencl, 2007; Ouellette & Beers, 2010; Protopapas, Mouzaki, Sideridis, Kotsolakou, & Simos, in press). In these studies, vocabulary made a contribution to reading comprehension beyond that made by decoding and listening comprehension. According to Perfetti and Hart’s (2002) lexical quality hypothesis, which posits
that skilled reading depends on high quality lexical representations, vocabulary is another powerful predictor of reading ability (e.g., Anderson & Freebody, 1981; Beck, McKeown, & Kucan, 2002; Kirby, Desrochers, Roth, & Lai, 2008; Tannenbaum, Torgesen, & Wagner, 2006). It has been proposed that traditional measures of vocabulary assess mainly breadth of vocabulary, because although they do capture a sense of how many words are known, they do not indicate how well or how deeply those words are known (Nagy & Herman, 1987; Wesche & Paribakht, 1996). Depth of vocabulary knowledge is a relatively new construct which has been claimed to measure deeper understanding of word meanings (Anderson & Freebody, 1981; Henriksen, 1999; Perfetti, 2007; Perfetti & Hart, 2002). Compared to breadth of vocabulary, depth of vocabulary is less researched. Furthermore, studies have suggested that other predictors such as students’ ability to make inferences and employ reading strategies need to be considered when understanding reading comprehension (Cain, Oakhill, & Lemmon, 2004; Cromley & Azevedo, 2007; Kirby & Savage, 2008).

Reading comprehension in English as a Second Language (ESL) has been given great attention in recent years because of the increase in numbers of children learning in ESL situations and the challenges they face in doing so. However, compared to the amount of research concerning the predictors of reading comprehension in EL1 students, relatively few studies have sought to explore predictors of reading comprehension in the ESL context. This is perhaps due to the complexity of examining ESL reading development. Moreover, most of the studies of ESL students’ reading ability have been conducted with Spanish-speaking ESL students (e.g., Gottardo & Mueller, 2009; Lindsey, Manis, Bailey, 2003; Proctor, Carlo, August, & Snow, 2005). Spanish (L1) and English (L2) are similar alphabetic scripts with many cognate words. Extending the research to those ESL students whose first language is very different from English may provide
us with a more comprehensive understanding of ESL students’ reading comprehension and thus a better basis for the design of effective instruction.

The large number of students learning English in China provides a useful opportunity to study reading comprehension in students whose first language is very different from English, especially because the orthographies of English and Chinese are also different. In this study we focus on middle school students rather than elementary school students, partly because ESL students usually acquire comprehension skills later than EL1 students due to less exposure to English, and partly because reading comprehension becomes more challenging in middle school and students struggle to comprehend academic texts. The National Literacy Panel on Language Minority Children and Youth (2006) indicated that most reading studies of ESL students have been conducted with elementary school students focusing on word-level skills. A limited amount of research has been done with ESL students focusing on text-level skills and reading comprehension. Thus, the overall purpose of the present study was to investigate the predictors of Chinese ESL middle school students’ reading comprehension. We investigate the application of two reading models, the simple view of reading and the lexical quality hypothesis. In evaluating the SVR model, we sought to examine whether decoding and linguistic comprehension are powerful predictors of English reading comprehension in Chinese ESL children and whether the SVR model needs to be expanded to include separate predictors for inference and strategies. In testing the lexical quality hypothesis, we attempted to explore the effects of different dimensions of vocabulary (breadth and depth) on different aspects of reading ability (i.e., word reading, and literal or detail versus inferential or main idea comprehension) after considering the two predictors in the SVR model.
Predictors of Reading Comprehension in EL1 and ESL Children

In EL1 reading research, various predictors, such as decoding, motivation, self-perception, listening comprehension, vocabulary, gender, inference and strategies, home literacy, and socioeconomic status, have been shown to contribute to reading comprehension (e.g., Adolf, Catts, & Little, 2006; Burgess, Hecht, & Lonigan, 2002; Cain, Oakhill, & Bryant, 2004; Cromley & Azevedo, 2007; Wigfield & Guthrie, 1997). Due to the different amount of exposure to English and the influence of L1, reading development of ESL students is somewhat different from that of EL1 students. When EL1 students enter Grade 1, they usually know many more words than ESL students because of extensive exposure to English (Nagy & Herman, 1987). Moreover, ESL students learning to read in an L2 are supported by a two-language processing system (L1 and L2 together) rather than just the L2 system because their L1 never completely turns off (Grabe & Stoller, 2009). Despite the differences in the reading development between ESL and EL1 learners, more similarities than differences between EL1 and ESL groups are found in the predictors’ contribution to reading comprehension (Geva, 2006; Lesaux, Geva, Koda, Siegel, & Shanahan, 2008). For example, tests of the SVR model have demonstrated that decoding and linguistic comprehension are strong predictors of reading comprehension in ESL children (Gottardo & Mueller, 2009; Hoover & Gough, 1990; Proctor, Carlo, August, & Snow, 2005; Yaghoub Zadeh, Farnia, & Geva, 2012); however it should be noted that most of the ESL children in these studies had L1s that are related to English and use alphabetic scripts. It needs to be investigated whether these same effects would hold true for children whose L1 is Chinese, sharing neither common linguistic history nor an alphabetic script.

Further, some researchers have argued that the two main components in the SVR model may not be sufficient to account for reading comprehension (see Kirby & Savage, 2008; Savage,
2006). For example, Kirby and Savage (2008) suggested that reading strategies may explain additional variance in reading comprehension. Inference is crucial in reading comprehension (Cain et al., 2004; Cromley & Azevedo, 2007). Vocabulary has been thought not to be fully reflected in the SVR model (Braze et al., 2007; Ouellette & Beers, 2010).

Reading comprehension theories in ESL are not yet well developed. Usually, the EL1 reading models are tested to see whether they apply to ESL readers but the degree to which EL1 reading comprehension models can be extended to ESL children reading English is still unclear. In this study, we attempt to explore how well the predictors in the SVR model are related to reading comprehension in ESL students, and then whether vocabulary, inference, and reading strategies, which have been shown to be related to reading comprehension in EL1 students, make additional contributions. We are interested in these specific factors because, compared to other factors, they are more plausible causes of reading comprehension and are not well represented in the reading comprehension in ESL students. We now review each of the predictor constructs in more detail.

**Decoding and Listening Comprehension**

In the SVR model, decoding and listening comprehension are the two predictors of reading comprehension (Gough & Tunmer, 1986). It has been well established that decoding is necessary for reading comprehension in both EL1 and ESL learners (Adolf et al., 2006; Geva & Wang, 2001; Gottardo & Mueller, 2009; Johnston & Kirby, 2006; Lesaux, Rupp, & Siegel, 2007; National Reading Panel, 2000). Slow decoding expends too many cognitive resources, leaving insufficient cognitive resources for higher-level processing in reading comprehension, which impedes reading comprehension (LaBerge & Samuels, 1974; Perfetti, 1985; Perfetti & Lesgold, 1977, 1979; Winne, 1995). In terms of measuring decoding, Tunmer and Greaney’s (2009)
argument that measures of decoding should be viewed as developmentally constrained can shed light on how decoding should be assessed. They indicated that pseudoword measures of decoding should be used during the early stages of learning to read, because letter-sound relationships play a crucial role in early literacy development; that context-free word recognition measures should be used at later stages of reading growth, because students have acquired sufficient orthographic knowledge; and finally that timed measures of word recognition would be the most appropriate measure at more advanced stages to assess the development of automaticity in word recognition (i.e., fluency).

When Gough and Tunmer initially proposed listening comprehension as a predictor of reading comprehension, they viewed it as “the process by which given lexical (i.e., word) information, sentences, and discourses are interpreted” (Gough & Tunmer, 1986, p. 7). Kirby and Savage (2008) argued that listening comprehension represents all of verbal ability, including vocabulary, syntax, and the construction of mental schemas. Studies with EL1 students have demonstrated the predictive role of listening comprehension in reading comprehension (Aarnoutse, van den Bos, & Brand-Gruwel, 1998; Adolf et al., 2006; Conners, 2009; Chen & Vellutino, 1997; Johnston & Kirby, 2006; Savage, 2001; Tiu, Thompson, & Lewis, 2003). Studies with ESL students have also found that listening comprehension is an important predictor of reading comprehension (e.g., Gottardo & Mueller, 2009; Hoover & Gough, 1990; Proctor et al., 2005; Yaghoub Zadeh et al., 2012). For example, Proctor et al. (2005) reported that 44% of the variance in reading comprehension was explained by listening comprehension in Grade 4 Spanish-speaking ESL learners.
Vocabulary

Although vocabulary has proved to be one of the strongest predictors of reading comprehension in both EL1 and ESL learners (e.g., Anderson & Freebody, 1981; Beck, McKeown, & Kucan, 2002; Farnia & Geva, 2011; Laufer, 1992; Nation, 2001; Read, 2000; Tannenbaum, Torgesen, & Wagner, 2006), the nature of vocabulary is not quite clear. Anderson and Freebody (1981) first made the distinction between two dimensions of vocabulary knowledge: breadth and depth. According to them, vocabulary breadth refers to “the number of words for which the person knows at least some of the significant aspects of meaning” (p. 93), while vocabulary depth is “the quality or depth of understanding” (p. 93). There is a consensus that breadth of vocabulary involves how many words are known but researchers have different views regarding depth of vocabulary. For example, Haastrup and Henriksen (2000) defined depth of vocabulary from the perspective of word meaning and collocation. Other researchers have suggested that morphological information is an aspect of depth of vocabulary because knowledge of affixes and roots can help students understand the formation of words, which enriches their understanding of the words in depth (Bowers & Kirby, 2010; Kieffer & Lesaux, 2008; Proctor, Uccelli, Dalton, & Snow, 2009). Others have said that depth of vocabulary should also include the understanding of words’ multiple meanings and how the meanings can be used within multiple contexts (Beck et al., 2002; Tannenbaum et al., 2006). Based on these definitions, depth of vocabulary can be seen to involve at least precision and multiplicity of meaning, collocational use of words, and information about word forms (i.e., phonological, morphological, orthographic information). There is a need for a full set of measures covering all of the components of vocabulary depth to measure it as a construct more validly.
Ouellette (2006) found that, for Grade 4 EL1 students, both vocabulary breadth and depth predicted reading comprehension after controlling for age, nonverbal IQ, decoding, and visual word recognition. Depth of vocabulary measured by word definition predicted reading comprehension beyond measures of vocabulary breadth, but the effect of vocabulary breadth on reading comprehension disappeared once depth of vocabulary was controlled. In contrast, Tannenbaum et al. (2006) found that breadth of vocabulary had a stronger relationship to reading comprehension than did depth in Grade 3 students but these two dimensions had significant overlapping variance that also contributed to reading comprehension. The depth of vocabulary was measured by a multiple meanings measure which required participants to define words with multiple meanings presented in different contexts, and an attributes measure which asked each participant to name different attributes of provided nouns. There is also one study examining the effects of breadth and depth of vocabulary in the SVR model. After controlling decoding, listening comprehension, phonological awareness, and irregular word reading, Ouellette and Beers (2010) indicated that only vocabulary breadth but not depth predicted reading comprehension in Grade 6 English-speaking students and this effect did not appear in Grade 1 students.

Similar to EL1 studies, ESL studies on the relationship between vocabulary and reading comprehension have mostly focused on vocabulary breadth rather than depth. However, there is growing attention to the effects of breadth and depth of vocabulary knowledge on reading comprehension in L2 students (Proctor et al., 2009; Qian, 1999, 2002; Vermeer, 2001). In his study with Canadian ESL university students, Qian (2002) found that depth measured by word associations test explained more variance (13%) than breadth (8%) in reading comprehension. To better understand how quality of vocabulary knowledge affected reading comprehension, Proctor
et al. (2009) examined the role of depth of vocabulary in reading comprehension in a group of bilingual and monolingual 5th-grade students. Results indicated that depth of vocabulary explained an additional 3% of the variance in reading comprehension beyond word decoding and breadth of vocabulary for both monolingual and bilingual students, which is consistent with Qian’s (1999, 2002) findings. It should be noted that the unique variance contributed by depth of vocabulary in Proctor et al.’s (2009) study is less than that reported in the previous studies because both breadth of vocabulary and word decoding were controlled.

Besides reading comprehension, vocabulary knowledge also helps to support the development of word reading in EL1 children (e.g., Bryant, Maclean, & Bradley, 1990; Garlock, Walley, & Metsala, 2001; Nation & Snowling, 1998). However, Ouellette (2006) found that only vocabulary breadth predicted word reading significantly after controlling age, nonverbal intelligence, and depth of vocabulary, but depth of vocabulary knowledge did not predict word reading after controlling all of the other predictors. This may indicate that only relatively shallow vocabulary knowledge is required to help in word identification in reading. But Ouellette and Beers (2010) reported that vocabulary depth measured by word definition explained additional variance in irregular word reading after accounting for decoding, phonological awareness, and vocabulary breadth in Grade 1 EL1 students but not in Grade 6 EL1 students. We know of no study that has used measures of vocabulary depth to assess the role of vocabulary knowledge in word reading in ESL learners. There is a need for research to examine the effect of vocabulary depth on word reading in ESL learners, taking into account the effect of vocabulary breadth.

Based on the reviewed studies, the extent to which breadth and depth of vocabulary contribute to reading comprehension and word reading is far from clear. The lack of clarity is due to varying grade levels of participants, the lack of theory-driven and broadly-based measures of
vocabulary depth, and the different reading measures that have been used. Participants’ grade levels are important because there is good reason to expect different effects of vocabulary breadth and depth at different levels of reading competence (Tannenbaum et al., 2006). The measurement issues around vocabulary depth have yet to be settled, and it is not even clear that vocabulary depth is reliably distinct from vocabulary breadth. Given the complexity of the vocabulary depth construct, it should be measured with a battery of measures, not with any single test.

Inference and Strategy

There is considerable evidence that readers’ inference-making skills improve their reading comprehension (e.g., Cain & Oakhill, 1999; van den Broek, Lorch, Linderholm, & Gustafson, 2001). Studies have indicated that readers must engage in extensive inferential processing to achieve successful EL1 and ESL reading comprehension (Barnes, Denis, & Haefele-Kalvaitis, 1996; Cain & Oakhill, 1999; Cain, Oakhill, & Bryant, 2004; Clarke & Silberstein, 1977; Coady, 1979; van den Broek, Risden, & Husebye-Hartman, 1995). Of the different types of inference skills which lead to substantial reading benefits, lexical inferencing skill (in which readers attempt to infer new word meanings from context) has been widely reported to facilitate ESL readers’ reading comprehension (Nassaji, 2003; Paribakht & Wesche, 1999; Pulido, 2007). In the present study, we investigated two types of text-level inferences, based on Cain et al. (2004) and Cromley and Azevedo (2007): coherence inferences, which require a combination of different pieces of information from within the text, and elaborative inferences, which require readers to integrate information from the text with prior knowledge (Cain et al., 2001).

Kirby and Savage (2008) suggested that reading strategies may explain some additional variance in the SVR model. Research has shown that improving readers’ strategy skills led to
substantial reading comprehension benefits (Guthrie et al., 2004; Rapp, van den Broek, McMaster, Kendeou, & Espin, 2007). For example, teaching readers to locate information, infer the meanings of new words from context, make inferences, find main ideas, and activate background knowledge can help readers gain a better understanding of text (Gascoigne, 2008; Graesser, Singer, & Trabasso, 1994; van den Broek, Fletcher, & Risden, 1993). Most studies of reading strategies have used self-report questionnaires, which may not reflect accurately the reading strategies used in reading. Cromley and Azevedo (2007) adopted a different approach, asking students to complete multiple-choice items which required the application of various reading strategies after reading short passages. They found that reading strategies had a significant effect, via inference skill, on reading comprehension. In ESL reading research, strategies have also been found to be significant predictors of reading comprehension (e.g., Cohen & Macaro, 2007; Hurd & Lewis, 2008; O’Malley & Chamot, 1990; Oxford, 1990; Phakiti, 2003, 2008; Purpura, 1997, 1998, 1999; Sheorey & Mokhtari, 2001). The evidence suggests that successful and unsuccessful ESL readers approach reading in different ways, the successful readers using more and higher-level strategies than unsuccessful readers (Cohen, 2007; Gascoigne, 2008; Sheorey & Mokhtari, 2001).

In spite of the evidence that inference and reading strategies contribute to reading comprehension, there has been no empirical research examining whether they should be added to the SVR model. Although it may be thought that inference and reading strategies would be part of the linguistic comprehension component, written material offers more opportunities for each to be productive (Kirby & Savage, 2008). Their role in predicting reading comprehension may depend upon how reading comprehension and inference and reading strategies are measured. Some
reading comprehension measures may undermine the possible benefits of inference and reading strategies, and others may require specific inference and strategy skills.

**Reading Comprehension**

Reading comprehension can be assessed by different measures and different reading comprehension tests measure different cognitive skills (Cutting & Scarborough, 2006; Keenan, Betjemann, & Olson, 2008; Nation & Snowling, 1997; Paris & Stahl, 2005). Using latent variable modeling, Francis, Fletcher, Catts, & Tomblin (2005) found that decoding was more related to a cloze test than to silent or oral passage reading with multiple-choice questions (see also Cutting and Scarborough, 2006). Research has suggested that some reading comprehension measures, such as cloze or picture selection tests, depend more on decoding skills than other comprehension skills but other reading comprehension tests, such as oral or silent reading passages with comprehension questions, required more language comprehension skills (Keenan et al., 2008). Not only do different reading comprehension measures assess different cognitive skills, but the same reading comprehension test may measure different skills (Keenan et al., 2008).

