

THE INFLUENCE OF CROSS-LINGUISTIC INPUT AND L2 PROFICIENCY ON L2
READING COMPREHENSION AMONG SPANISH-SPEAKING ADULTS
LEARNING ENGLISH AS A SECOND LANGUAGE

by

ASTRID SUSSETTE RODRÍGUEZ

A dissertation submitted to the Graduate Faculty in Educational Psychology in partial
fulfillment of the requirements for the degree of Doctor of Philosophy,
The City University of New York

2010

© 2010

ASTRID SUSSETTE RODRÍGUEZ

All Rights Reserved

This manuscript has been read and accepted for the Graduate Faculty in Educational Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Date	Dr. Linnea C. Ehri Chair of Examining Committee
------	--

Date	Dr. Mario A. Kelly Program Executive Officer
------	---

Dr. Linnea C. Ehri, Distinguished Professor, GSUC-CUNY

Dr. Jay Verkuilen, Assistant Professor, GSUC-CUNY

Dr. Emilia C. Lopez, Professor, QC-CUNY

Dr. Helen L. Johnson, Professor, GSUC-CUNY

Dr. Miriam Eisenstein Ebsworth, Professor, NYU

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

THE INFLUENCE OF CROSS-LINGUISTIC INPUT AND L2 PROFICIENCY ON L2
READING COMPREHENSION AMONG SPANISH-SPEAKING ADULTS
LEARNING ENGLISH AS A SECOND LANGUAGE

by

Astrid Sussette Rodríguez

Adviser: Distinguished Professor Linnea C. Ehri

Developing literacy and language proficiency in English is essential to thrive in school and in the workforce in American society. Research on cross-linguistic influences on text-level skills is scant, especially studies investigating reading comprehension among language-minority adults. The present study investigated the effects of cross-linguistic input and second language proficiency on second language reading comprehension among Spanish-speaking adults enrolled in English as a Second Language (ESL) classes at a community college in New York City. The main research question was whether language-minority adults would comprehend printed text better if they read it in Spanish (L1) followed by English (L2) than if they read the text twice only in English. An experimental study using a 2 (lower L2 proficiency *vs.* higher L2 proficiency) x 2 (think-aloud *vs.* no think-aloud) x 2 (Spanish/English *vs.* English/English) crossed factorial design was conducted. Eighty students were randomly assigned to the Spanish/English or English/English reading conditions within each L2 proficiency level and think-aloud conditions. Reading comprehension was assessed via a

verbal recall task, eight open-ended questions, and a 43-item cloze task. The results revealed that students with lower L2 proficiency benefitted substantially from reading the text in both languages as evident by their performance on the recall and Q&A reading comprehension tasks. In contrast, L1 input did not provide an added advantage over reading the text twice in English for students with higher L2 proficiency. A reactive effect to the think-aloud procedure was found for students with lower L2 proficiency on the Q&A and cloze tasks. On the Q&A task, among those who completed the think-aloud procedure, students who read the text in both languages outperformed students who read the text only in English. On the cloze task, students who did not perform the think-aloud procedure outperformed students who did the think-aloud. In addition, it was found that the reading comprehension assessment methods had a differential effect on students' ability to demonstrate the degree to which they comprehended the text. Results showed that performance was lowest on the recall task across all groups. Other findings and implications are discussed.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to Dr. Linnea C. Ehri for serving as the Chair of my dissertation committee. Thank you very much for guiding me and sharing so much of your time, as well as for your consideration, patience, and support throughout this arduous process. I deeply admire your scholarship and demand for excellence. I am very fortunate to have had you as my advisor.

I am also thankful to Dr. Jay Verkuilen and Dr. Emilia C. Lopez for agreeing to join the dissertation committee and for sharing your expertise pertaining to statistical analysis and knowledge of literacy development among English language learners, respectively. Your suggestions really strengthened my study approach and analysis. To Dr. Helen L. Johnson and Dr. Miriam Eisenstein Ebsworth, thank you for agreeing to read my dissertation and for your input. To all, thank you for your contributions and critiques.

Of utmost importance, this dissertation could have not been completed without the participation of the faculty and staff of the Borough of Manhattan Community College (BMCC) where I collected the data. Special thanks to Professor Gay Brookes, Chair of the Developmental Skills Department and its faculty for letting me into your classrooms so that I could recruit participants. I also wish to express my deepest gratitude to the library staff, in particular to Guerda Baucicaut, for accommodating me during the two years of the data collection process, for such long periods of time, and many times, on such short notice. Very special thanks to the students that participated; I wish you all the best!