Building on theory, reading comprehension occurs at different levels of processing, proceeding from superficial (literal) to deeper (more inferential) levels of representation (Biggs & Collis, 1982; Craik, 2002; Kintsch, 1998; Kirby, Cain & White, 2012; Kirby & Woodhouse, 1994). The deeper level of representation may depend more on effortful, deliberate, and strategic processing whereas the superficial level of representation may rely on more automatic processing (Kirby & White, 2011). Research is needed to examine whether levels of representation can be identified through different reading comprehension measures and whether different levels are predicted by different skills.
The Present Study

Of the few studies that have been done to examine the predictors of ESL reading comprehension and test the SVR model in ESL students, most have examined bilingual children whose L1 and L2 are similar in orthography, such as Spanish (L1) and English (L2) (Gottardo & Mueller, 2009; Hoover & Gough, 1990; Proctor et al., 2005). Although Zadeh et al.’s (2010) sample included some students whose L1 and L2 were orthographically and linguistically distant, such as Chinese (L1) and English (L2), it had students from many different linguistic backgrounds, many of which were not that distant. There have been no studies investigating whether the SVR model can be applied to the reading comprehension of the majority of the samples that are Chinese students learning English. Therefore, one purpose of this study was to test whether the SVR model is applicable to Chinese students who learn English as their L2. More precisely, we are interested in whether decoding and linguistic comprehension can explain as much variance in reading comprehension in Chinese ESL students as in EL1 students and whether other predictors account for additional variance. Because previous studies examining the SVR model had not included inference and strategy measures and most had not assessed the effect of the students’ L1 reading comprehension ability, we included these two measures. Our goal in doing so was to provide a deeper understanding of the relationship between linguistic skills and reading comprehension performance in ESL children.

Research in the SVR model has emphasized the crucial role of vocabulary in addition to decoding and listening comprehension in reading comprehension. Vocabulary can be classified into two dimensions: breadth and depth. In the studies which examined the effect of vocabulary on reading comprehension in the SVR model, only Ouellette and Beers (2010) went one step further to divide vocabulary into breadth and depth and investigate their effects on reading.
comprehension. Although breadth and depth of vocabulary are described as two different dimensions of vocabulary knowledge, there has been evidence that they are highly correlated (e.g., Nurweni & Read, 1999; Qian, 1999, 2002; Vermeer, 2001). The precise nature of the relationship between these two different aspects of vocabulary is still far from clear. Although vocabulary is related to word reading, there is no research examining the effects of breadth and depth of vocabulary on word reading in L2 learners. In addition, most L2 studies examining the contributions of breadth and depth of vocabulary to L2 reading comprehension have not controlled L1 proficiency, L2 word reading, and listening comprehension which also influence L2 reading comprehension. Furthermore, measures of vocabulary depth have not always been broadly based or well designed. The second purpose of the present study was thus to explore the relationship between breadth and depth of vocabulary and their effect on word reading and, beyond the SVR model, on reading comprehension. Based on the literature reviewed, vocabulary depth involves three aspects: elaborated meaning, morphology, and collocational use of words. Therefore, measures of each of these were given to students to obtain a fuller assessment of vocabulary depth. To measure vocabulary breadth, we used the Gates-MacGinitie Vocabulary subtest (MacGinitie & MacGinitie, 1992), which is a standardized vocabulary breadth measure for EL1 students.

As different reading comprehension measures depend upon different cognitive skills and even the same measure assesses different skills, we gave the students two reading comprehension measures. One was a commonly used measure, in which students answered multiple-choice questions after silently reading passages. The other was a text-absent summary writing task to assess their deep understanding of a text. Each reading comprehension measure was divided into lower and higher levels. The multiple-choice reading comprehension questions were classified
into literal and inferential levels and the summary writing measure into detail and main idea
levels.

This study addresses the following questions regarding Chinese ESL children: (a) How
well do the two predictors in the SVR model (i.e., decoding and linguistic comprehension) predict
English reading comprehension? (b) Do measures of inference and strategy account for additional
variance? (c) What are the effects of breadth and depth of vocabulary on English reading
comprehension? and (4) What are the effects of breadth and depth of vocabulary on English word
reading?

Method

Participants

The participants were 248 students from the four English-immersion classes in Grade 8
from a middle school in Xi’an, China. This research was cleared through Queen’s University
General Research Ethics Board (see Appendix A). A Letter of Information (see Appendix B) was
provided to parents requesting permission to allow their children to participate in this research.
All students agreed to participate and their parents signed the consent form (see Appendix C).
Two students who were absent for several tests were deleted from the study. Of the total 246
participants, there were 69 boys and 177 girls. The gender imbalance may be due to the better
performance of girls on the English entrance exam that the school uses to select students.

Measures

Nonverbal Ability. Raven’s progressive Matrices (Raven, Raven, & Court, 1998) was
used to measure students’ nonverbal ability as a control measure. It comprises 60 items, divided
into five sets of increasing complexity. Each item presents a geometric design with a part
missing, and students have to choose one of six or eight alternatives to complete the pattern. The score was the number of correct answers.

**English Decoding.** The Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999) was used to assess children’s ability to pronounce isolated words accurately and quickly. A reading fluency measure was used to measure decoding as Tumner and Greaney (2009) indicated that automaticity in word recognition becomes essential at the advanced stage of reading development. Students were asked to read as many of the 104 words as quickly and as accurately as possible within 45 seconds. The score was the number of words read correctly in 45 sec.

**English Listening Comprehension.** 19 items\(^1\) were chosen from the *Woodcock Listening Comprehension Test* (Woodcock, 1998), the first 9 items and odd items from items 10 to 30; some items were changed to adapt to Chinese culture and the students’ background knowledge. This is a cloze-type assessment in which the student listened to tape-recorded passages in order of increasing difficulty and produced an oral response to unfinished sentences. The score was the number of correct answers.

**Gates-MacGinitie vocabulary subtest.** (Level C, Grade 3) was given to the students to measure their vocabulary breadth. There are 45 items to be completed within 20 minutes. Each of 45 items shows a target word in a phrase or short sentence and students have to choose which of four options is closest in meaning. The score was the number of correct answers.

**Word Definitions.** This task was based on Ouellette’s (2006) measure to assess students’ oral vocabulary and depth of vocabulary (see Appendix D). 12 words (bus, hand, apple, cow, tree,  

\(^1\) The other items were translated into Chinese and used to make a Chinese Listening Comprehension measure. However, this measure did not yield an acceptable level of reliability and so was dropped.
noodle, shirt, farm, television, island, drum, and envelope) were chosen from the Peabody Picture Vocabulary Test (Dunn & Dunn, 1997). Students were presented each word in both written and spoken form and were asked to provide definitions for the word. Students were encouraged to say as much as they could about each word. Responses were recorded by a digital recorder for later scoring.

Scoring was based on whether student’s definitions meet four features of the word: category, function, description, and value. For example, the definition of the word “apple” could include information about category (fruit), function (can be eaten), description (different colors, round shape, has skin and seeds, sour and sweet, grows on tree, etc.), and value (good for health, gives vitamin C, etc.). One point was given for each feature, to a maximum of 4 per word; the maximum total score was 36.

**Multiple-meaning Vocabulary.** This test was based on Qian’s (2002) depth-of-vocabulary-knowledge measure and Tannenbaum et al.’s (2006) Multiple Meanings test. It was used to measure students’ depth of vocabulary knowledge about the multiple meanings words can have in different contexts. There were 15 target words and each word was presented in 4 different sentence contexts (see Appendix E). Students were asked to choose the sentences in which the target word was used correctly. For example, for the target word “rose,” the sentences were: (a) The rose of the tree is very rough; (b) He bought a dozen roses for his girlfriend; (c) She is sitting in a rose watching TV; and (d) The sun rose over the mountains (the correct answers are b and d). The number of correct sentences per word varied from 0 to 4. The score was the number of correct answers selected and incorrect answers rejected, with a maximum score of 60.

**Morphological Awareness.** The Base Identification task from Bowers and Kirby (2010) was used to measure students’ depth of vocabulary knowledge about their ability to identify the
bases in multi-morphemic words presented in written form (see Appendix F). The students were asked to circle the smallest part (i.e., the base) which is related to the main word in meaning. The scoring was 2 for circling the smallest base, 1 for circling any word which removed at least one affix but did not identify the smallest base, and 0 for any other response. For example, in the word *scarred*, 0 points were given for *red* which is not a morpheme in *scarred*, 1 point was given for circling *scarr* because the student accurately peeled off a suffix *ed*, and 2 points was given for circling *scar* which is the base of *scarred*.

**Inference and Strategy.** Two 10-item multiple-choice tests based on two passages chosen from *HIP Reading Assessment* (Jamison, 2007) and *Reading Inventory for the Classroom* (Flynt & Cooter, 2004) were constructed to measure students’ ability to draw inferences and apply various reading strategies in reading (see Appendix G). The two tests were modelled on those used by Cromley and Azevedo (2007). After reading each passage, students answered 5 multiple-choice items that assessed coherence and elaborative inferences, and 5 multiple-choice items that required the application of various reading strategies, such as finding main ideas, predicting, inferencing, and summarizing. The score in each test was the number of correct answers.

**Chinese Reading Comprehension.** The Gates-McGinitie Reading Comprehension Test (MacGinitie & MacGinitie, 1992) Level E was translated into Chinese to measure students’ Chinese reading comprehension so that the Chinese reading measure would be similar to the English reading measure. The Chinese translation was also translated back into English to be compared to the original English version to ensure the adequacy of the translation. According to the manual, Level E is appropriate for Grade 7 to 9 English native-speaking children. The score was the number of correct answers.
**English Reading Comprehension.** The Gates-MacGinitie Reading Comprehension Test (Level C, Grade 3) (MacGinitie & MacGinitie, 1992) was used to measure students’ general reading comprehension. According to the test manual, Level C is used for Grade 3 EL1 students; we selected this level after consulting school staff members and carrying out a brief pilot study. Participants read 12 short passages and with the passages still in view answered multiple-choice questions within the 35-min time limit. There were 48 items and the score was the number of correct answers. The test manual indicates that some of the questions can be answered from information that is explicitly stated in the passage, which are considered literal questions; others require constructing an understanding based on information that is implicitly stated in the passage, which are considered inferential questions. Because the manual does not state which questions fall in each category, we classified the questions into literal and inferential levels. The literal items required readers to locate and recognize information that is presented in a direct way and comprehend the text at surface level, while inferential items required readers to link sentences in the text, draw inferences beyond the text, identify ideas implied by passage, and combine prior knowledge to passage information (Paris & Stahl, 2005). The literal and inferential items were classified by two raters; inter-rater reliability was .83. Any disagreements were resolved through discussion. 58% of the items were classified at the literal level (items 1, 2, 4, 7, 8, 10, 11, 13, 14, 15, 16, 21, 22, 24, 28, 30, 31, 32, 34, 35, 36, 37, 39, 40, 42, 44, 46, and 47) and 42% of the items at the inferential level.

**Summary Writing.** This task, based on Kirby and Pedwell’s (1990) and Stein and Kirby’s (1992) summary and recall tasks, was used to measure students’ deep understanding of text (see Appendix H). Before this task began, students were reminded that a good summary gives the main ideas of the text, that it is shorter, and that it is in one’s own words. The students
were informed that the text and any notes they took while reading the text would be collected before they wrote their summaries. They were then given 15 minutes to read a 254 word narrative passage entitled Cities that discussed the problems of cities in the Middle Ages. After a 5 minute interval, in which the original text and notes taken by the students were collected, they were given 10 minutes to write a text-absent summary. The students’ summaries were scored for the number of main ideas, which required readers to integrate details included in the text with a main idea, and the number of important details. The text was analyzed to have 5 main idea units and 19 important detail units; less important details were not included (see Appendix I). The maximum score was 24 points. One point was awarded for the inclusion of each unit. Twenty of the summaries were scored by two raters; inter-rater reliability was .92. Any disagreements were resolved through discussion, and the remaining summaries were scored by the first author.

**Procedure**

All 12 tests took place in the students’ school at times which were convenient to the school. There were four testers including the first author, all graduate students in Psychology or Education, and fluent in English and Chinese. They were trained for three days in the administration of all tests. Individual testing lasted about 50 minutes (1 session) during which the participant was tested by one tester in a quiet room. Group testing took approximately 200 minutes (spread over 4 sessions, spaced over several weeks), with a classroom teacher and two testers present in the classroom.

**Results**

**Descriptive Statistics**

The means, standard deviations, and reliability coefficients of all predictor and outcome measures are shown in Table 1. Measures whose skewness or kurtosis values fell outside of the
acceptable range (i.e., the absolute value of Skewness/SE or Kurtosis/SE >3.09) were transformed according to the guidelines of Tabachnick and Fidell (2007). Nonverbal ability and multiple-meaning vocabulary test were negatively skewed. They were first reflected and then a square root transformation was applied. They were then reflected again so that high scores represented better performance. These transformations eliminated the skewness problem, but, did not change their correlations with other variables substantially; therefore, all following analyses were conducted on the raw scores.

The reliabilities of morphological awareness, inference, and strategy were below 0.6 (see Table 1). We combined the inference and strategy measures because they were measured in the same way after the same passages, and inference is one aspect of reading strategies. The Cronbach’s alpha reliability of the combined test was .58. Although item analysis suggested minor improvements to these reliabilities, results based on revised scales were not meaningfully different from those with the original scales. Therefore we decided to maintain the original scales.
<table>
<thead>
<tr>
<th>Measures</th>
<th>M</th>
<th>SD</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonverbal Ability</td>
<td>50.60</td>
<td>4.78</td>
<td>0.79(^a)</td>
</tr>
<tr>
<td>Chinese Reading Comprehension</td>
<td>36.88</td>
<td>3.31</td>
<td>0.61(^a)</td>
</tr>
<tr>
<td>Decoding</td>
<td>67.07</td>
<td>6.32</td>
<td>0.93(^b)</td>
</tr>
<tr>
<td>Listening Comprehension</td>
<td>9.35</td>
<td>2.20</td>
<td>0.63(^b)</td>
</tr>
<tr>
<td>Gates-MacGinitie Vocabulary</td>
<td>19.28</td>
<td>5.43</td>
<td>0.74(^a)</td>
</tr>
<tr>
<td>Word Definition</td>
<td>25.22</td>
<td>4.51</td>
<td>0.68(^a)</td>
</tr>
<tr>
<td>Multiple-meaning</td>
<td>51.75</td>
<td>4.31</td>
<td>0.68(^a)</td>
</tr>
<tr>
<td>Morphological Awareness</td>
<td>50.80</td>
<td>5.27</td>
<td>0.54(^a)</td>
</tr>
<tr>
<td>Inference</td>
<td>6.76</td>
<td>1.87</td>
<td>0.49(^a)</td>
</tr>
<tr>
<td>Strategy</td>
<td>6.36</td>
<td>1.79</td>
<td>0.43(^a)</td>
</tr>
<tr>
<td>English Reading Comprehension - Literal</td>
<td>19.42</td>
<td>3.44</td>
<td>0.71(^a)</td>
</tr>
<tr>
<td>English Reading Comprehension - Inferential</td>
<td>11.78</td>
<td>3.23</td>
<td>0.54(^a)</td>
</tr>
<tr>
<td>Summary Writing - Detail</td>
<td>7.81</td>
<td>3.75</td>
<td>0.89(^c)</td>
</tr>
<tr>
<td>Summary Writing –Main Idea</td>
<td>1.94</td>
<td>1.42</td>
<td>0.92(^c)</td>
</tr>
</tbody>
</table>

Note. \( N = 246. \) \(^a\) indicates Cronbach’s alpha reliabilities; \(^b\) indicates split-half reliabilities; \(^c\) indicates inter-rater reliabilities
Correlations

The correlations between all measures are shown in Table 2. Decoding and listening comprehension are correlated significantly with both English reading comprehension and summary writing. The four vocabulary measures are all correlated significantly with English reading comprehension. Furthermore, the correlations between inference and strategy and both components of English reading comprehension are significantly different from those with either part of summary writing \( t(243) = 3.18 \) for difference of correlations between inference and strategy and English reading comprehension literal and correlations between inference and strategy and summary details, \( t(243) = 3.13 \) for difference of correlations between inference and strategy and English reading comprehension literal and correlations between inference and strategy and summary main ideas, \( t(243) = 2.45 \) for difference of correlations between inference and strategy and English reading comprehension inferential and correlations between inference and strategy and summary details, \( t(243) = 2.41 \) for difference of correlations between inference and strategy and English reading comprehension inferential and correlations between inference and strategy and summary main ideas, \( p < .05 \). The reading comprehension and summary writing measures are correlated \( (rs = .17 \text{ to } .42, \text{ all } ps < .01) \).

\[ t = \frac{(r_{12} - r_{13})\sqrt{(N-3)(1-r_{23})}}{\sqrt{2(1-r_{12}^2-r_{13}^2-r_{23}^2+2r_{12}r_{13}r_{23})}} \] was used to calculate the difference between two correlations for correlated samples, see Ferguson (1971), p.189).
### Table 2

*Correlations between all Measures*

<table>
<thead>
<tr>
<th>Measure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonverbal Ability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chinese Reading Comprehension</td>
<td>.34**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decoding</td>
<td>.01</td>
<td>.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English Listening Comprehension</td>
<td>.10</td>
<td>.15*</td>
<td>.31**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gates-MacGinitie Vocabulary</td>
<td>.20**</td>
<td>.17**</td>
<td>.18**</td>
<td>.17**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word Definition</td>
<td>.06</td>
<td>.05</td>
<td>.18**</td>
<td>.16*</td>
<td>.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple-meaning</td>
<td>.39**</td>
<td>.38**</td>
<td>.22**</td>
<td>.33**</td>
<td>.35**</td>
<td>.24**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morphological Awareness</td>
<td>.09</td>
<td>.22**</td>
<td>.14*</td>
<td>.18**</td>
<td>.01</td>
<td>.21**</td>
<td>.22**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inference and Strategy</td>
<td>.19**</td>
<td>.31**</td>
<td>.16*</td>
<td>.22**</td>
<td>.23**</td>
<td>.20**</td>
<td>.34**</td>
<td>.22**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading Comprehension – Literal</td>
<td>.40**</td>
<td>.36**</td>
<td>.18**</td>
<td>.32**</td>
<td>.41**</td>
<td>.12*</td>
<td>.50**</td>
<td>.22**</td>
<td>.39**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading Comprehension - Inferential</td>
<td>.17**</td>
<td>.11*</td>
<td>.14*</td>
<td>.32**</td>
<td>.34**</td>
<td>.20**</td>
<td>.27**</td>
<td>.15*</td>
<td>.37**</td>
<td>.55**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Summary Writing - Detail</td>
<td>.25**</td>
<td>.44**</td>
<td>.27**</td>
<td>.27**</td>
<td>.18**</td>
<td>.29**</td>
<td>.16*</td>
<td>.18**</td>
<td>.38**</td>
<td>.15*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Summary Writing - Main Idea</td>
<td>.22**</td>
<td>.35**</td>
<td>.21**</td>
<td>.29**</td>
<td>.17**</td>
<td>.22**</td>
<td>.35**</td>
<td>.22**</td>
<td>.19**</td>
<td>.42**</td>
<td>.21*</td>
<td>.75**</td>
<td></td>
</tr>
</tbody>
</table>

*Note. N = 246. **p < .01, *p < .05*
Exploratory Factor Analysis

Because there are no standardized or well-established vocabulary depth measures and the three vocabulary depth measures used in the present study were researcher-designed, it was necessary to verify the constructs they measure. An exploratory principal axis factor analysis with oblique (direct oblimin) rotation was conducted. Two factors were extracted after examining eigenvalues greater than 1.0, the scree plot, and interpretability. These two factors explained 65.77% of the total variance and were correlated $r = .51$, $p < .01$ (see Table 3).

Table 3

*Principal Axis Factor Analysis with Direct Oblimin Rotation of Vocabulary Measures*

<table>
<thead>
<tr>
<th></th>
<th>Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Word Definition</td>
<td>.06</td>
</tr>
<tr>
<td>Morphological Awareness</td>
<td>-.07</td>
</tr>
<tr>
<td>Multiple-meaning</td>
<td>.53</td>
</tr>
<tr>
<td>Gates-MacGinitie Vocabulary</td>
<td>.63</td>
</tr>
</tbody>
</table>

*Note. N = 246. Factor loadings greater than .30 are shown in bold type. The correlation between factors was .51 ($p < .01$).*

The three measures that were intended to measure vocabulary depth all loaded on factor 2, therefore this factor was labeled as *vocabulary depth*. The Gates-MacGinitie vocabulary test used to measure vocabulary breadth loaded highly on factor 1. However, the Multiple-meaning vocabulary measure which was intended to measure vocabulary depth also loaded on factor 1. This suggests that Multiple-meaning vocabulary assesses both vocabulary depth and breadth.
Factor 1 is thus labeled as *vocabulary breadth*. Regression factor scores were calculated and used in the subsequent analyses.

**Prediction of Reading Comprehension**

A series of hierarchical regression analyses was conducted to assess the contribution of the predictors to English reading comprehension.

First, the simple view of reading model was examined (see Table 4). Decoding measured by word reading efficiency and linguistic comprehension measured by listening comprehension significantly predicted the literal level of English reading comprehension and both levels of the summary writing. Linguistic comprehension explained more variance in both levels of English reading comprehension and in the main idea level of summary writing than did decoding. However, the variance accounted for by these two predictors only ranged between 10-11%, which is considerably lower than in EL1 studies.
Table 4

*Summary of Hierarchical Regression Analyses Predicting Different Levels of Two Reading Comprehension Measures from the Simple View of Reading Model (N = 246)*

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Reading Comprehension</th>
<th>Summary Writing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Literal</td>
<td>Inferential</td>
</tr>
<tr>
<td>β R²</td>
<td>β R²</td>
<td>β R²</td>
</tr>
<tr>
<td>Decoding</td>
<td>.10*</td>
<td>.05</td>
</tr>
<tr>
<td>Listening</td>
<td>.11**</td>
<td>.11**</td>
</tr>
<tr>
<td>Comprehension</td>
<td>.29**</td>
<td>.31**</td>
</tr>
</tbody>
</table>

*Note. **p < .01, *p < .05*

Next, a series of hierarchical regression analyses was conducted to examine the contributions of inference and strategy and vocabulary breadth and depth to the two levels of the two reading comprehension outcomes (see Table 5). In the first step of these analyses, Nonverbal ability and L1 reading comprehension were entered to control for background ability; together they accounted for 3 to 22% of the variance. In the second step, decoding and listening comprehension were added; the results here were similar to those shown in Table 4, though generally slightly weaker. When the inference and strategy score was entered in step 3, it predicted both literal and inferential levels of reading comprehension but had no effect on either level of the summary writing task. When inference and strategy was added later (at step 5, after the vocabulary measures), the same pattern was found, though again somewhat weaker in strength. The two vocabulary scores, breadth and depth, were entered initially at steps 4 and 5, then in the opposite order in steps 4A and 5A, later at steps 3’ and 4’, and in the opposite order in steps 3’A and 4’A. Vocabulary breadth predicted both literal and inferential levels of reading...
comprehension over and above vocabulary depth, whereas vocabulary depth only predicted the main idea level of summary writing after controlling vocabulary breadth; this pattern remained constant regardless of whether the vocabulary scores were entered before or after inference and strategy.
Table 5

Summary of Hierarchical Regression Analyses Predicting Different Levels of Two English Reading Comprehension Measures (N = 246)

<table>
<thead>
<tr>
<th>Predictors/Models</th>
<th>RC Literal</th>
<th></th>
<th></th>
<th></th>
<th>RC Inferential</th>
<th></th>
<th></th>
<th></th>
<th>SW Details</th>
<th></th>
<th></th>
<th></th>
<th>SW Main Ideas</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>β</td>
<td>Δ R²</td>
<td>β</td>
<td>Δ R²</td>
<td>β</td>
<td>Δ R²</td>
<td>β</td>
<td>Δ R²</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Nonverbal</td>
<td>.32**</td>
<td>.14*</td>
<td>.12*</td>
<td>.11*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chinese RC</td>
<td>.25**</td>
<td>.22**</td>
<td>.07</td>
<td>.03*</td>
<td>.40**</td>
<td>.20**</td>
<td>.31**</td>
<td>.13**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Decoding</td>
<td>.11</td>
<td>.05</td>
<td>.21**</td>
<td>.14*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LC</td>
<td>.23**</td>
<td>.08**</td>
<td>.29**</td>
<td>.10**</td>
<td>.14*</td>
<td>.08**</td>
<td>.20**</td>
<td>.08**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Inference &amp; Strategy</td>
<td>.23**</td>
<td>.04**</td>
<td>.30**</td>
<td>.08**</td>
<td>.00</td>
<td>.00</td>
<td>.02</td>
<td>.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Vocabulary Breadth</td>
<td>.33**</td>
<td>.08**</td>
<td>.24**</td>
<td>.04**</td>
<td>.13*</td>
<td>.01*</td>
<td>.14*</td>
<td>.01*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Vocabulary Depth</td>
<td>.02</td>
<td>.00</td>
<td>.05</td>
<td>.00</td>
<td>.06</td>
<td>.00</td>
<td>.22**</td>
<td>.03**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4A. Vocabulary Depth</td>
<td>.13*</td>
<td>.01*</td>
<td>.12*</td>
<td>.01*</td>
<td>.10</td>
<td>.00</td>
<td>.24**</td>
<td>.04**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5A. Vocabulary Breadth</td>
<td>.32**</td>
<td>.06**</td>
<td>.22**</td>
<td>.03**</td>
<td>.11</td>
<td>.01</td>
<td>.06</td>
<td>.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3’. Vocabulary Breadth</td>
<td>.36**</td>
<td>.10**</td>
<td>.29**</td>
<td>.06**</td>
<td>.13*</td>
<td>.01*</td>
<td>.14*</td>
<td>.01*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4’. Vocabulary Depth</td>
<td>.05</td>
<td>.00</td>
<td>.10</td>
<td>.01</td>
<td>.05</td>
<td>.00</td>
<td>.21**</td>
<td>.03**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3’A. Vocabulary Depth</td>
<td>.17**</td>
<td>.02**</td>
<td>.18**</td>
<td>.03**</td>
<td>.09</td>
<td>.01</td>
<td>.23**</td>
<td>.04**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4’A. Vocabulary Breadth</td>
<td>.34**</td>
<td>.07**</td>
<td>.26**</td>
<td>.04**</td>
<td>.11</td>
<td>.01</td>
<td>.05</td>
<td>.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5’. Inference &amp; Strategy</td>
<td>.16**</td>
<td>.02**</td>
<td>.25**</td>
<td>.05**</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td>.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. **p < .01, *p < .05, RC = English Reading Comprehension, SW = Summary Writing
β is the standardized beta coefficient for the step at which the predictor first entered the model

Hierarchical regression analysis was employed again to assess the prediction of breadth and depth vocabulary to word reading. Nonverbal ability and Chinese (L1) reading
comprehension were controlled in the first step. Vocabulary breadth and depth was entered in steps 2 and 3. Then, this order was alternated to see what each vocabulary dimension contributed over and above the other in steps 2A and 3A. The results, shown in Table 6, indicate that the control variables do not account for a meaningful amount of variance in word reading. Vocabulary breadth and depth each predict word reading, with breadth being slightly stronger.

Table 6

Summary of Hierarchical Regression Analyses Predicting English Word Reading from Breadth and Depth of Vocabulary (N = 246)

<table>
<thead>
<tr>
<th>Predictors</th>
<th>English Word Reading</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>β</td>
</tr>
<tr>
<td>1. Nonverbal Ability</td>
<td>.00</td>
</tr>
<tr>
<td>Chinese Reading Comprehension</td>
<td>.04</td>
</tr>
<tr>
<td>2. Vocabulary Breadth</td>
<td>.30**</td>
</tr>
<tr>
<td>3. Vocabulary Depth</td>
<td>.17*</td>
</tr>
<tr>
<td>2A. Vocabulary Depth</td>
<td>.26**</td>
</tr>
<tr>
<td>3A. Vocabulary Breadth</td>
<td>.23**</td>
</tr>
</tbody>
</table>

Note. **p < .01, *p < .05

β is the standardized beta coefficient for the step at which the predictor first entered the model.

Discussion

This study explored the predictors of ESL reading comprehension. We investigated how well the two predictors of the SVR model, decoding and linguistic comprehension, account for reading comprehension in Chinese ESL students, and then whether or not other predictors such as
inference and reading strategies and vocabulary should be added to the model. We examined these effects at different levels of the outcome measures, one literal or focused on details, and the other more inferential or focused on main ideas. The research questions are addressed in turn.

**How Well Do Decoding and Linguistic Comprehension Predict English Reading Comprehension?**

Because reading comprehension is complex and multifaceted, researchers should not rely on a single measure of it (Carpenter & Paris, 2005; Kintsch, 1998; Pressley & Hilden, 2005; Tapiero, 2007). Studies have shown that different measures of reading comprehension tap different cognitive processes (Cutting & Scarborough, 2006; Francis, Fletcher, Catts, & Tomblin, 2005; Keenan, Betjemann, & Olson, 2008; Nation & Snowling, 1997). Furthermore, any one reading comprehension test can measure different levels of reading comprehension (Kirby & White, 2011). In this study, we employed two reading comprehension measures. One involved silent passage reading with multiple-choice questions, which we classified as being at the literal or inferential levels. The other was a text-absent summary writing task which required deeper or more extensive processing of the text to extract main ideas and form a situation model (Kintsch, 1998) as a basis for summary writing; we classified idea units into details and main ideas. Compared to the details level, the main ideas level would require deeper processing as students link ideas and integrate information from the text (Kirby & Woodhouse, 1994).

When we tested the original “simplest” view of reading model, consisting only of decoding and linguistic comprehension constructs, only 10% to 11% of the variance in two levels of two reading comprehension measures was accounted for. These results are low compared to the 30% to 70% found in previous studies with L1 readers and L2 readers whose L1 is a related language (e.g., Adlof et al., 2006; Conners, 2009; Gottardo & Mueller, 2009; Johnston & Kirby,
We employed a variety of other regression models to explore the roles of other constructs in predicting this “missing” variance.

Not surprisingly, nonverbal ability and Chinese reading comprehension accounted for an initial 6 to 20% of the variance. Reading comprehension should be related to general ability, and those who are more skilled readers in their first language should have an advantage in learning to read a second. However, these may be underestimates of the effects of these variables; because of the selection process used to accept students into the school, students with less ability or motivation may have been screened out, reducing the range of variation within our sample. In the original SVR, the effect of general ability would be subsumed under the linguistic comprehension term.

One question is whether the decoding and linguistic comprehension measures we used were adequate. Previous studies examining the SVR model in L2 readers whose L1 is also an alphabetic language have included L2 phonological awareness (e.g., Gottardo & Mueller, 2009; Yaghoub Zadeh et al., 2012), L2 naming speed (e.g., Yaghoub Zadeh et al., 2012), or L2 reading fluency (e.g., Proctor et al., 2005; Yaghoub Zadeh et al., 2012) to explain part of the decoding construct. Because our participants were older and more skilled in English, we only used a word reading speed measure of decoding; it is likely that adding measures of skills that contribute to word reading, such as phonological awareness, would have improved our prediction of reading comprehension.

To measure linguistic comprehension, we used measures of listening comprehension, as suggested by the original SVR model. However, our listening comprehension did not make a large contribution to reading comprehension. This may have been because the reliability of our
listening comprehension measure was not very high ($\alpha = 0.63$). We had divided the Woodcock Listening Comprehension measure into two halves, making parallel English and Chinese listening comprehension measures; in the end, we did not use the Chinese measure, because of unacceptably low reliability. In retrospect, a more reliable and more comprehensive assessment of linguistic comprehension would have been helpful.

In spite of low reliability, listening comprehension made a larger contribution to reading comprehension than decoding (see Tables 4 and 5), as has been found in previous studies of EL1 children in upper elementary grades (Catts, Hogan, & Adolf, 2005; Torgesen, Wagner, & Rashotte, 1997) and ESL children (Gottardo & Mueller, 2009; Proctor et al., 2005). Moreover, decoding explained more variance in the literal or details level than in the inferential or main ideas level of reading comprehension but listening comprehension accounted for more variance in the inferential or main ideas level than in the literal or details level of reading comprehension (Tables 4 and 5). These results shows that deeper processing requires more higher-level skills while surface level processing depends more on basic level skills.

**Should the SVR Model Include Inference and Strategy?**

Although it has been well accepted that inference and strategy exert important influence on reading comprehension, there has been no research examining their roles in the SVR model, perhaps because it has been assumed that these factors are included in the linguistic comprehension term. We gave reading passages with multiple-choice questions to measure how well students drew inferences and applied reading strategies in comprehending passages. The results showed that inference and strategy significantly predicted both literal and inferential English reading comprehension measure after controlling nonverbal IQ, L1 reading comprehension, L2 decoding and listening comprehension, and vocabulary (see Table 5). This is
consistent with literature illustrating that reading strategies contributed significantly to L2 reading comprehension (Cohen & Macaro, 2007; Hurd & Lewis, 2008; Phakiti, 2003; Purpura, 1997), and is also in line with L1 studies indicating that text inferences are required for successful reading comprehension (Cain et al., 2001, 2004; van den Broek et al., 1995).

Our inference and strategy measures were designed based on those of Cromley and Azevedo (2007), and like them we found they predicted Gates-MacGinitie reading comprehension, which is a measure of silent reading passages with multiple-choice questions. We further classified the Gates-MacGinitie reading comprehension measure into literal and inferential levels. The contributions of inference and strategy to inferential reading comprehension were larger than to literal reading comprehension, supporting their role in deeper processing of text. However, inference and strategy did not contribute to either the detail or main idea level of summary writing. Text-absent summary writing is a measure which should in principle require deeper processing and more inference and strategy than the Gates-MacGinitie measure. Because inference and strategy were measured in the same way as literal and inferential reading comprehension in the Gates-MacGinitie, through reading passages with multiple-choice questions, this suggests the inference and strategy findings are test-specific and the relationship may reflect shared method variance. More broadly-based inference and strategy measures are required to generalize the importance of inference and strategies in predicting reading comprehension after accounting for other factors in L2 students.

So, should the SVR model include inference and strategy? Even if linguistic comprehension represents all verbal ability and it is the process by which words, sentences, and discourses are interpreted (Gough & Tunmer, 1986; Kirby & Savage, 2008), written text provides further opportunities for inference and strategy. A more comprehensive measure of listening
comprehension may be more successful in assessing inference and strategy, but we suggest that there will continue to be an additional contribution made by inference and strategy regarding written text. More comprehensive measures of both linguistic comprehension and inference and strategy in reading are required.

**What are the Effects of Breadth and Depth of Vocabulary on English Reading Comprehension?**

It is important to consider the relationship between breadth and depth of vocabulary before we look into their effects on reading comprehension. The correlation between the vocabulary breadth and depth factors was .51 ($p < .01$), which, although not as high as the .8 or .9 reported in some studies (Qian, 1999; Vermeer, 2001) between individual measures, demonstrates that there is substantial overlap between these two constructs. Instead of being two different entities, breadth and depth of vocabulary are better seen as two dimensions of the same construct which facilitate each other. When language development begins, a small number of words can be recognized and known in terms of basic meanings. Knowledge of more and more words accumulates, providing some breadth of vocabulary. With increased experience, these words can be defined in greater detail and are associated with other words in different contexts, providing greater depth of vocabulary knowledge. This greater depth then supports the learning of other new words. Therefore, breadth contributes to depth and vice versa. Both are acquired by extensive exposure to language. We suggest that these two dimensions are interdependent, developing together, and influenced by extensive exposure to language.

Vocabulary breadth significantly predicted both literal and inferential English reading comprehension after controlling vocabulary depth, nonverbal ability, L1 reading, decoding, and listening comprehension. In contrast, vocabulary depth predicted unique variance in main idea
level summary writing, after considering the effects of vocabulary breadth and the other control variables (see Table 5). The relationship between vocabulary breadth and the multiple-choice reading comprehension test may be due to the relatively constrained nature of the latter; even though inferential questions are involved and were significantly predicted by vocabulary breadth, the passages are relatively brief and remain in sight while the questions are answered. Thus participants may use pragmatic test-taking strategies, such as reading the questions first and then searching the passage for key words, to avoid creating a situation model or even much of a macrostructural hierarchy of propositions.

The vocabulary depth measures required students to provide definitions of words, choose correct usages, and identify the morphological bases of words, none of which could be done if they did not know the words well. The detail level of summary writing may only require relatively shallow vocabulary knowledge (i.e., breadth), while the main idea level which represents deeper processing in reading comprehension (Kirby, Cain & White, 2012), may demand deeper vocabulary knowledge. The multiple-choice format with passages visible of the standardized reading comprehension test may not put as much stress on the richness of vocabulary knowledge as a summary-writing task in which words must be found to explain one’s situation model. Superficial knowledge of words, which may suffice on a vocabulary breadth test and for many multiple-choice comprehension items, would not be adequate for comprehending longer texts, constructing integrated long-term memory structures, and then describing those structures in a written summary. Therefore, different reading comprehension measures rely on different skills and different dimensions of vocabulary knowledge. When researchers choose reading comprehension measures, they should pay attention to what skills they are measuring because different reading comprehension tests measure different skills.
What are the Effects of Breadth and Depth of Vocabulary on English Word Reading?