I would also like to thank the various entities that provided the financial support, small and large, for the completion of this study. Many thanks to the Office of Student Affairs for the Sue Rosenberg Zalk Student Travel and Research Fund Award and for the Doctoral Student Research Grant; and to the Alliance for Graduate Education and the Professoriate (AGEP) and the National Science Foundation (NSF) for the Continuing Student Research Award.

I am forever grateful to the three *jefes* I worked with since this journey began. Thank you for your unwavering support and for accommodating my work schedule so that I could attend school whenever necessary, collect data during the pilot and dissertation studies, and write this dissertation. Thank you very much John G. O'Brien, may God have you in his glory; Rafael Hilerio, and Jimmy Jung. I could have not completed this dissertation without you.

Last, but not least, I want to thank my family, specially my mother and sister, and my closest friends, Jerry and Linda, for celebrating even the smallest accomplishment and for encouraging me to forge ahead and not letting me dismay during the most difficult times. I love you all.

*Ask and it will be given to you;
seek and you will find;
knock and the door will be opened to you.
For everyone who asks receives;
he who seeks finds;
and to him who knocks,
the door will be opened.
Matthew 7:7-8*

Table of Contents

Chapter 1: Introduction	1
Chapter 2: Literature Review	11
<i>Bilingual Education</i>	11
<i>Hypotheses About Cross-Linguistic Transfer</i>	15
<i>Reading Comprehension Processes</i>	25
<i>Reading Comprehension Instruction</i>	29
<i>Cross-Linguistic Input and Reading Comprehension</i>	31
<i>Reading Comprehension Assessment</i>	40
<i>An Inside Look Into Reading Comprehension</i>	44
<i>Summary</i>	48
Chapter 3: Rationale and Hypotheses	51
<i>Preliminary Study</i>	51
<i>Dissertation Study</i>	54
Chapter 4: Method	58
<i>Participants</i>	58
<i>Materials</i>	59
<i>Procedure</i>	70
<i>Design and Statistical Analyses</i>	73
Chapter 5: Results	78

<i>Characteristics of Participants</i>	78
<i>L1 and L2 Language Proficiency</i>	79
<i>Time of Recruitment Effects</i>	80
<i>Comparison of Experimental Groups Prior to Treatment</i>	86
<i>Effects on Reading Comprehension Among Students With Lower L2 Proficiency</i>	88
<i>Effects on Reading Comprehension Among Students With Higher L2 Proficiency</i>	96
<i>Comparison of Effect Sizes Between English Language Proficiency Groups</i>	100
<i>Comparison of Performance Across Comprehension Measures</i>	102
<i>Correlations and Regression Analyses</i>	106
<i>Online Reading Processes as Revealed by the Think-Aloud</i>	116
<i>Exit Interview</i>	122
Chapter 6: Discussion	126
<i>Effect of L1 Input and L2 Proficiency on Reading Comprehension</i>	128
<i>Effect of Think-Aloud on Reading Comprehension</i>	132
<i>Performance Across Reading Comprehension Tasks</i>	135
<i>Strategy Use as Revealed by Think-Aloud Protocols</i>	139
<i>Summary Comments</i>	142
<i>Implications for Instruction</i>	144
Appendixes	148
References	169

List of Tables

		Page
Table 1	Mean Performance and Standard Deviation on the Component Tests of the Woodcock-Muñoz Language Survey-Revised by Level of English Proficiency	81
Table 2	Demographic Characteristics of Students With Lower English Proficiency by Year of Recruitment	83
Table 3	Mean Performance and Standard Deviation of Students With Lower English Proficiency on the Component Tests of the Woodcock-Muñoz Language Survey-Revised by Year of Recruitment	84
Table 4	Demographic Characteristics of Students With Higher English Proficiency by Year of Recruitment	85
Table 5	Mean Performance and Standard Deviation of Students With Higher English Proficiency on the Component Tests of the Woodcock-Muñoz Language Survey-Revised by Year of Recruitment	87
Table 6	Results of MANOVA for the Component Tests of the Woodcock-Muñoz Language Survey-Revised as a Function of English Proficiency, Think-Aloud Procedure, and Language of Text for Students Prior to Treatment	89
Table 7	Mean Performance and Standard Deviation on the Three Measures of Reading Comprehension by the Four Experimental Groups on Each Think-Aloud Condition and Overall	90
Table 8	Results of ANOVAs and Partial Eta-Squared Estimates for the Three Measures of Reading Comprehension as a Function of Think-Aloud Procedure and Language of Text for Students With Lower English Proficiency	91
Table 9	Results of ANOVAs and Partial Eta-Squared Estimates for the Three Measures of Reading Comprehension as a Function of Think-Aloud Procedure and Language of Text for Students With Higher English Proficiency	98