The fourth research question, about the effects of breadth and depth of vocabulary on word reading, has been examined in EL1 reading, but not in ESL reading. Both vocabulary breadth and depth contributed to word reading but breadth explained more variance in word reading after controlling for depth than vice versa (see Table 6). This is somewhat inconsistent with Ouellette’s (2006) finding that vocabulary breadth significantly predicted word reading after controlling vocabulary depth but depth did not make a significant contribution after controlling breadth. The difference may be due to the measures and participants involved. Ouellette (2006) used word definition as the vocabulary depth measure, but we also used multiple-meaning and morphological awareness tests to measure vocabulary depth. Ouellette’s participants were Grade 4 English-speaking children, but our participants were Grade 8 Chinese ESL children. Ouellette and Beers (2010) found that vocabulary depth predicted word reading significantly after controlling vocabulary breadth in Grade 1 but not in Grade 6 EL1 children. They suggested that vocabulary depth may be the more sensitive predictor for younger children who may tend to store new words rapidly without a full and deep meaning of each word. Although our participants were older, compared to Ouellette and Beers’ participants, as ESL students, they may still be less able to learn many new words with deep understanding, because of their reduced exposure to English. As a result, vocabulary depth turns out to be a significant predictor of word reading for Grade 8 ESL students. There has been no research examining the effects of vocabulary breadth and depth on word reading in ESL students. As far as we know, this study is the first to investigate the effects of vocabulary breadth and depth on word reading in L2 learners.
Limitations

Several limitations should be noted. First, the reliabilities of four measures, Chinese reading comprehension, English listening comprehension, morphological awareness, and inference and strategy, were low ($\alpha = .61, .63, .54,$ and $.58,$ respectively). Item analysis did not improve reliabilities greatly. It is not clear why these reliabilities were low, but it may be that bilingual participants approach such measures, which were designed for EL1 students, in different ways, making the test items less consistent measures of the same construct. For instance some students may rely more on L1 skills than L2 skills for some measures, and may switch approach from item to item. Given the large number of Chinese students learning English as a second language, there is a need to investigate how these students process individual tests and items, and a further need to develop standardized measures of key abilities. Second, only word reading fluency was used to measure decoding in the present study. Tunmer and Greaney (2009) have suggested that a composite measure of pseudoword reading, context-free word recognition, and timed word recognition would be the best measure of decoding. Therefore, future research should include multiple and more extensive measures to assess decoding. Third, only listening comprehension was used to measure the linguistic comprehension construct. As linguistic comprehension is a process that involves word, sentence, and text-level skills (Gough & Tunmer, 1986; Kirby & Savage, 2008), more comprehensive measures are needed to obtain a more complete understanding of it.

Conclusions and Implications

This study investigated the cognitive predictors of reading comprehension in Chinese ESL students. The findings suggest that the two predictors in the SVR model do not predict ESL reading comprehension very well. The complexity of ESL reading comprehension involves many
cognitive skills that operate in two languages. Although inference and strategy uniquely predicted
English reading comprehension beyond the SVR model, further research is needed to develop
more broadly-based inference and strategy measures. Both breadth and depth of vocabulary made
contributions to word reading and reading comprehension, but their contributions differed
depending on what level of reading comprehension was being predicted. We suggest that both
aspects of vocabulary knowledge are required for adequate ESL word reading and reading
comprehension; deeper processing of text may require vocabulary depth, whereas general
understanding of text may rely more on vocabulary breadth.

The findings from this study have both theoretical and practical implications.
Theoretically, the findings of the current study test the application of EL1 models of reading
comprehension to ESL students, especially those whose L1 does not share similarities with their
L2. The results reinforce the need to build a comprehensive reading comprehension model for
ESL reading. Practically, our findings suggest that ESL reading teachers should focus on
improving students’ decoding skills and oral language skills, especially vocabulary knowledge.
ESL learners clearly need to begin with relatively superficial knowledge about new words
(breadth), but the sooner depth can be added the better. Developing depth may be more
challenging in ESL learners, because of their lack of exposure to the L2. This is a challenge that
educators need to accept and address. Successful reading comprehension is complex, taking place
at a variety of cognitive levels. It requires a complex set of oral language, word reading, and
strategic skills to operate efficiently. As the number of L2 readers increases, much research and
development are required to improve the ways we teach and assess reading comprehension and
the skills that support it.
Chapter 3

Unexpected Poor Comprehenders among Adolescent ESL Students

Miao Li and John R. Kirby
Queen’s University, Canada

Abstract
This study explored characteristics of reading comprehension difficulties among Chinese students learning English as a second language (ESL). Two hundred and forty-six Grade 8 English-immersion students in China were administered a battery of reading-related and reading comprehension tests. Three groups of comprehenders matched on age, nonverbal intelligence, and word reading speed were identified: unexpected poor comprehenders, expected average comprehenders, and unexpected good comprehenders. The three groups differed in vocabulary and higher-level processes. Vocabulary breadth and depth distinguished between the unexpected poor comprehenders and the expected average comprehenders. Inference, strategy, listening comprehension, summary writing, and morphological awareness distinguished between the expected average comprehenders and the unexpected good comprehenders. The findings suggest that vocabulary is the main source of reading comprehension difficulties in ESL unexpected poor comprehenders. The advantage of the unexpected good comprehenders group is primarily due to discourse comprehension and strategic processes, and is only possible with high language proficiency.
Reading comprehension is a complex process that is essential to academic success. The National Reading Panel (2000) reported that basic skills (e.g., phonological awareness, decoding, fluency, and vocabulary knowledge) are important for successful reading comprehension. At the same time, it has become clear that higher-level reading skills (e.g., inferencing skills, comprehension monitoring, integration of text information) are also essential (Oakhill, Cain, & Bryant, 2003; Kirby, Cain, & White, 2012; Snow, 2002; van den Broek et al., 2005).

Research on English as a second language (ESL) has been growing rapidly in recent years because more and more people are learning a second language (L2). However, most ESL studies have focused on literacy skills (e.g., Chen-Bumgardner, 2008; Geva, 2006; McBride-Chang, Shu, Ng, Meng, & Penney, 2006; Wang, 2011), oral language skills (e.g., Ellis, 1994), sociocultural factors (e.g., Jiménez, 2000; Pucci & Ulanoff, 1998), metacognitive skills (e.g., Carlisle, Beeman, Davis, & Spharim, 1999; Phakiti, 2008), and strategy use (e.g., Jiménez, García, & Pearson, 1995; Purpura, 1999). Comparatively little research has explored ESL reading comprehension difficulties and very few studies have been done to investigate why some ESL students have problems in reading comprehension in spite of normal word reading ability (Lesaux & Kieffer, 2010). In the upper elementary and middle school years, the need for higher levels of reading skill is a challenge to ESL students and many of them experience reading comprehension difficulties and, in turn, academic failure (Murnane & Levy, 2004). Most of the empirical research based on the nature of reading comprehension difficulties is limited to English as a first language (EL1) children and children in elementary grades. The population of ESL older students with reading comprehension difficulties has been ignored and underserved (August & Shanahan, 2006).
The present study aims to explore the nature of reading comprehension difficulties in a sample of adolescent ESL students in middle school. We are particularly interested in ESL students who have adequate word reading skills but still have difficulties in reading comprehension. We investigate the characteristics of reading comprehension difficulties among these students. Recent studies have identified native English speaking children with reading comprehension difficulty in the absence of word reading problems (Cain, Oakhill, & Lemmon, 2004; Nation & Snowling, 1998, 1999; Oakhill et al., 2003). These children are called poor comprehenders or less skilled comprehenders or unexpected poor comprehenders by different researchers (Cain, Oakhill, & Lemmon, 2004; Nation & Snowling, 1998, 1999; Tong, Deacon, Kirby, Cain, & Parrila, 2011; White & Kirby, 2008). In this study, the term unexpected poor comprehenders is used because these children have unexpectedly poor reading comprehension given that their word reading ability is normal.

**Unexpected Poor Comprehenders in EL1 Children**

Research conducted with EL1 children with reading comprehension difficulties has demonstrated different profiles. Some children are weak in lower level skills such as phonological awareness, decoding, and semantic processing (e.g., Nation & Snowling, 1998, 1999; Perfetti, 1988; Shankweiler, 1989). Some are weak in higher level comprehension skills such as working memory, inferencing, and integration (e.g., Cain & Oakhill, 1999; Cain, Oakhill, & Bryant, 2004), and some are weak in both (e.g., Buly & Valencia, 2002; Catts, Hogan, & Adlof, 2005; Hock et al., 2009). For example, Catts et al. (2005) reported that in a sample of 152 poor readers in Grade 8, 36% of the students had word reading difficulties but adequate listening comprehension, 15% had adequate word reading but poor listening comprehension skills, and 36% had difficulties in both word reading and listening comprehension skills. Furthermore, some
children demonstrate reading difficulties in later elementary school year but not in the initial stages of reading acquisition (e.g., Compton, Fuchs, Fuchs, Elleman, & Gilbert, 2008; Leach, Scarborough, & Rescorla, 2003; Lipka, Lesaux, & Siegel, 2006). A heterogeneous profile was found in these late-emerging reading disability students who were not identified as at risk of reading until fourth grade. For example, Leach et al. (2003) reported that, in their sample of fourth- and fifth-grade late-emerging reading disability students, 35% showed word-level processing deficits with normal comprehension skills, 32% had weak comprehension skills but good word reading skills, and 32% had difficulties in both word-level and comprehension skills.

Investigations into the sources or correlates of specific comprehension difficulties in unexpected poor comprehenders in EL1 children have been carried out in recent years. However, most of the findings showed the correlates rather than the sources of reading comprehension difficulties in unexpected poor comprehenders. These studies have compared unexpected poor comprehenders to the age-matched good comprehenders in some reading skills, which do not demonstrate the causal relationship between the reading skills and reading comprehension. Only longitudinal studies, intervention studies, and reading-age-matched comparisons can determine whether the weakness in one reading skill is the cause or source of difficulties in reading comprehension (Cain, Oakhill, & Bryant, 2000). Cain, Oakhill, and colleagues have matched their unexpected poor comprehenders to skilled comprehenders on age, word reading accuracy, and vocabulary. They have found that unexpected poor comprehenders, relative to controls, have difficulties in higher level skills, such as inferencing, integration of text information, working memory, comprehension monitoring, and strategy use (see Cain & Oakhill, 1999, 2006; Cain, Oakhill, Barnes, & Bryant, 2001; Cain, Oakhill, & Elbro, 2003; Cain, Oakhill, & Lemmon, 2004). Cain and Oakhill (1999) even compared a group of 7-8 year old unexpected poor
comprehenders with a group of good comprehenders, matched for age and word reading accuracy, and a group of younger children of equivalent comprehension ability, the reading-age matched group. They found that unexpected poor comprehenders performed more poorly than the good comprehender and reading-age matched groups on different types of inferences. They interpreted the reading-age matched group’s superior performance on inference tasks as indicating that low inferencing skill causes reading comprehension failure. Compared to good comprehenders, unexpected poor comprehenders were even poorer at making inferences when knowledge base was equal to both groups (Cain et al., 2001) and they were deficient at tasks depending more on working memory (Cain et al., 2000). Cain and colleagues have also examined the role of vocabulary in reading comprehension for poor comprehenders. For example, Cain et al. (2003) found that children with poor reading comprehension were less able to infer the meanings of unfamiliar words embedded in text. Cain, Oakhill, and Lemmon (2004) also found that poor comprehenders with weak vocabulary knowledge were less able to learn new vocabulary from direct instruction.

Other researchers, Nation, Snowling and colleagues, have matched their unexpected poor comprehenders to controls with good reading comprehension on decoding skill and performance IQ (Nation, Clarke, Marshall, & Durand, 2004; Nation & Snowling, 1998, 1999; Nation, Snowling, & Clarke, 2005, 2007). They have found that the unexpected poor comprehenders, relative to controls, have weaker semantic skills. For example, Nation and Snowling (1998) compared a group of 9-10 year old unexpected poor comprehenders with a control group matched for decoding skills and nonverbal ability and found that unexpected poor comprehenders performed more poorly than control children on semantic tasks of receptive and expressive vocabulary, synonym judgement, and semantic fluency. This finding is consistent with the lexical
quality hypothesis (Perfetti, 2007; Perfetti & Hart, 2001), which argues that reading comprehension depends on the quality of the lexical representations of words. Nation and colleagues have also found that the unexpected poor comprehenders have difficulties in morphological processing (Nation et al., 2004, 2005). For example, Nation et al. (2005) found that eight-year old unexpected poor comprehenders, matched to normal comprehenders for age and nonword reading, were less able to inflect irregular words (producing past tense forms of verbs). Nation and colleagues suggested that the weaknesses of semantic and morphological processing were due to the poor comprehenders’ weak language skills. Tong et al. (2011) examined the sources of reading comprehension difficulties in Grade 3 and Grade 5 students by identifying unexpected poor comprehenders, expected average comprehenders, and unexpected good comprehenders who performed similarly on reading-related skills such as phonological awareness, naming speed, and orthographic processing at both grade levels. They found that these three groups performed differently on morphological awareness, with the unexpected poor comprehenders performing worst, which was consistent with Nation et al.’s (2004, 2005) finding that poor comprehenders’ weak morphological processing may be the source of their reading comprehension difficulties.

Some researchers have also found that the problems of reading comprehension difficulties of unexpected poor comprehenders are related to general language comprehension (Catts, Adlof, & Weismer, 2006; Compton et al., 2008; Hock et al., 2009). For example, Catts et al. (2006) identified 57 Grade 8 unexpected poor comprehenders who scored below the 25th percentile in reading comprehension and above the 40th percentile in word recognition. The unexpected poor comprehenders performed poorly on vocabulary, grammatical understanding, and discourse inference compared to typical readers and poor decoders. Hock et al. (2009) found
that 12% of their sample of grade 8 and 9 students had difficulties in reading comprehension but not in word-level and these unexpected poor comprehenders scored significantly lower than the proficient readers in vocabulary and fluency.

In sum, the evidence suggests that the possible sources of reading comprehension difficulties of unexpected poor comprehenders are related to children’s semantic and morphological processing skills on the one hand, and higher-level processes such as inferencing skills, working memory, comprehension monitoring, and strategy use on the other. These characteristics can all be seen as aspects of general language comprehension. The Simple View of Reading Model states that reading comprehension can be explained by word reading and language comprehension (Gough & Tunmer, 1986). If unexpected poor comprehenders by definition have normal word recognition abilities, it is reasonable that the sources of their reading comprehension difficulties should be in the area of language comprehension. However, the matching technique in the majority of the reviewed studies may not be as reliable as the one based on the regression method used by Tong et al. (2011). Whereas most previous studies have selected unexpected poor comprehenders with word reading scores above some cutoff score and reading comprehension scores below some cutoff score, Tong et al. (2011) used regression to identify them. Reading comprehension was predicted from age, nonverbal cognitive ability, word reading accuracy, and word reading speed, and this predicted value was plotted against reading comprehension scores. Students well below the regression line were identified as unexpected poor comprehenders, those well above the regression line as unexpected good comprehenders, and those close to the regression line as expected average comprehenders. The comparison of these three groups provides us with a clearer understanding of the characteristics of unexpected poor comprehenders because it examines relative as opposed to absolute discrepancies between
skills and because it distinguishes between expected average comprehenders (the most appropriate comparison group) and unexpected good comprehenders (who may differ from expected average comprehenders in other ways).

**Reading Comprehension Difficulties in ESL Children**

Research on reading comprehension difficulties in EL1 children has demonstrated that unexpected poor comprehenders have low performance in semantic and morphological processing skills as well as higher-level processes. However, it is not clear whether the model can be applied to ESL children. There have not been many studies in ESL reading difficulties to support the unexpected poor comprehenders model. Testing the model of EL1 unexpected poor comprehenders in ESL students can discover whether they have the same deficits in reading skills as EL1 unexpected poor comprehenders. This would help educators understand ESL reading comprehension difficulties and guide them in designing effective interventions.

There have been growing concerns about reading comprehension difficulties in ESL children (Kieffer, 2010; Lesaux & Kieffer, 2010; Lesaux, Lipka, & Siegel, 2006; Samson & Lesaux, 2009). These studies compared ESL children to EL1 children in terms of reading-related skills and suggested that ESL children experienced reading comprehension difficulties similar to the EL1 children and the characteristics of reading comprehension difficulties were also similar for the two groups.

However, only two studies have investigated the characteristics of ESL unexpected poor comprehenders. Lesaux et al. (2006) examined reading performance of Grade 4 students and found three groups in both ESL and EL1 children: poor readers who were poor at word reading and reading comprehension, unexpected poor comprehenders who had difficulties in reading comprehension in the absence of word reading difficulties, and good comprehenders who were
skilled in word reading and reading comprehension. There were more students identified as unexpected poor comprehenders (21.2% of the ESL students) than poor readers (4.7% of the ESL students) among ESL students, which was the same pattern for EL1 students. Lesaux et al. (2006) found that ESL students performed poorly on syntactic awareness and verbal working memory compared to their EL1 peers within the good comprehender and unexpected poor comprehender groups.

In another study, Lesaux and Kieffer (2010) explored the possible sources of reading comprehension difficulties among ESL students and their peers in Grade 6. Two hundred and one ESL and 61 EL1 students who all performed below the 35th percentile on a standardized reading comprehension measure were selected as the struggling comprehenders and were administered oral language and reading measures. Latent class analysis demonstrated three profiles which were evenly distributed between ESL and EL1 students. One of the three profiles was characterised by above-average word reading accuracy and average fluency, and these students were considered unexpected poor comprehenders. The unexpected poor comprehenders demonstrated low vocabulary and semantic working memory scores, which were considered the possible sources of ESL students’ reading comprehension difficulties (Lesaux & Kieffer, 2010).

Studies of reading comprehension difficulties in ESL children suggest that vocabulary, syntactic awareness, and working memory are the main sources of reading comprehension difficulties in unexpected poor comprehenders. The findings of these studies are consistent with those of EL1 reading comprehension difficulties studies. However, none of these studies investigated higher-level processes such as inferencing and strategies in the ESL unexpected poor comprehenders. Similar to EL1 reading comprehension difficulties studies, the technique for
selecting unexpected poor comprehenders may not have been as reliable as the one based on the regression used by Tong et al. (2011).

The Present Study

The purpose of this study was to explore the possible sources of reading comprehension difficulties of Chinese ESL unexpected poor comprehenders. There have been no studies examining the characteristics of Chinese ESL unexpected poor comprehenders. Given that there are a large number of Chinese students learning English in China, there is a need to explore the nature of English reading comprehension difficulties in these students and investigate both lower-level and higher-level reading skills in these students. Previous studies of EL1 unexpected poor comprehenders in elementary schools have shown that semantic and morphological processing or higher-order skills such as inference, working memory, and comprehension monitoring might be the sources of poor reading comprehension (Cain & Oakhill, 1999; Cain et al., 2003; Catts et al., 2006; Hock et al., 2009; Nation & Snowling, 1998, 1999; Nation et al., 2005; Tong et al., 2011). Relatively little research has examined ESL unexpected poor comprehenders. Research on the literacy development of ESL learners has indicated that although they develop decoding skills equivalent to L1 peers, they usually lag behind them in reading comprehension (Lesaux, Koda, Siegel, & Shanahan, 2006). Many ESL students in middle schools face challenges in English reading comprehension, thus it is important to investigate the characteristics of adolescent ESL unexpected poor comprehenders.

We aim to explore whether unexpected poor comprehenders differ, when compared to expected average comprehenders and unexpected good comprehenders, in vocabulary and higher level skills such as inference and strategy, listening comprehension, and summary writing. To our knowledge, this study is the first attempt to investigate the sources of reading comprehension
difficulties of Chinese ESL unexpected poor comprehenders. Therefore, the findings may provide us with a better basis for designing interventions to improve achievement in ESL students from Chinese background with reading comprehension difficulties.

Method

Participants

The participants for this study were selected from a sample of 246 students in four English-immersion classes in Grade 8 from a middle school in Xi’an, China. This research was cleared through Queen’s University General Research Ethics Board (see Appendix A). A Letter of Information (see Appendix B) was provided to parents requesting permission to allow their children to participate in this research. All students agreed to participate and their parents signed the consent form (see Appendix C). Two students who were absent for several tests were deleted from the study. About 30% of the curriculum is taught in English, including English language arts, social studies, and science. The rest of the curriculum is taught in Chinese.

Three groups of students: unexpected poor comprehenders, expected average comprehenders, and unexpected good comprehenders were identified within this sample using regression (Tong et al., 2011; White & Kirby, 2008). Students’ reading comprehension scores were regressed upon their age, nonverbal ability, and word reading scores, accounting for 16% of the variance. The standardized predicted values from this analysis were then plotted against actual reading comprehension scores (see Figure 1). To define the groups more precisely, and to avoid having students close to the boundaries between groups, confidence intervals around the regression line were used: students below the lower 70% confidence interval of the regression line were defined as unexpected poor comprehenders and those above the upper 70% confidence interval were defined as unexpected good comprehenders. Those students who scored within the
20% confidence intervals were identified as expected average comprehenders (see Figure 1).

Figure 1. The scatterplot for the distribution of unexpected poor comprehenders (UPCs), expected average comprehenders (EACs), and unexpected good comprehenders (UGCs), as well as those who were not selected for analyses (NSC), in the regression predicting reading comprehension from age, nonverbal intelligence, and word read efficiency.

The presence of multiple predictors in the regression equation made it possible that there were members of the groups with, for instance, very low word reading scores but very high nonverbal ability. Because the essential characteristic of unexpected poor comprehenders is poor reading comprehension in the presence of adequate word reading skills (e.g., Cain & Oakhill,
1999), we excluded students from the unexpected poor and unexpected good comprehender groups whose word reading scores were more than one SD below the overall sample mean. To ensure that the expected average comprehenders did not include students with extreme word reading scores, we eliminated those with scores more than one SD above or below the mean. This eliminated 2 participants from the unexpected poor comprehender group, 7 participants from the unexpected good comprehender group, and 16 participants from the expected average comprehender group. These procedures identified 33 unexpected poor comprehenders, 28 expected average comprehenders, and 30 unexpected good comprehenders, as shown in Figure 1.

The means and standard deviations of age and performance on nonverbal ability, word reading efficiency, and reading comprehension for the three groups are reported in Table 7. One-way ANOVA results, also reported in the Table 7, showed that unexpected poor comprehenders, expected average comprehenders, and unexpected good comprehenders were similar in age, nonverbal ability, or word reading efficiency. However, as expected, there was a significant difference in reading comprehension.
Table 7
Means, Standard Deviations, and Analysis of Variance Results for Unexpected poor comprehenders, Expected average comprehenders, and Unexpected good comprehenders on Age, Nonverbal Intelligence, Word Reading Efficiency, and Reading Comprehension

<table>
<thead>
<tr>
<th>Measures</th>
<th>Unexpected poor comprehenders (N = 33)</th>
<th>Expected average comprehenders (N = 28)</th>
<th>Unexpected good comprehenders (N = 30)</th>
<th>F(2, 88)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>162.67 4.65</td>
<td>162.39 4.86</td>
<td>162.20 4.57</td>
<td>0.08</td>
</tr>
<tr>
<td>Nonverbal Ability</td>
<td>51.00 2.87</td>
<td>52.43 3.35</td>
<td>51.20 3.57</td>
<td>1.65</td>
</tr>
<tr>
<td>Word Reading Efficiency</td>
<td>66.97 4.25</td>
<td>66.82 3.42</td>
<td>68.83 4.19</td>
<td>2.37</td>
</tr>
<tr>
<td>Reading</td>
<td>23.15 2.79</td>
<td>31.68 1.57</td>
<td>39.97 2.43</td>
<td>403.31**</td>
</tr>
</tbody>
</table>

Note. ** p < .001.

Measures

All reliability estimates given are for the complete sample (N = 246).

**Nonverbal Ability.** Raven’s Progressive Matrices (Raven, Raven, & Court, 1998) was used to measure students’ nonverbal ability. It comprises 60 items, divided into five sets of increasing complexity. All items have a similar format: A matrix of geometric designs with one cell of the matrix left blank is presented with six or eight alternatives for the matrix’s completion. The score was the number of correct answers. The Cronbach’s alpha reliability of this test was .79.
**Word Reading.** The Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999) was used to assess children’s ability to pronounce isolated words accurately and quickly. It is an individual test and student was asked to read as many of the 104 words as quickly and as accurately as possible within 45 seconds. The score was the number of words read correctly in 45 sec. The split-half reliability of this test was .93.

**Vocabulary Breadth.** The Gates-MacGinitie Vocabulary subtest (MacGinitie & MacGinitie, 1992) Level C, Grade 3 was given to the students to measure their vocabulary breadth. There are 45 items and students are required to finish them within 20 minutes. Each of 45 items shows a target word in a phrase or short sentence and students have to choose which of four options is closest in meaning. The score was the number of correct answers. The Cronbach’s alpha reliability of the test was .74.

**Word Definitions.** This individually administered task, based on that of Ouellette (2006), was used to measure vocabulary depth (see Appendix D). 12 words: bus, hand, apple, cow, tree, noodle, shirt, farm, television, island, drum, and envelope were chosen from the Peabody Picture Vocabulary Test (Dunn & Dunn, 1997). According to the manual, these words are used to test EL1 children 2 to 9 years old. Students were presented each word in both written and spoken form and were asked to provide definitions for the word. Students were encouraged to say as much as they could about each word. Responses were recorded by a digital recorder for later scoring.

Scoring was based on whether the student’s definitions addressed four features of each word: its category, function, description, and value. For example, the definition of the word “apple” would include its category (fruit), function (can be eaten or thrown), description (different colors, round shape, has skin and seeds, juicy, sour and sweet, grows on tree, etc.), and
value (good for health, give vitamin C, etc.). As long as the student could provide one aspect of each feature, 1 point was awarded, for a maximum score of 4 for each item and 48 for the test. The Cronbach’s alpha reliability of this test was .68.

**Multiple-meaning Vocabulary.** This test was based on Qian’s (2002) depth-of-vocabulary-knowledge measure and Tannenbaum, Torgesen, and Wagner’s (2006) Multiple Meanings test. It was used to measure students’ depth of vocabulary knowledge about the multiple meanings words can have in different contexts (see Appendix E). There were 15 target words (kind, can, light, miss, kiss, line, date, mind, play, ring, present, rose, run, second, well) and each word was presented in 4 different sentence contexts. Students were asked to choose the sentences in which the target word was used correctly. For example, for the target word “rose,” the sentences were: (a) The rose of the tree is very rough; (b) He bought a dozen roses for his girlfriend; (c) She is sitting in a rose watching TV; and (d) The sun rose over the mountains (the correct answers are b and d). The number of correct sentences varied from zero to four. The score was the number of correct sentences selected and incorrect sentences rejected, with a maximum score of 60. The Cronbach’s alpha reliability of this test was .68.

**Morphological Awareness.** The Base Identification task from Bowers and Kirby (2010) was used to measure students’ depth of vocabulary (see Appendix F). Students were asked to identify the bases of multi-morphemic words presented in written form by circling the smallest part that is related to the word (i.e., the base). The scoring was: 2 for circling the smallest base, 1 for removing at least one affix (without identifying the smallest base), and 0 for all other responses. For example, in the word “scarred”, 2 points were given for circling “scar” (the base), 1 point for circling “scarr” (because the suffix “ed” had been removed), and 0 points for “red” (because it is not related morphologically). The Cronbach’s alpha reliability of this test was .54.
Item analysis indicated that deleting three items would increase the reliability.60. However, this small increase did not produce differences in results, so it was decided to retain the three items.

**Listening Comprehension.** 19 items were chosen from the *Woodcock Listening Comprehension Test* (Woodcock, 1998), the first 9 items and the odd items from items 10 to 30; some items were changed to adapt to Chinese culture and the students’ background knowledge. This is a cloze-type assessment in which the student listened to tape-recorded passages in order of increasing difficulty and produced an oral response to unfinished sentences. The score was the number of correct answers. The split-half reliability of this test was .63.

**Inference and Strategy.** Two 10-item multiple-choice tests based on two passages chosen from *HIP Reading Assessment* (Jamison, 2007) and *Reading Inventory for the Classroom* (Flynt & Cooter, 2004) were constructed to measure students’ ability to draw inferences and apply various reading strategies in reading (see Appendix G). The two tests were modelled on those used by Cromley and Azevedo (2007). After reading each passage, students answered 5 multiple-choice items that assessed coherence and elaborative inferences, and 5 multiple-choice items that required the application of various reading strategies, such as finding main ideas, predicting, inferencing, and summarizing. The score was the number of correct answers. The reliabilities of these two tests were low, so to the tests were combined. The Cronbach’s alpha reliability of the combined test was .58. Item analysis indicated that two items could be deleted to increase the reliability to .60. However, this small improvement did not produce differences in results and the 2 items were retained.

**Summary Writing.** This task, based on Kirby and Pedwell’s (1990) and Stein and Kirby’s (1992) summary and recall tasks, was used to measure students’ deeper understanding of text (see Appendix H). Before this task began, students were reminded that a good summary
gives the main ideas of the text, that it is shorter, and that it is in one’s own words. The students were informed that the text and any notes they made while reading the text would be collected before they wrote their summaries. They were then given 15 minutes to read a 254 word passage entitled Cities that discussed the problems of cities in the Middle Ages. After a 5 minute interval, in which the original text and notes taken by the students were collected, they were given 10 minutes to write a text-absent summary. The students’ summaries were scored for the number of main ideas and the number of important details. Main ideas were defined as propositions which required integration of several ideas within the text, and important details were defined as those which supported the main ideas. The text was analyzed to have 5 main ideas and 19 important details (see Appendix I). One point was awarded for the inclusion of each unit, thus the maximum score was 24 points. Twenty of the summaries were scored by two raters; inter-rater reliability was .92. Any disagreements were resolved through discussion, and the remaining summaries were scored by the first author.

Reading Comprehension. The Gates-MacGinitie Reading Comprehension Test (MacGinitie & MacGinitie, 1992), Level C, Grade 3, was used to measure students’ reading comprehension. Participants read up to 12 short passages and with the passages still in view answered multiple-choice questions within the 35-min time limit. There were 48 items and the score was the number of correct answers. The Cronbach’s alpha reliability of this test was .77.

Procedure

All tests took place in the students’ school at times which were convenient to the school. There were four testers including the first author and they were all graduate students in Psychology or Education, fluent in English and Chinese. They were trained for three days in the administration of all tests. Individual testing lasted about 50 minutes (1 session) during which the
participant was tested by one tester in a quiet room. Group testing took approximately 200 minutes (spread over 4 sessions, spaced over several weeks), with a classroom teacher and two testers present in the classroom.

**Results**

**Descriptive Statistics**

The means and standard deviations of all measures are shown in Table 8. Measures whose skewness or kurtosis values fell outside of the acceptable range (i.e., the absolute value of Skewness/SE or Kurtosis/SE >3.09) were transformed according to the guidelines in Tabachnick and Fidell (2007). Nonverbal intelligence and multiple-meaning vocabulary were negatively skewed. They were first reflected and then square root transformations were applied. They were then reflected again so that high scores represented better performance. However, transformations of these two variables did not change their correlations with other variables substantially; therefore, all following analyses were conducted on the raw scores.
Table 8

Means, Standard Deviations, and Analysis of Variance Results for Unexpected poor comprehenders, Expected average comprehenders, and Unexpected good comprehenders on Vocabulary Measures and Higher Level Processes

<table>
<thead>
<tr>
<th>Measures</th>
<th>UPC (N = 33)</th>
<th>EAC (N = 28)</th>
<th>UGC (N = 30)</th>
<th>Pairwise Comparisona</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Vocabulary Breadth</td>
<td>16.27</td>
<td>4.18</td>
<td>19.93</td>
<td>6.73</td>
</tr>
<tr>
<td>Word Definition</td>
<td>25.21</td>
<td>3.70</td>
<td>24.32</td>
<td>5.29</td>
</tr>
<tr>
<td>Multiple-meaning</td>
<td>49.52</td>
<td>4.10</td>
<td>52.79</td>
<td>4.01</td>
</tr>
<tr>
<td>Morphological Awareness</td>
<td>50.15</td>
<td>4.18</td>
<td>49.18</td>
<td>5.39</td>
</tr>
<tr>
<td>Inference and Strategy</td>
<td>11.88</td>
<td>2.72</td>
<td>13.00</td>
<td>2.23</td>
</tr>
<tr>
<td>Listening Comprehension</td>
<td>8.55</td>
<td>2.08</td>
<td>8.96</td>
<td>1.75</td>
</tr>
<tr>
<td>Summary Writing</td>
<td>9.73</td>
<td>4.91</td>
<td>9.75</td>
<td>5.11</td>
</tr>
</tbody>
</table>

Note. UPC = unexpected poor comprehenders; EAC = expected average comprehenders; UGC = unexpected good comprehenders

a Equal sign indicates nonsignificant difference, and less-than symbol indicates \( p < .05 \).
Group Differences in Vocabulary and Higher Level Processes

Two multivariate analyses of variance (MANOVAs) were conducted to test the differences among the three groups in vocabulary and higher level processes. In the first MANOVA, the dependent variables were vocabulary breadth (Gates-MacGinitie Vocabulary) and three measures of vocabulary depth (word definition, multiple meanings, and morphological awareness). There was a statistically significant difference between the three groups on these vocabulary measures, Wilks’ λ = 0.67, \( F(8, 170) = 4.76, p < .001 \). Univariate one-way analyses of variance indicated that group had a significant effect on vocabulary breadth, \( F(2, 88) = 11.04, p < .001 \), partial \( \eta^2 = .20 \), multiple meanings, \( F(2, 88) = 8.55, p < .001 \), partial \( \eta^2 = .16 \), and morphological awareness, \( F(2, 88) = 4.44, p < .05 \), partial \( \eta^2 = .09 \). Bonferroni post-hoc tests showed that vocabulary breadth and multiple meanings distinguished between unexpected poor and expected average comprehenders, whereas morphological awareness distinguished between expected average and unexpected good comprehenders (see Table 8).

In the second MANOVA, the dependent variables were higher level skills (inference and strategy, listening comprehension, and summary writing). There was a statistically significant difference among three groups on these higher level skills, Wilks’ λ = 0.58, \( F(6, 172) = 8.99, p < .001 \). Group has a significant effect on inference and strategy, \( F(2, 88) = 16.42, p < .001 \), partial \( \eta^2 = .27 \), listening comprehension, \( F(2, 88) = 9.81, p < .001 \), partial \( \eta^2 = .18 \), and summary writing, \( F(2, 88) = 4.62, p < .05 \), partial \( \eta^2 = .10 \). Bonferroni post-hoc tests showed that inference and strategy, listening comprehension, and summary writing were significantly different between expected average comprehenders and unexpected good comprehenders but not between unexpected poor comprehenders and expected average comprehenders (see Table 8).
Discussion

This study investigates the possible sources of reading comprehension difficulties of Chinese ESL students. Although our method does not establish the causes of reading comprehension difficulties, which can only be done in longitudinal studies, intervention studies, and reading-age-match comparisons, our results indicate a plausible explanation for the sources of reading comprehension difficulties in ESL unexpected poor comprehenders. Within the larger sample of Chinese ESL Grade 8 students, we identified three groups of comprehenders: unexpected poor comprehenders, expected average comprehenders, and unexpected good comprehenders, which had been identified in previous EL1 research (e.g., Tong et al., 2011; White & Kirby, 2008). There were no significant differences among these groups in age, nonverbal intelligence, and word reading ability, but they differed substantially in reading comprehension. About 13% of the children were identified as unexpected poor comprehenders in this study. This finding is consistent with previous studies in which 10-15% of EL1 children were found to be unexpected poor comprehenders in elementary schools (e.g., Catts et al., 2005; Leach et al., 2003; Nation et al., 2004). The current finding is close to the 17-18% of ESL children who were found to be unexpected poor comprehenders in elementary schools in Canada and the United States (e.g., Lesaux & Kieffer, 2010; Lesaux et al., 2006).