Table 10	Correlations Between the Woodcock-Muñoz Language Survey-Revised Pretest Measures in English and Spanish and the Three Measures of Reading Comprehension	107
Table 11	Correlations Between the Woodcock-Muñoz Language Survey-Revised Pretest Measures in English and Spanish and the Three Measures of Reading Comprehension as a Function of Language of Text	108
Table 12	Variance Explained by Woodcock-Muñoz Language Survey-Revised Measures in English (L2) and Spanish (L1) on the Three Measures of Reading Comprehension for the Overall Sample and as a Function of Language of Text	112
Table 13	Number and Total Percent of Strategies Verbalized by Type for the Four Experimental Groups During the Think-Aloud Procedure	119
Table 14	Responses to the Questions on the Exit Interview	123

List of Figures

		Page
Figure 1	Study Design	72
Figure 2	Mean Scores and 95% Confidence Intervals on the Recall Comprehension Test by Students With Lower English Proficiency as a Function of Think-Aloud and Language of Text	93
Figure 3	Mean Scores and 95% Confidence Intervals on the Q&A Comprehension Test by Students With Lower English Proficiency as a Function of Think-Aloud and Language of Text	95
Figure 4	Mean Scores and 95% Confidence Intervals on the Cloze Comprehension Test by Students With Lower English Proficiency as a Function of Think-Aloud and Language of Text	97
Figure 5	Mean Percent Scores on the Three Reading Comprehension Measures by Students With Lower English Proficiency Who Read the Text in Spanish and English (SE) or Twice in English (EE)	104
Figure 6	Mean Percent Scores on the Three Reading Comprehension Measures by Students With Higher English Proficiency Who Read the Text in Spanish and English (SE) or Twice in English (EE)	105
Figure 7	Decomposition of Variance Accounted for by Woodcock-Muñoz Language Survey-Revised Composite Scores in the Spanish (L1) and English (L2) Forms on the Three Measures of Reading Comprehension	113
Figure 8	Decomposition of Variance Accounted for by Woodcock-Muñoz Language Survey-Revised Composite Scores in the Spanish (L1) and English (L2) Forms on the Three Measures of Reading Comprehension by Students Who Read the Text in Spanish and English (SE) or Twice in English (EE)	114

List of Appendixes

	Page	
Appendix A	Written Informed Consent Form	149
Appendix B	Demographics and Language Background Interview	153
Appendix C	Reading Passage	155
Appendix D	General Instructions: Meeting # 2	157
Appendix E	Recall Task Scoring Sheet	162
Appendix F	Q&A Task and Scoring Key	163
Appendix G	Cloze Task and Scoring Key	165
Appendix H	Exit Interview	167

Chapter 1: Introduction

The purpose of this study was to investigate the effect of native-language (L1) input and second-language (L2) proficiency on L2 reading comprehension among language-minority adults whose native language was Spanish. The primary question of interest was whether adults who were learning English as a second language would be able to comprehend printed text more effectively if they read the text first in L1 and then in L2, than if they read the text twice only in L2. It was expected that reading a text in one's native language would provide a scaffold in the form of background knowledge that would allow readers to make connections between their first language and the new information in the second language, thus facilitating L2 reading. That is, learners would use what they had learned in L1 to increase their understanding of the L2 input in terms of its overall content as well as to better understand its surface features.

Adult Spanish-speakers constitute about 9.5% of the United States population (American Fact Finder, 2008). They comprise the largest percentage of language-minority adults in the U.S. as well as the largest growing language-minority population. According to the U.S. Department of Education (February, 2006), over 2.5 million adults across the U.S. receive ESL classes through federally-funded programs such as English as a Second Language (ESL), Adult Basic Education (ABE), and Adult Secondary Education (ASE). A literacy skill that has proven particularly difficult to acquire for this population is reading, and the continuing educational underachievement of this population has been referred to as an "urgent and perplexing" issue in American education (Education Commission of the States, 2004, p. 1).