Vocabulary Skills Distinguish Unexpected Poor Comprehenders from Expected Average Comprehenders

When we explored the possible sources of English reading comprehension difficulties among these Chinese ESL students, the results demonstrated that unexpected poor comprehenders scored significantly lower than expected average comprehenders, and also much lower than unexpected good comprehenders on vocabulary breadth and multiple meanings in vocabulary
depth. This pattern suggests that unexpected poor comprehenders’ inferior reading comprehension may arise from vocabulary difficulties. However, there was no significant difference between expected average comprehenders and unexpected good comprehenders in vocabulary breadth and multiple meanings. The lower vocabulary performance of unexpected poor comprehenders is consistent with the findings of others who have examined language abilities in EL1 unexpected poor comprehenders. These findings indicate that English native speaking poor comprehenders have weaker vocabulary and semantic processing skills than children with normal reading comprehension (Catts et al., 2006; Hock et al., 2009; Nation et al., 2004; Nation & Snowling, 1998). The deficits in vocabulary could lead to problems inferring the meaning of new words from text and learning new vocabulary from direct instruction (Cain et al., 2003, 2004; Nation et al., 2007). Our results are also consistent with Lesaux and Kieffer’s (2010) conclusion that vocabulary was the main source of reading comprehension difficulties in Grade 6 ESL students. In addition, our findings support the lexical quality hypothesis that reading comprehension depends on the quality of lexical representation of words (Perfetti & Hart, 2001). Not only the number of words (breadth), but also the quality of the knowledge of words (depth) is crucial to reading comprehension. Skilled reading comprehension relies upon the richness of word knowledge, which was assessed by the multiple-meaning vocabulary depth measure. Knowing multiple meaning of the words is more than simple vocabulary knowledge; it gets at the quality or depth of lexical representation. Developing vocabulary depth may be more challenging to them, because of their lack of exposure to the L2. However, these results suggest that ESL unexpected poor comprehenders will need to improve both vocabulary breadth and depth to improve their reading comprehension.
Surprisingly, vocabulary depth measured by word definition did not distinguish between unexpected poor comprehenders and expected average comprehenders. It is possible that the words in the word definition task are not difficult enough to distinguish between the two groups. Morphological awareness also did not distinguish between unexpected poor comprehenders and expected average comprehenders but it did distinguish between expected average comprehenders and unexpected good comprehenders. Tong et al. (2011) also found that morphological awareness distinguished between expected average comprehenders and unexpected good comprehenders, in Grade 3 EL1 students, using an oral word analogy measure of morphological awareness. However, in Grade 5, Tong et al. (2011) found morphological awareness distinguished between unexpected poor comprehenders and expected average comprehenders. This indicates that the differences between unexpected poor comprehenders and expected average comprehenders in morphological awareness emerged during the course of development, as the expected average comprehenders became more proficient. Thus morphological awareness may be a more advanced or later developing form of vocabulary knowledge, perhaps because it is not explicitly taught. The expected average comprehenders of the present study, although older than those of Tong et al., may not have had a rich enough experience of English to improve beyond the level of the unexpected poor comprehenders; in contrast, only the unexpected good comprehenders have had this rich experience of English, which may contribute to their being unexpected good comprehenders.

**Higher-level Skills Distinguish Expected Average Comprehenders from Unexpected Good Comprehenders**

However, vocabulary is necessary but not sufficient for good comprehension (Cain, Oakhill, & Lemmon, 2004). Our results show that unexpected good comprehenders performed
higher than expected average and unexpected poor comprehenders on inference and strategy, listening comprehension, and summary writing. Previous studies of EL1 unexpected poor comprehenders have found them to have difficulties in higher level comprehension skills, such as inferencing, working memory, and strategy use, compared to undifferentiated groups of good comprehenders (e.g., Cain, 1999; Cain & Oakhill, 1999, 2006; Cain et al., 2000, Cain et al., 2001; Catts et al., 2006). The present results suggest that some of those previous results may be due to the inclusion of unexpected good comprehenders in the comparison group. Unexpected poor comprehenders do have weaknesses in higher level skills, but in this they may not be different from expected average comprehenders. Higher level skills may instead be the defining characteristic of unexpected good comprehenders, those whose reading comprehension exceeds what would be expected from the word reading and general mental abilities.

This suggests the hypothesis that higher level skills may only contribute effectively to children’s reading comprehension once their English language proficiency has reached a certain level. Low vocabulary skills may constrain the use of higher level processes and morphological awareness in the expected average comprehenders because vocabulary and other basic language skills act in a bottom-up manner in reading comprehension, providing word- and sentence-level input to higher comprehension processes. If ESL students have not acquired adequate vocabulary, more cognitive resources are needed to engage in lower level processing, thus leaving few cognitive resources for higher level skills processing (Perfetti, 1985). Although the present findings are from ESL students, this hypothesis is worth investigating with EL1 students.

Conclusions, Limitations, and Implications

Taken together, these results indicate that Grade 8 ESL unexpected poor comprehenders had deficits in vocabulary, compared to expected average comprehenders. Previous studies
mainly compared unexpected poor comprehenders to good comprehenders but not to average comprehenders in reading-related skills. Our distinction between unexpected poor and expected average comprehenders allowed us to further show that vocabulary skills, including both depth and breadth, are what prevent the unexpected poor comprehenders from performing at an average level. Higher level skills of making inferences, strategy, and discourse comprehension prevent both the unexpected poor and expected average comprehenders from proceeding to more advanced levels of comprehension. Although the evidence is not strong enough to prove causation, it does show plausible causes of reading comprehension difficulties. Such findings can then be followed up with longitudinal or intervention studies to provide a more rigorous test of the hypothesis.

Several limitations should be noted. First, the constructs considered here are complex and multi-faceted, and we were only able to use a limited number of measures. The present results need to be replicated with other and more varied measures. Furthermore, several of our measures had low reliabilities (inference and strategy, and morphological awareness), suggesting that the results be treated with some caution. Second, we only used word reading efficiency to measure word reading ability; other measures such as word reading accuracy and word decoding are needed to provide a full range of word reading skills. Third, other constructs such as syntactic awareness and working memory should be considered for inclusion in the future ESL reading comprehension difficulties research (Lesaux et al., 2006). Finally, the participants in the present study came from a particular background in which native English-speaking models were not widely available, so research needs to be done to investigate whether these results generalize to other groups of ESL students.
This study provides some guidance for instruction for ESL students who have reading comprehension difficulties but normal word recognition ability. The evidence that vocabulary is the key source to differentiate ESL poor comprehenders from average and good comprehenders highlights the need for effective classroom vocabulary instruction. Less attention is paid to vocabulary instruction in middle school classrooms than in elementary school classrooms, and most ESL teachers may only provide a brief definition of new word without rich and deep elaboration (Ivey, 1999; Lesaux & Kieffer, 2010). Therefore, systematic and explicit vocabulary instruction is required in the classroom to help students understand and acquire new knowledge of the text. Our results that unexpected poor comprehension have deficits in both vocabulary breadth and depth suggest that teachers should target both the size (breadth) and the depth of vocabulary by introducing multiple meanings of new words and strategies for learning words independently. Morphological instruction may be an efficient and productive means of teaching vocabulary (Bowers & Kirby, 2010).

Once ESL students have acquired well-developed vocabulary and language proficiency, higher level processes appear to be the main sources of variability in reading comprehension. There is a need to focus on the training of higher level processes to improve intermediate ESL students’ reading comprehension. It is important to note that these higher level processes include both strategies and the processes of making inferences, identifying and constructing main ideas, and situation model construction (Kirby et al., 2012). Further research is required to determine how to optimally balance lower and higher level skill instruction, and when to introduce the latter. This research will allow us to help ESL students overcome the challenges they face in becoming skilled reading comprehenders.
Chapter 4

General Discussion

In this chapter, I discuss the findings of my two studies by highlighting the contributions this research makes to our knowledge about ESL reading comprehension. First, I provide a broad picture of English reading comprehension in Chinese English-immersion students by revisiting the purpose of my research, briefly summarizing my results, and making connections between my two studies and the literature. Second, I raise several critical issues around ESL reading comprehension based on my findings. Third, I reflect on the implications of my findings for the practice of teaching and future research. Finally, I offer concluding remarks on my research.

A Picture of English Reading Comprehension in Chinese Students

More and more people have started to learn English as an L2 because English is not only a global language but also the language of science, technology, and advanced research (Grabe & Stoller, 2011). In China, enthusiasm for English has become more and more apparent. This can be seen from the policy change of lowering the starting age for English learning. In 2001, the Ministry of Education mandated a formal policy to begin English instruction in China in Grade 3 instead of Grade 7 (Hu, 2007). At the same time, new methods of teaching English have appeared that challenge the traditional methods of teaching English in China, which emphasized grammar and rote memory. For example, English-immersion programs were set up in several cities in China to bring in a new approach to teaching English, which uses English to teach non-language subjects. It is important to know how these English-immersion students read in English and why some of them have difficulties in English reading. The goal of my research was to understand the
cognitive processes underlying English reading comprehension in these Chinese ESL students who are in English-immersion programs.

Two studies were conducted to provide different perspectives on the cognitive development of English reading comprehension in Chinese English-immersion students in the middle school. The first study examined which cognitive factors significantly predict English reading comprehension in these students. I found that decoding and linguistic comprehension, which are two powerful predictors of reading comprehension in EL1 children, do not predict reading comprehension well in Chinese ESL children. Vocabulary as well as inference and strategy explained additional variance in reading comprehension beyond decoding and listening comprehension. The breadth and depth of vocabulary knowledge were measured, each contributing differently to different levels of reading comprehension measures. Breadth of vocabulary contributed more to surface understanding of the text and depth of vocabulary contributed more to deep understanding of the text. The second study explored characteristics of reading comprehension difficulties among Chinese English-immersion students. Using a regression method, I identified a group of students whose reading comprehension was much less than would be expected from their word reading and general ability. These ESL unexpected poor comprehenders, when compared to average comprehenders, performed poorly on vocabulary measures; thus vocabulary may be the main source of reading comprehension difficulties. However, higher-level processes may be the reason why unexpected good comprehenders have an advantage in reading comprehension over average comprehenders.

The results from the two studies provide a picture of how students read in English as their L2 in China. Generally speaking, although the cognitive development of Chinese ESL students may be similar to that of EL1 students, their English performance is much lower than that of EL1
students. The main vocabulary breadth test and the English reading comprehension test, which are standardized tests designed for Grade 3 EL1 students, were given to Grade 8 Chinese ESL students and were found appropriate in my research. According to the Gates-MacGinitie Reading Test manual, the mean scores of vocabulary and reading comprehension of Chinese students are equal to those of North American students at Grade 2.8 and 3.7 levels. This finding suggests that the English proficiency level of Grade 8 Chinese students, at least in some respects, is equal to that of Grade 3 EL1 students, presumably due to less exposure to English and different learning contexts. The finding is consistent with the results of studies cited in the literature review that even ESL students in North American schools have low vocabulary and reading comprehension compared to their EL1 peers, although they develop decoding skills equivalent to their EL1 peers (August, Carlo, Dressler, & Snow, 2005; Lesaux, Geva, Koda, Siegel, & Shanahan, 2008). The present finding suggests that vocabulary is the key factor to understand why some Chinese ESL students do poorly in reading comprehension. ESL students need to have enough vocabulary knowledge so that their higher-level skills can be used efficiently to help their reading.

**Critical Issues around ESL Reading Comprehension**

**The Crucial Role of Vocabulary in ESL Reading Comprehension**

The findings from my research show that vocabulary plays a crucial role in ESL students’ reading comprehension. The Threshold Hypothesis provides a theoretical foundation for the important role of L2 vocabulary in reading (e.g., Clarke, 1980; Cummins, 1979). This hypothesis stated that L2 students must have a sufficient amount of L2 vocabulary to use skills from their L1 reading comprehension abilities. When they attain a certain threshold level of linguistic competence in L2, their L1 reading strategies and skills can be used to help comprehend the L2 text. This hypothesis has been supported by a number of L2 reading studies (see Bernhardt, 2011;
Grabe, 2009; Koda, 1989; Nation, 1990; Yamashita, 2002). But how can we determine a threshold score of vocabulary for L2 learners to be able to comprehend L2 text? Laufer and Ravenhorst-Kalovski (2010) claimed that a threshold of 95% lexical coverage of a text is needed for minimum comprehension. Lexical coverage refers to the percentage of the words in a text that a reader understands, so it is probably closest in meaning to breadth of vocabulary. As emphasized in Study One, vocabulary depth is also important: a student who knew the basic meanings of most words in a text but did not know their possible multiple meanings or alternative usages would be unlikely to comprehend the text well.

ESL students who experience slow vocabulary development are less able to comprehend text than those with average vocabulary development. Previous research in reading comprehension difficulties of ESL children has indicated that these ESL unexpected poor comprehenders’ weaknesses in reading comprehension are caused by slowness or inability to access word meanings (Lesaux & Kieffer, 2010). Vocabulary can affect reading processing at either word levels including phonological, orthographic, and morphosyntactic processes, or sentence and text levels such as grammatical processing, construction of text meaning, and text models (Alexander, & Fox, 2004; Carlisle & Nomanbhoy, 1993; Chall, 1987; Geva & Farnia, 2011; Perfetti, 2007; Snow, 2010). Knowing a word implies knowing many things about the word, such as its primary literal meaning, alternative or multiple meanings, collocations, morphological structure, use in syntactic constructions, and semantic associates. These various aspects are related to depth of vocabulary, which is important in addition to learning many words (breadth of vocabulary). In Study One, vocabulary depth was found to explain additional variance in deeper level reading comprehension after considering vocabulary breadth. Knowing more about the meanings and use of words should improve the quality of knowledge about the words,
and thus, according to the lexical quality hypothesis (Perfetti, 1985), enhance the depth of understanding of the text.

**Assessment of Vocabulary Depth**

Most ESL studies of the relationship between vocabulary and reading comprehension have focused on vocabulary breadth rather than depth. Vocabulary depth is a relatively new construct which is considered to measure deep understanding of word meanings (Anderson & Freebody, 1981). Although there is growing attention to the effect of depth of vocabulary knowledge on reading comprehension in L2 students (Proctor et al., 2009; Qian, 1999, 2002; Vermeer, 2001), there has not been a consensus on how depth of vocabulary should be measured. Unlike the measures of vocabulary breadth, there is no standardized measure of vocabulary depth. Only a few tasks have been developed to assess depth of vocabulary knowledge (Kieffer & Lesaux, 2008; Ouellette, 2006; Tannenbaum et al., 2006; Qian, 1999, 2002). There are several possible reasons for the lack of measures, including the newness of the construct, the lack of a clear theory-driven construct definition, and the complexity of assessing the full range of meanings and uses of words.

I argue that none of the individual vocabulary depth measures captures the entire scope of the construct. Most studies of the relationship between vocabulary depth and reading comprehension have included only one vocabulary depth measure, and so the full construct of vocabulary depth has not been assessed. In my research, based on the previous measures used to assess depth of vocabulary, I designed three measures: word definition, multiple-meanings, and morphological awareness. The relatively small correlations among the three measures (in the .21 to .24 range; see Table 2 in Study One) suggest that they assess different components of vocabulary depth and that vocabulary depth requires multiple measures to be fully assessed.
However, when I employed exploratory factor analysis to verify the construct measured by the three vocabulary depth measures, the multiple-meanings measure had split loadings on both vocabulary depth and breadth factors. This measure, which was intended to only measure vocabulary depth, assessed both depth and breadth. Clearly more work is required to measure vocabulary depth as a construct more validly and there is a need for a full set of measures covering all of the components of vocabulary depth. It is also likely that vocabulary breadth and depth are so inter-related that it may not be possible to design pure measures of either.

Assessment of Reading Comprehension

Among various reading comprehension measures, multiple-choice tests are frequently used because they are easy to administer and score. However, it has been found that multiple-choice reading comprehension measures allow students to answer some test questions without reading the passage (see Keenan & Betjemann, 2006). For example, having given children and adults the passage-less questions from the Gray Oral Reading Test (Wiederholt & Bryant, 1992), a standardized multiple-choice reading test, Keenan and Betjemann (2006) found that the majority of the questions could be answered with above-chance accuracy without participants having read the passages. In addition, multiple-choice reading comprehension measures have been criticized because students can still obtain the correct answer by guessing or eliminating the distracters in the choices even if they do not quite understand the text (Katz & Lautenschlager, 2001; Roediger & Marsh, 2005).

Reading comprehension measures should address both lower and higher-levels of processing. The multiple-choice reading comprehension measure is a test that measures reading comprehension without usually activating much higher-level processing (Haladyna, Downing, & Rodriguez, 2002). Students are found not to engage in deep learning when they are preparing for
multiple choice exams (Scouller, 1998). In theories of reading comprehension, higher-level processing is crucial to ensuring a high quality of learning. A valid reading comprehension measure is expected to reflect the extent to which students are able to construct meaning from the text (Campbell, 2005). In my research, in addition to the multiple-choice reading comprehension measure, text-absent summary writing was used to measure whether or not students understand the passage which they had read. This test measures students’ deep understanding of the passage because they are not able to write a competent summary if they do not understand the main ideas of text. Text-absent summary writing is perceived as assessing higher-level cognitive processing (Kirby & Woodhouse, 1994; Nuttall, 1996). However, the summary writing task has not been often used to measure reading comprehension due to more demanding scoring. There is also a possibility that students may understand the text but have difficulties in writing the summary, which is typical for ESL students. More generally, text-absent summary writing may reflect writing more than reading problems. It is open to rote memory strategies, reliance on prior knowledge rather than what was read, and may require more motivation than other measures. As reading comprehension is a complex construct, no single measure can assess all processes and different tests have different purposes. The assessment of reading comprehension should be multidimensional (Francis, Fletcher, Catts, & Tomblin, 2005), and based upon multiple measures.

**Higher-level Processes in ESL Reading Comprehension**

The two studies performed in my research together emphasize the importance of higher-level processes through inference and strategy, summary writing, and listening comprehension in ESL reading comprehension. Higher-level processes enable readers to make the necessary integrative and inferential links to construct a meaningful and productive representation of the text, and therefore are necessary to ensure advanced reading comprehension (Cain & Oakhill,
2006). These higher-level skills involve deeper processing which is required to integrate the meanings of individual words and sentences within the text as a whole (Kintsch, 2004). This is particularly shown in my first study in which inference and strategies still make additional contributions after the effects of vocabulary and other cognitive variables were controlled. It is also shown by the advantage unexpected good comprehenders had in higher-level processes in the second study.