Developing literacy and language proficiency in English is essential not only for academic success but also for life-long success. There is a positive relationship between literacy levels and many socioeconomic-related factors (Kutner et al., 2007). For example, individuals with higher levels of literacy earn higher incomes than individuals with low literacy levels. In addition, individuals with higher levels of literacy are more likely to be employed in full-time jobs and have professional occupations, and are less likely to receive public assistance. On the other hand, individuals with lower literacy levels are more likely to be employed in part-time jobs or in service occupations, and are more likely to receive public assistance and for a longer period of time. According to the National Center for Education Statistics (Kutner et al.), Hispanic adults have the lowest literacy levels compared to Caucasian and African-American adults, and for many, their substandard reading skills limit their job opportunities “a lot.”

Reading is a complex activity. Becoming a skillful reader in the native language is a difficult task; reading in a second language can certainly become an even more complicated process. Alderson (1984) suggested that poor reading in L2 is a function of either inadequate L2 proficiency or poor L1 reading ability, or both, whereas good reading in L2 is a result of good L1 reading ability or good L2 competence, or both. The concept of cross-linguistic transfer as conceptualized by Cummins (1979, 2000) inspired much of the research on transfer of cognitive-academic skills, and it has been pivotal in providing a rationale for bilingual education. According to Cummins’ linguistic interdependence hypothesis, the development of competence in L2 is partly a function of competence already developed in L1 at the time when intensive exposure to L2 begins. That is, literacy-related language skills and knowledge acquired in L1 (i.e., *cognitive-*

academic language proficiency [CALP]) can positively influence the development of literacy skills in L2 since they are developmentally linked to a *common underlying proficiency*. In addition, Cummins' linguistic threshold hypothesis suggests that there *may be* threshold levels of linguistic competence in L2 that must be attained in order for the L1 to positively influence cognitive and academic functioning in L2.

Research investigating the language and literacy development of second language learners has shown that some L1 and L2 literacy skills and sub-skills are related in important ways; and that well-developed literacy skills in L1 can facilitate the development of certain literacy skills in L2 -- especially when the L1 and L2 languages are typologically close such as Spanish-English (August & Shanahan, 2006). Furthermore, the use of L1 in the classroom has shown to present no adverse effects on the development of L2 academic skills (Cummins, 2000). Quite the contrary, bilingual children have shown consistent cognitive and metalinguistic advantages early in the process of development (Diaz & Klinger, 1991). With such positive evidence you would expect policy-makers to make bilingual education one of their top priorities.

In countries such as United States, Belgium, Bolivia, Canada, Nigeria, and Sweden, various types of bilingual programs have been implemented (Paulston, 1988). In the United States, however, the use of bilingual programs is a topic of increasing controversy in spite of the evidence that such programs do help children develop literacy skills in English (e.g., Slavin & Cheung, 2004). It is not possible to draw generalizations about classroom practices in bilingual programs because they vary between and within countries. In the New York City public school system, for instance, the Department of Education offers English Language Learners (ELLs) three programs: English as a Second

Language (ESL), Transitional Bilingual Education (TBE), and Dual Language (DL).

Whatever the program, the objective is for ELLs to become literate and fluent in English within three years or less (NYCDOE, 2009). The present study sought to add to our understanding regarding the potential benefits of bilingual instruction on reading comprehension among young adults who were ESL students.

Reading comprehension entails a complex combination of cognitive and linguistic processes. Various reading models have been proposed, but most reading models emphasize the interaction of lower- and higher-level processes and knowledge sources as the reader attempts to make sense of the text (e.g., Ehri, 1998; Rumelhart, 1977; Stanovich, 1980, 1984). One of the most influential models of text comprehension is Kintsch's (1988, 2004, 2005) construction-integration model of text comprehension. According to Kintsch, the reader may form three qualitatively different representations of the text, each representing different levels of understanding. At the first level of representation, understanding is superficial (i.e., *surface-level*) and entails word-level processing of the text. At the second level (i.e., *semantic-level*), the reader attains understanding of the explicit meaning of the text (i.e., *textbase*), however, comprehension is still shallow. In order to reach deep understanding, the reader must construct a *situation model*, which corresponds to the third level of text representation, by integrating the information in the text with information from his/her prior knowledge.

Based on these interactive models of reading, we can expect reading comprehension among readers with low levels of L2 competence to be compromised by a deficiency of lower-level skills (e.g., lexical knowledge); yet, they may compensate for this deficiency by using higher-level skills (e.g., activating prior knowledge).