EL1 reading research indicates that some children have weaknesses in making inferences requiring the integration of information and they use inappropriate and inefficient strategies to understand what they have read (Cain & Oakhill, 1999, 2003, 2006; Cain, Oakhill, Barnes, & Bryant, 2001). When attempting to remember a text, these EL1 unexpected poor comprehenders are also found to use a surface approach that focuses on the reading of individual words rather than a deep approach that focuses on the main points of the text because they have not developed strong word recognition (Cain, 1999). The ESL unexpected poor comprehenders in my study also showed a similar trend: they were poor at finding important details and main ideas to obtain a global understanding of the text. However, further research is necessary to establish how impairments in higher-level skills are related to reading comprehension difficulties. The causal link between weaknesses in higher-level skills and inadequate reading comprehension is not clear. It may be that unexpected poor comprehenders’ poor ability in higher-level skills leads to reading comprehension difficulties. On the other hand, it may be that their low reading comprehension is due to poor word reading skills, which in turn leads to lesser experience of reading, which in turn weakens the acquisition of higher-level processing skills. Resolving these causal issues will require longitudinal and experimental studies.
Implications for Teaching and Future Research

This research has implications for both classroom teaching and future reading research. The findings suggest the need for more attention to reading instruction for older ESL students in middle schools as well as the need to develop reading models that take into account the lower-level and higher-level skills of ESL students.

Vocabulary Teaching

The findings from the two studies demonstrate the importance of vocabulary in reading comprehension in Chinese English-immersion students but raise the question of how to effectively teach them vocabulary. In EL1 reading, Biemiller and colleagues (Biemiller & Slonim, 2001; Biemiller & Boote, 2006) suggested a wide but shallow approach of vocabulary instruction which favours direct teaching of many words without targeting deeper knowledge of words. In contrast, Beck and colleagues (Beck & McKeown, 1983; Beck et al., 2002) suggested a deep but narrow approach of vocabulary instruction which goes beyond establishing a definition for a word and develops complex, in-depth knowledge about the words being taught. Bowers and Kirby (2010) proposed a deep and wide approach by advocating the teaching of morphology, arguing that once bases are known, vocabulary knowledge can grow productively around these kernels by adding affixes.

In English-immersion classes in China, where English is not encountered often outside of school, vocabulary instruction should be paid even more attention. Students’ access to the environment of using English is more limited than that of EL1 and ESL students in English-speaking countries. Deliberate teaching of vocabulary may be the most effective way to enlarge their vocabulary and raise their awareness of particular words (Genesee & Riches, 2006; Nation, 2001; Stoller & Grabe, 1995). The basic words which may not need to be taught to EL1 students
may have to be taught in class to ESL students in China given that their low chance of encountering even those words outside of English class.

However, according to my observation of English-immersion classes in one middle school in China, I found most teachers place an emphasis on spoken English and group discussion without focusing much on explicit vocabulary instruction. This has also been found in ESL classes in North America (Scott, Jamieson-Noel, & Asslin, 2003; Watts, 1995). English classes taught in this way incorporate little systematic and explicit vocabulary instruction in the curriculum. In discussion with English immersion teachers, they indicated that it is the students’ responsibility to learn vocabulary on their own before class. There may be a misconception among teachers that teaching vocabulary is an out-dated method for English teaching and that they should focus instead more on subject content and conversational English. However, if students do not accumulate adequate vocabulary knowledge, how can they build up foundations to learn other skills? My findings suggest that Chinese ESL students would need direct vocabulary instruction as part of the classroom curriculum to prevent at least some of the reading comprehension difficulties. The vocabulary instruction should address both breadth and depth. Students clearly need to begin with relatively superficial knowledge about new words (breadth). It is important for them to learn the meaning of many basic words that EL1 students already know, and these words may require deliberate instruction (Beck et al., 2002; Nation, 2001; Stoller & Grabe, 1995). The sooner depth can be added the better. Depth can be introduced through morphology instruction. Once students acquire the bases, their vocabulary can be enlarged by adding different affixes. Depth need not be contrary to breadth. When students know more words, they are able to develop a matrix showing relationships among the words and make semantic links to other words. As a result, they gain rich understanding of the words. Furthermore, depth
can be developed by teaching multiple meanings and various usages of the word (Graves, 2006; Stahl & Nagy, 2006). Studies that provided rich and deep vocabulary instruction have shown gains in both accuracy of word knowledge and comprehension of text containing the taught words in L1 and L2, and the gains were even greater for L2 learners (Blachowicz & Fisher, 2000; Carlo et al., 2004; Milton, 2009).

**Teaching of Higher-level Skills**

When ESL students have achieved higher language proficiency in their L2, higher-level skills appear to be important in their reading comprehension. Research in EL1 has shown that students who have reading difficulties benefit from training programs that focus on making inferences and reading strategies (Cain et al., 2004). For example, teaching how to recognize and use the implicit and explicit cues to generate inferences, and teaching reading strategies such as locating information, finding main ideas, and activating background knowledge can help in comprehending the text. The use of inferences and strategies may compensate for weaker low-level skills in reading comprehension because readers can employ certain reading strategies to infer the meanings of unknown words and connect information to establish an adequate sense of the text. This process may be particularly useful for ESL readers who have low language proficiency in their L2.

However, teaching how to make inferences and employ strategies is not enough to lead to deep understanding of the text. Higher-level processing is more than inferencing and using strategies; it is deeper processing which is embodied in the construction of situation models. A situation model should be an elaborated interpretation of the text that goes beyond the text, and it should involve a great deal of higher-level processing. Students need to construct a situation model by integrating text information with prior knowledge to create a coherent representation of
the text (Kintsch, 1998). Teachers should not expect great instructional benefits from simply teaching strategies. They need to provide instruction on how to establish an integrated model of text. Not only do they need to teach students how to integrate information, but they also need to teach students to construct arguments, evaluate, monitor, and develop critical thinking to lead to deeper learning (Britt & Rouet, 2012). The deeper goal of reading is to achieve high quality learning. Higher-level processing or deeper processing is essential for high quality learning (Kirby, Cain, & White, 2012). Higher-level processing is difficult for many students because it requires more basic aspects of text processing, prior knowledge, elaboration on important information, construction of meanings, and integration of information. But educators must help students overcome the difficulty and accept the challenge to increase students’ higher-level skills for deeper understanding and higher quality learning.

**ESL Reading Comprehension Model**

The major findings of this research highlight the need for a comprehensive theoretical framework for ESL reading comprehension in early adolescents, and emphasize the need to develop theories about the reading comprehension difficulties that some ESL students encounter. The results from the present research clearly demonstrate that Chinese English-immersion students’ reading comprehension depends upon important lower-level skills such as decoding, speed of processing, and vocabulary. The findings add to the earlier ESL literature which focused mainly on higher-level skills. As information is processed, the lower-level processing is integrated into higher-level processing (Kirby, 2007; Perfetti, Landi, & Oakhill, 2005). Deeper comprehension relies critically upon higher-level skills. Therefore, a comprehensive reading model should include a full range of lower-level and higher-level processes and their inter-relations.
This research examined the reading related skills and reading comprehension of a group of Grade 8 Chinese ESL students. To fully understand the English reading comprehension development of ESL students, longitudinal and experimental studies are required. It is important to note that my results are limited to Chinese ESL students learning English in China. Further research is needed to explore whether the findings can be extended to other ESL students from different L1 backgrounds. It is also necessary to investigate how the reading trajectories of these Chinese ESL learners compare not only to ESL learners from other backgrounds but also to EL1 learners. The comparison of different groups learning to read English would provide a better understanding of the complexities of English reading comprehension and would offer guidance for classroom instruction.

**Concluding Remarks**

Research in EL1 reading has demonstrated that lower-level skills (e.g., phonological awareness, decoding, fluency, and vocabulary knowledge) and higher-level skills (e.g., inferencing skills, comprehension monitoring, integration of text information) are essential for success in English reading comprehension (National Reading Panel, 2000; Oakhill, Cain, & Bryant, 2003; Snow, 2002; van den Broek et al., 2005). The findings from my research indicate that decoding and vocabulary, as well as higher-level skills such as inference and strategy, listening comprehension, and summary writing, are also important to reading comprehension of Chinese English-immersion students. However, lower-level skills may be more crucial than higher-level skills to students who have reading comprehension difficulties because the higher-level skills will be limited in their effects until those lower-level skills are automated.
References


Children Receiving English Instruction. *Journal of Educational Psychology, 102*, 712-728.


111
English-as-a-Second-Language children: The role of oral proficiency and underlying


Specialist, 6*, 126-135.

second-language reading comprehension? A study of Spanish-speaking children
acquiring English as a second language from first to second grade. *Journal of
Educational Psychology, 101*, 330-344.

and Special Education, 7*, 6-10.

Cambridge University Press.

Education Limited.


Bilingual education in China: Practices, policies, and concepts (pp. 94-126). Clevedon:
Multilingual Matters.


Language Policy, 6, 359-376.

Huang, X., Tube, B., & Yu, C. (2011). Meeting the dual goals of content knowledge and English
language learning: A study of the CCUEI curriculum materials. Frontiers of Education in
China, 6, 37-67.

Bristol, UK: Multilingual Matters.


Johnston, T. C. & Kirby, J. R. (2006). The contribution of naming speed to the simple
view of reading. Reading and Writing, 19, 339–361.


item performance on the SAT reading task. Educational Assessment, 7, 165-176.


118


Savage, R. S. (2006) Reading comprehension is not always the product of decoding and listening comprehension: evidence from teenagers who are very poor readers. *Scientific Studies of Reading, 10*, 43–164.


Appendix A

Research Ethics Board Approval

October 21, 2010

Ms. Maio Li
Ph.D. Candidate
Faculty of Education
Duncan McArthur Hall
Queen’s University

Dear Ms. Li:

GREB Ref #: GEDUC-525-10
Title: “Understanding Reading Comprehension in Chinese English-immersion Students”

The General Research Ethics Board (GREB), by means of a full board review, has cleared your proposal entitled “Understanding Reading Comprehension in Chinese English-immersion Students” for ethical compliance with the Tri-Council Guidelines (TCPS) and Queen’s ethics policies. In accordance with the Tri-Council Guidelines (article D.1.6) and Senate Terms of Reference (article G), your project has been cleared for one year. At the end of each year, the GREB will ask if your project has been completed and if not, what changes have occurred or will occur in the next year.

You are reminded of your obligation to advise the GREB, with a copy to your unit REB, if applicable, of any adverse event(s) that occur during this one-year period (details available on webpage http://www.queensu.ca/ors/researchethics/GeneralREB/forms.html - Adverse Event Report Form). An adverse event includes, but is not limited to, a complaint, a change or an unexpected event that alters the level of risk for the researcher or participants or situation that requires a substantial change in approach to a participant(s). You are also advised that all adverse events must be reported to the GREB within 48 hours.

You are also reminded that all changes that might affect human participants must be cleared by the GREB. For example, you must report changes in study procedures or implementations of new aspects into the study procedures on the Ethics Change Form that can be found at http://www.queensu.ca/ors/researchethics/GeneralREB/forms.html – Research Ethics Change Form. These changes must be sent to the Ethics Coordinator, Gail Irving, at the Office of Research Services or irvingg@queensu.ca prior to implementation. Mrs. Irving will forward your request for protocol changes to the appropriate GREB reviewers and/or the GREB Chair.

On behalf of the General Research Ethics Board, I wish you continued success in your research.

Yours sincerely,

Joan Stevenson, Ph.D
Professor and Chair
General Research Ethics Board

c.c.: Dr. John Kirby, Faculty Supervisor
Dr. Lesly Wade-Woolley, Chair Unit REB
E-REB: c/o Graduate Studies & Bureau of Research, Attn.: Celina Freitas

JS/gi
Appendix B

Letter of Information

Dear Parent or Guardian,

I am inviting you and your child to participate in a research entitled *Understanding Reading Comprehension in Chinese English-immersion Students*. My name is Miao Li and I am a doctoral student from the Faculty of Education at Queen’s University in Canada, and I am conducting this research under the supervision of Dr. John Kirby. This research is a joint project between the Faculty of Education at Queen’s University in Canada and the China-Canada-USA English Immersion (CCUEI) collaboration program. This research has been granted clearance according to the recommended principles of Canadian ethics guidelines, and Queen's policies. The director of English-immersion programs of your child’s school has also given approval for this research.

The purpose of this research is to explore the cognitive processes underlying English reading comprehension in Chinese students in English-immersion programs. Reading comprehension is the process of constructing meaning from text. Research with students whose first language is English has shown that decoding, vocabulary, listening comprehension, working memory, inference, and strategy skills are significant predictors of reading comprehension. However, it is not clear whether these factors also influence reading comprehension of students who are learning English as a second language. It is important to study these factors in the development of reading so that we can design the best type of instruction to improve students’ English reading comprehension.

For this research, individual and group tests will be given to your child to measure his or her knowledge of vocabulary, listening comprehension, working memory, inference, and strategy skills as well as his or her ability to read words and comprehend texts in English and Chinese. There will be four testers including me and three graduate students from the Faculty of Education at Shanxi Normal University. Individual testing will take approximately 50 minutes (1 or 2 sessions) and your child will be individually tested by one tester in a quiet room with the door open. Group testing will take approximately 200 minutes (spread over 4 sessions), with a classroom teacher and at least two testers present in the classroom. Non-participants will go to the school library to read. For some of the tests, your child’s answers will be recorded as a digital audio file using a digital recorder. In this way, the responses can be scored at a later time. All tests will take place at school during times which are convenient to the school and will not affect their regular lessons. Each session will only begin if your child agrees to take part, and your child may decide to end his or her participation in the study at any time.

There will be no risk, discomfort or inconvenience for your child. You child will not have to answer any questions that he or she does not want to answer. Each session will only begin if your child agrees to take part, and your child may decide to end his or her participation at any
time. The activities will be similar to the activities in a normal English lesson. Participation is voluntary. You may take your child out of the study at any time before, during or after each test has been given by contacting the child’s classroom teacher, the director of English-immersion programs or myself. This will have no effect on your child’s grades in school. If you decide to withdraw your child from the study, you can ask me to remove part or all of his or her information from the study. The students who withdraw from the study part way through will be sent to school library during the rest of the tests.

It is important to protect the privacy of the children who participate in this research. The information of you and your child will be kept confidential. Only I and my supervisor and committee members will have the access to the data. Only group results will be reported. Your child’s confidentiality will be maintained to the extent possible. The results from this study will be published in my thesis, academic and professional journals, and presented at academic conferences. In every case, results will be combined across many participants, so no individual will be identifiable. If the data are made available to other researchers for secondary analysis, your child’s identity will still be protected. In accordance with Queen’s policy, all test papers, children’s name file, digital audio files, and data will be retained for a minimum of five years; the data in this study will be retained indefinitely.

If you give your consent for your child to participate in this research study, please sign the attached consent form and return it to your child’s teacher. Please also answer the questions on the attached parental questionnaire. This questionnaire will help me to understand about your child’s English experiences outside of school. Please return this questionnaire with the consent form. Your input will contribute to greater understanding of the nature of reading comprehension in English-immersion students in China. The results of this research may help English teachers to know what skills they should focus on when teaching reading in English.

Thank you very much for considering your child’s participation in this research. Any questions about study participation may be directed to Miao Li at miao.li@queensu.ca, cell phone (China): 13654528661, or my supervisor Dr. John Kirby at John.Kirby@queensu.ca. Any ethical concerns about the study may be directed to the Chair of the General Research Ethics Board at 613-533-6081 or chair.GREB@queensu.ca

Sincerely,

Miao Li
Ph.D Candidate
Faculty of Education, Queen’s University
CCUEI Program
Appendix C
Consent Form

Please sign one copy of this Consent Form and return to Miao Li (by way of your child’s classroom teacher). Retain the second copy for your records.

I have read and kept a copy of the Letter of Information and the Consent Form. I have had any questions answered to my satisfaction.

I understand that I am being asked to allow my child to participate in the research project entitled *Understanding Reading Comprehension in Chinese English-immersion Students* and the purpose is to explore the cognitive processes underlying English reading comprehension in Chinese students who are in the English-immersion programs.

I understand that my child’s participation in this research will involve (1) individual testing which will take about 50 minutes, and (2) group testing which will take about 200 minutes. My child will be asked to complete a variety of oral language, reasoning, and reading activities, which will only begin and continue as long as he or she is willing and interested.

I understand that there are no known risks, discomforts or inconveniences associated with participation in the research study.

I understand that my child’s participation is voluntary. I can withdraw my child from the study at any time and that withdrawal will not affect my child’s grade in school. If I withdraw my child, I understand that I can ask for all or part of his or her information to be removed from the study, without consequences.

I understand that the combined information from this study will be published in the researcher’s thesis and academic and professional journals, and presented at academic conferences. If the information is made available to other researchers for further study, my child’s privacy will be protected.
I understand that my child will be audio recorded. I have been assured that the information collected will be kept confidential to the extent possible. All test papers, child’s name file, digital audio file, and data will be retained for a minimum of five years; the data in this study will be retained indefinitely.

I understand that by completing and submitting this questionnaire, I consent to my participation in the research.

Any questions about study participation may be directed to Miao Li at miao.li@queensu.ca, or my supervisor Dr. John Kirby at John.Kirby@queensu.ca. Any ethical concerns about the study may be directed to the Chair of the General Research Ethics Board at 613-533-6081 or chair.GREB@queensu.ca

☐ I consent for my child to participate.

Child’s name: ____________                     Child’s class _________________________
Child’s birthday: Year ____________  Month ____________ Day ____________
Name of Parent/Guardian: (please print) ______________________________________
Signature of Parent/Guardian: ________________________     Date: _______________
☐ I wish to receive a copy of the summary of findings (please write email or postal address below)
## Appendix D

### English Word Definition Test and Scoring Guidelines

<table>
<thead>
<tr>
<th>Word</th>
<th>Category</th>
<th>Function</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bus</td>
<td>vehicle, transportation method</td>
<td>people ride on, carries people, go somewhere, get to school/work</td>
<td>large/wheels/seats/passengers/bus driver/numbered, city/school, pay for it, on the street, public</td>
<td>protect environment, better than cars</td>
</tr>
<tr>
<td>Hand</td>
<td>part of body</td>
<td>get/grab/hold/write/eat/catch/do things, work,</td>
<td>has five fingers, at the end of arm, two hands, yellow, we should clean it every day</td>
<td>to make friendship, to say goodbye, make money</td>
</tr>
<tr>
<td>Apple</td>
<td>fruit</td>
<td>can be eaten</td>
<td>red/yellow/green color, round shape, has skin and seeds, sour and sweet, delicious, grows on trees, the king of the fruit</td>
<td>good for health, give vitamin C, get water and energy</td>
</tr>
<tr>
<td>Cow</td>
<td>(farm) animal</td>
<td>produce milk or meat, beef, help farmers plant crops</td>
<td>female, four legs, eat grass, white and black</td>
<td>make money, transportation in old times, give us leather</td>
</tr>
<tr>
<td>Tree</td>
<td>plant</td>
<td>give us oxygen, make wood, fire, table, paper, house, make air fresh, protect environment, photosynthesise, keep noise down</td>
<td>has trunk or stem, and branches, leaves are green in spring and yellow in autumn, tall or short, birds live in it, under it in the rain – dangerous, some produce fruits</td>
<td>beautiful, climb trees for fun, good for environment and nature, make money</td>
</tr>
<tr>
<td>Noodle</td>
<td>food, traditional</td>
<td>can be eaten</td>
<td>white, made from flour, long/thin, boiled/stir-fry/soup/with sauce on,</td>
<td>good for health, make us strong, give us</td>
</tr>
<tr>
<td></td>
<td>Chinese food</td>
<td>have many kinds, tastes good, eat it on birthday</td>
<td>energy</td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>--------------</td>
<td>-----------------------------------------------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td><strong>Shirt</strong></td>
<td>clothes</td>
<td>worn for upper body in summer or spring</td>
<td>has collar, long sleeve, many colors and sizes, buy in shops</td>
<td>wear for warm/cool, beautiful</td>
</tr>
<tr>
<td><strong>Farm</strong></td>
<td>a piece of land where the farmers and animals lives</td>
<td>used for growing crops or keeping animals</td>
<td>grow rice, plough land, fence, have grass, very big, many plants, in countryside or mountains, beautiful</td>
<td>relax ourselves, go there on holiday, farmers make money</td>
</tr>
<tr>
<td><strong>Television</strong></td>
<td>electronic instrument or tool</td>
<td>for entertainment, to get news, information and knowledge</td>
<td>has many programs or channels, has sound, pictures, the shape is like a box, different size, buy it in shop, expensive, every home owns it, many people like to watch it</td>
<td>for relax, bad for eyes and study, make people happy, open eyes or minds, make people not lonely</td>
</tr>
<tr>
<td><strong>Island</strong></td>
<td>a piece of land, part of our earth surrounded by water</td>
<td>used for vacation or people living,</td>
<td>surrounded by water/sea, in the middle of the oceans, name of a specific island, Taiwan island (the biggest island in China), has many beaches, large or small, go there by ship</td>
<td>beautiful, for holidays</td>
</tr>
<tr>
<td><strong>Drum</strong></td>
<td>instrument for making music or noise</td>
<td>can be used in many activities, making music or noise, for fun</td>
<td>beat with sticks or hands, skin, body, different sizes, made of wood, played in the concert, air inside</td>
<td>play for fun</td>
</tr>
<tr>
<td><strong>Envelope</strong></td>
<td>something to put letters/cards/money in</td>
<td>to send letters to others</td>
<td>usually made of paper, has stamps, address on, get it from post office</td>
<td>protect letters, not as useful/fast/popular as email</td>
</tr>
</tbody>
</table>
Appendix E
Multiple-meaning Vocabulary Test

In this test, please choose the sentences in which the given word is used correctly (circle the correct letter). For each word, there could be one or more right answers. There could even be no right answers. (Practice example: jam)

1. kind
   A. Our English teacher is very nice and kind.
   B. There are different kinds of fruit in the market.
   C. The children are sent to the kind by their parents.
   D. The kind of my chair is soft.

2. can
   A. I will drive a can to meet my mom.
   B. Can you answer this question?
   C. They are going to climb the can of mountain.
   D. The can of the river was very cool.

3. light
   A. The light is not bright enough to read.
   B. Let’s light a candle in the room.
   C. The coat is light but warm.
   D. This computer is lighter than that one.

4. miss
   A. Miss Wang is sitting at the table.
   B. Will you miss the teacher when she leaves?
   C. The table is miss of food.
   D. One of my books is missing.

5. case
   A. Push the case to start the computer.
   B. The case of the dog is very loud.
   C. Put the case on noodles.
   D. There are three books in my case.

6. line
   A. He lines a cigarette in the dark.
   B. She put a line of air in the room.
   C. I heard the line on the telephone.
   D. I like to eat ice-cream with the line.

7. date
   A. He dates his bike to the building.
   B. The date is made of wood.
   C. You should put your money in the date.
   D. What’s the date today?

8. mind
   A. Would you mind if I go to the washroom?
   B. Be careful not to drop the mind.
   C. I will keep it in mind.
   D. She has mind to bike.
9. play
A. The children play a football game on the field.
B. We are going to see a new play at the theatre.
C. I could hear music playing on the radio.
D. The play is very hard to colour.

10. ring
A. She opened the ring because she was angry yesterday.
B. He bought her a diamond ring for the wedding.
C. Will you answer the phone if it rings?
D. The ring is made of water.

11. present
A. This book was a present from my brother.
B. There were 200 people present at the meeting.
C. At present, I’m afraid I can’t help you.
D. The women wore beautiful dresses to the present.

12. rose
A. The rose of the tree is very rough.
B. He bought a dozen of roses for his girlfriend.
C. She is sitting in a rose watching TV.
D. The sun rose over the mountains.

13. run
A. She had a run of beef and rice for dinner.
B. I put on a run to keep warm.
C. Can you run fast?
D. The dog is shaped like a run.

14. second
A. The children were playing with a second.
B. February is the second month of the year.
C. I will be ready in a second.
D. The second is too sharp for the boys.

15. well
A. I hope you are well.
B. She speaks English very well.
C. Are they coming as well?
D. Vegetables well taste good.
## Appendix F

### Morphological Awareness Test and Scoring

<table>
<thead>
<tr>
<th></th>
<th>Word</th>
<th>Scoring</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Circled letters for 2 points</td>
</tr>
<tr>
<td>1</td>
<td>desks</td>
<td>desk</td>
</tr>
<tr>
<td>2</td>
<td>ruder</td>
<td>rud</td>
</tr>
<tr>
<td>3</td>
<td>caring</td>
<td>car</td>
</tr>
<tr>
<td>4</td>
<td>classes</td>
<td>class</td>
</tr>
<tr>
<td>5</td>
<td>scarred</td>
<td>scar</td>
</tr>
<tr>
<td>6</td>
<td>disease</td>
<td>ease</td>
</tr>
<tr>
<td>7</td>
<td>staring</td>
<td>star</td>
</tr>
<tr>
<td>8</td>
<td>because</td>
<td>cause</td>
</tr>
<tr>
<td>9</td>
<td>changing</td>
<td>chang</td>
</tr>
<tr>
<td>10</td>
<td>socially</td>
<td>soci</td>
</tr>
<tr>
<td>11</td>
<td>starring</td>
<td>star</td>
</tr>
<tr>
<td>12</td>
<td>finished</td>
<td>fin</td>
</tr>
<tr>
<td>13</td>
<td>usefully</td>
<td>use</td>
</tr>
<tr>
<td>14</td>
<td>educating</td>
<td>duc</td>
</tr>
<tr>
<td>15</td>
<td>carefully</td>
<td>care</td>
</tr>
<tr>
<td>16</td>
<td>unmarried</td>
<td>marri</td>
</tr>
<tr>
<td>17</td>
<td>reproduce</td>
<td>duce</td>
</tr>
<tr>
<td>18</td>
<td>condensing</td>
<td>dens</td>
</tr>
<tr>
<td>19</td>
<td>reelected</td>
<td>lect</td>
</tr>
<tr>
<td>20</td>
<td>conscious</td>
<td>sci</td>
</tr>
<tr>
<td>21</td>
<td>dictation</td>
<td>dict</td>
</tr>
<tr>
<td>22</td>
<td>impossible</td>
<td>poss</td>
</tr>
<tr>
<td>23</td>
<td>unpleasant</td>
<td>pleas</td>
</tr>
<tr>
<td>24</td>
<td>refereeing</td>
<td>fer</td>
</tr>
<tr>
<td>25</td>
<td>assignment</td>
<td>sign</td>
</tr>
<tr>
<td></td>
<td>decreasing</td>
<td>creas</td>
</tr>
<tr>
<td>---</td>
<td>------------</td>
<td>------</td>
</tr>
<tr>
<td>26</td>
<td>adaptation</td>
<td>apt</td>
</tr>
<tr>
<td>27</td>
<td>constructed</td>
<td>struct construct</td>
</tr>
<tr>
<td>28</td>
<td>insensitive</td>
<td>sense</td>
</tr>
<tr>
<td>29</td>
<td>precautions</td>
<td>caut caution</td>
</tr>
<tr>
<td>30</td>
<td>undiscovered</td>
<td>cover</td>
</tr>
<tr>
<td>31</td>
<td>accompanying</td>
<td>pan company</td>
</tr>
<tr>
<td>32</td>
<td>victoriously</td>
<td>victor</td>
</tr>
<tr>
<td>33</td>
<td>incorruptible</td>
<td>rupt corrupt</td>
</tr>
<tr>
<td>34</td>
<td>insignificance</td>
<td>sign</td>
</tr>
<tr>
<td>35</td>
<td>acknowledgement</td>
<td>know</td>
</tr>
<tr>
<td>36</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix G

Inference and Strategy Tests

Please read the passages and choose the one best answer for each question and fill in the answer sheet.

1. “You’re going – that’s all there is to it,” Tom said, staring at Mike.
   “In your dreams,” his friend replied. Mike scowled across the table at Tom with an angry expression on his face.
   “What kind of chicken are you?” Tom went on.
   “Okay, I admit it. I’m a chicken, if you really have to know,” Mike shot back.
   The two friends had been through this conversation many times before. Tom was always pushing to try something adventurous. Mike would always say no, and then give in.
   It was a pattern that had started back in Grade One, back when Tom dared Mike to ride his bike down on Jefferson Street hill.
   Mike kept saying no. But Tom kept pushing and daring him. That’s how it always went, until Mike gave in. And Mike ended up with a sprained wrist.
   “You know you want to go on this trip,” Tom told him. “Look at it this way: You can be back at school, doing math problems or a spelling test – or you can be walking in the mountains, in the fresh air.”
   “Right,” Mike replied, laughing. “Freezing in the fresh air. With my luck, I’ll probably fall off a cliff.” He paused and looked right at Tom. “You can’t talk me into it this time. I’m not going and that’s final.”
   Tom sat back in his chair. “Well, at least come to the information meeting with me. Keep me company, okay?”

1. What does Mike mean when he says, “In your dreams”?
   A. Tom forgets his dreams.
   B. Tom needs to make dreams.
   C. Tom needs a good sleep.
   D. It won’t happen in real life.

2. When Mike “scowled” (line 2) how did he feel?
   A. happy      B. annoyed     C. eager      D. sad

3. Why doesn’t Mike want to go on the trip?
   A. Mike is tired that wants a good sleep
   B. Mike thinks he would get hurt or cold.
   C. Mike plans to enter a bike race.
   D. Mike has to go to a trip with his family.
4. Why do you think Tom asked Mike to come to the meeting?
   A. Because Tom wants Mike to meet more people at the meeting.
   B. Tom wants Mike to eat some good food at the meeting.
   C. Tom doesn’t want to go by himself.
   D. Tom feels safer if Mike came with him.

5. In what ways are the two friends different?
   A. Tom is more adventurous than Mike.
   B. Mike loves cycling more than Tom does.
   C. Tom is more talkative than Mike.
   D. Mike is lazier than Tom.

6. Which of the following is most likely to happen in the end?
   A. Mike will stop being friends with Tom.
   B. Mike will go on the trip with Tom.
   C. Mike will have a fight with Tom.
   D. Tom will decide not to go to the trip.

7. Which of the following is the best summary of the passage?
   A. Tom is having an argument with his friend Mike.
   B. Mike is asking Tom not to go on the trip.
   C. Tom is having a fight with his friend Mike.
   D. Tom is trying to talk his friend Mike into going on the trip with him.

8. Which of the following can’t be true after what is said in the passage?
   A. Tom really wants to go on the trip with Mike.
   B. Mike doesn’t want to go on the trip.
   C. The trip will be very boring.
   D. Tom and Mike are good friends.

9. What is another way to say “I’m a chicken” (line 5)?
   A. I’m not brave.
   B. I’m very young.
   C. I was born in the year of the chicken.
   D. I like eating chicken.

10. Which of these sentences from the passage would tell you the most about the passage?
   A. “What kind of chicken are you?” (line 4)
   B. “I’m a chicken, if you really have to know.” (line 5)
   C. It was a pattern that had started back in Grade One. (line 9)
   D. You know you want to get on this trip. (line 13)
Lots of people all over the world have heard of the soft drink called Coca-Cola. But not many people know the real story about how this drink was invented.

Coca-Cola was the invention of a Mr. John Pemberton. Although he wasn’t a doctor, most people called him Dr. Pemberton. He was a druggist in a town in the South. Dr. Pemberton liked to invent new things. He lived during the time just after the Civil War.

One day, Dr. Pemberton decided to make a headache medicine. He made it from nuts, fruits, and leaves. He also added the drugs necessary to cure a headache. Dr. Pemberton now thought he had something to sell that tasted good.

In the summer of 1886, Dr. Pemberton took a jug of this headache syrup to one of the best drugstores in Atlanta, Georgia. He told the manager of the drugstore to mix some of the syrup with water and have just people with headaches drink it. At first, it did not sell very well. Then one day a clerk sold some of the new medicine to a customer with a bad headache. But instead of using regular water, he used carbonated water by accident. Carbonated water has bubbles in it. Everyone loved this new change, and carbonated water is still used in Coca-Cola today.

Most of the medicine that cures headaches was taken out of Coca-Cola as time went on. But Dr. Pemberton’s drink is still one of the world’s favourite soft drinks.

1. Why didn’t the medicine sell very well at first?
   A. Because Dr. Pemberton was not a real doctor.
   B. Because the medicine was not effective.
   C. Because the syrup was mixed with regular water.
   D. Because the medicine was too expensive.

2. Why could one say that Coca-Cola became popular because of a mistake?
   A. Because a clerk mixed the syrup with regular water.
   B. Because a clerk accidentally mixed the syrup with carbonated water.
   C. Because the medicine was taken out of Coca-Cola.
   D. Because a patient had an accident after taking the drug.

3. What is different between today’s Coca-Cola and the original version?
   A. No headache medicine
   B. No carbonated water
   C. No syrup
   D. No sugar

4. What’s the difference between regular water and carbonated water?
   A. Carbonated water tastes salty.
   B. Regular water tastes bad.
   C. Carbonated water tastes sweet.
   D. Regular water doesn’t have bubbles.
5. In “Everyone loved this new change” (line 15), what does this new change refer to?
   A. Adding carbonated water.
   B. Taking out headache medicine.
   C. Adding fruits.
   D. Adding nuts.

6. Which of the following is the best summary of the passage?
   A. How Dr. Pemberton made his headache drug
   B. How Coca-Cola was invented
   C. How headache medicine was taken out of Coca-Cola
   D. How Coca-Cola cured headaches

7. Which of the following is most important about the passage?
   A. How Coca-Cola protected people from headaches.
   B. Dr. Pemberton made a headache medicine from nuts, fruits, and leaves.
   C. The headache medicine was originally sold in Atlanta, Georgia.
   D. Coca-Cola became popular when carbonated water was added by accident.

8. Which of the following questions could NOT be answered from the passage?
   A. How did Coca-Cola become popular?
   B. How was the headache medicine taken out of the Coca-Cola?
   C. What was the original purpose of Coca-Cola?
   D. Did Dr. Pemberton’s headache medicine sell very well at the beginning?

9. What would be the best strategy for finding the date the headache medicine was originally sold?
   A. Re-read the entire passage
   B. Ask a friend
   C. Skim the paragraph for dates
   D. Re-read the first sentence

10. What is the best title for this passage?
    A. The Invention of Coca-Cola
    B. The Success of Coca-Cola
    C. The Popularity of Coca-Cola
    D. The Failure of Coca-Cola
Appendix H  
Summary Writing Test

Read the following text. You may make notes, but when you are finished reading, we will collect the texts and your notes. Then we will ask you to write a summary of the text.

Cities

In the Middle Ages, cities were growing larger and larger, but had many problems. The Middle Ages were about 1000 years ago. You probably know them as the time of knights fighting with long lances on horseback. These fights were called jousts.

The cities had grown larger because so many people wanted to move there to make more money or to have an easier life. But there were not enough houses. People had to live in very crowded conditions, so sickness spread very quickly. How would you like to share your house with five other families?

Sometimes new houses were built poorly. The houses were close together, so that when a fire started, many houses caught fire. Many people died. The city of London in England had a fire in which 70,000 homes were destroyed – that fire had started in a bakery!

Garbage was another problem. They did not have any garbage pick-up, so people just tossed their garbage out the window into the street. Herds of wild pigs ran through the streets eating the garbage! There were rats everywhere, and the rats also spread disease.

Cities in the Middle Ages did not have any police. Powerful people could do anything they wanted, and the poor people just had to suffer. Criminals could walk into your house and take anything they wanted!

There were many problems, but people worked hard to solve them. Cities still have many problems, but they are not as bad as the ones the cities in the Middle Ages had.
Appendix I

Summary Writing Scoring

M2 = Main Ideas, M1 = Important Details

M2-1 Cities in Middle Ages had problems, problems then and now
M1-1 cities growing larger
M1-2 people moving there
M1-3 1000 years ago

M2-2 Problem of crowding leads to disease
M1-4 crowding
M1-5 more money/easier life
M1-6 not enough houses
M1-7 so sickness spread

M2-3 Problem of poorly built/crowded houses leads to fire
M1-8 poorly built
M1-9 Fires spread
M1-10 People died

M2-4 Problem of garbage leads to disease
M1-11 garbage
M1-12 no garbage removal
M1-13 garbage thrown in street
M1-14 rats spread disease

M2-5 Problem of no police leads to criminal taking
M1-15 no police
M1-16 Powerful people could do as they wanted
M1-17 Poor/weak people suffered

M1-18 people worked to solve problems
M1-19 still problems