Furthermore, according to Cummins' (1979, 2000) linguistic interdependence and threshold hypotheses, with respect to reading, we should expect second-language learners to be able to easily transfer their ability to extract meaning from printed text from L1 to L2, as long as experiences in L1 have afforded the development of strategic reading skills (e.g., Goldman, Reyes, & Varnhagen, 1984). However, we should also expect that an individual's ability to transfer literacy-related language skills and knowledge from the native language to the target language will be hampered by L2 proficiency when is below a threshold. That is, low L2 proficiency would create a *linguistic ceiling* that "short-circuits" the individual's ability to transfer effective comprehension processes (e.g., use of prior knowledge and strategic reading skills) that are used during L1 reading (e.g., Clarke, 1979; Lee & Schallert, 1997). This in turn may lead to the formation of an incomplete textbase and an inadequate situation model. Consequently, the meaning of the text will not be fully understood.

Taking these propositions into account, a question of interest in the present study was whether exposure to print in one language could provide the necessary scaffold to facilitate reading performance in another language. That is, is it possible that concurrent exposure to print in the native- and second-language would facilitate connections among vocabulary and content prior knowledge, thus, allowing readers to form a mental representation of the L2 text irrespective of the reader's level of L2 proficiency? Comprehending the meaning of texts is often difficult for language-minority students because they do not have the extensive vocabulary that native-speakers have. For instance, the vocabulary of adults at the beginning stages of learning English only consists of about 2,000 to 7,000 words compared to 10,000 to 100,000 words among

fluent English-speakers (Hadley, 1993). In the review of the literature, only a handful of studies were identified that specifically investigated how use of the native language may help readers understand input in a second language.

In a recent study, for instance, Friesen and Jared (2007) examined whether cross-language transfer effects were present at both the semantic- and surface-level, as measured by overall reading times and duration of eye gazes and fixations on specific words, when students read pairs of passages that were identical, translations, or unrelated. Among other things, they found that exposure to a text written in the native language (English) facilitated reading performance of the same text in the second language (French) compared to reading unrelated texts for both high- and low-ability students. The authors interpreted the results as suggesting that students transferred knowledge from L1 to reading L2 translations of the texts.

The *bilingual reading method* is not a novel concept. In 1970, Parent and Belasco investigated whether exposure to parallel translation modes of excerpts of a novel in English-French would be beneficial for reading comprehension compared to L2-only materials across high- and low-ability students. The results revealed that among low-ability students, the bilingual reading groups outperformed the L2-only reading group on weekly multiple-choice tests. In contrast, they found no significant differences in the performance of high-ability students.

Along similar research lines, Jarvis and Jensen (1982) examined the extent to which reading parallel translations of dialogs from a film script (English-Russian) would facilitate reading comprehension in the second language across beginning- and advanced-level learners. The passages, which were rated as “easy” or “difficult,” were presented

with one of four reading aids: occasional side glosses, standard parallel translations, word-for-word parallel translations, or standard plus word-for-word parallel translations. The results showed that the translation modes resulted in higher scores on multiple-choice cloze tests for both beginning- and advanced-level students across all four tasks, with the exception of advanced-level students on the cloze tasks for “easy” content. On this task, advanced-level students performed better when occasional side glosses were available than when word-for-word parallel translations were available.

The most recent study that investigated the effectiveness of bilingual repeated-readings on reading comprehension was conducted by Lasisi, Falodun, and Onyehalu (1988). They investigated whether reading culturally unfamiliar stories in the native language (Yoruba) would facilitate reading comprehension in the first- and second-language (English). Passages were “plot-based” and “theme-based” prose presented in L1, L2 or in combination. Although students tended to perform better in the L1-L2 condition than in the L2-only condition, no significant differences in performance on L2 multiple-choice tests were found when children read the passages in both languages compared to when they read the passages only in L2.

Taken together, the studies reviewed provide support for the hypothesis that L1 input may serve as a scaffold for increasing L2 reading comprehension. Although the evidence provided by these studies is inconclusive, it is also possible that the effects might be stronger for students with lower L2 proficiency, whereas students with higher L2 proficiency may have reached a threshold for which exposure to L1 no longer provides an added advantage.

With respect to assessing reading comprehension, however, even when students have comprehended what they read, test type can positively or negatively affect their ability to demonstrate what they have understood (Bachman, 1990). Reading comprehension may be assessed using several informal methods requiring different levels of linguistic knowledge and production. The available research suggests that not all reading comprehension assessments are equivalent and that there is no one best method for assessing reading comprehension. In fact, students' ability to demonstrate comprehension may be affected by testing method (e.g., Shohamy, 1984; Wolf, 1993). Therefore, it is important to use multiple measures when testing reading comprehension. According to Bachman, *constructed* response tasks such as open-ended questions, cloze tasks, and recall protocols may provide a more authentic measure of comprehension compared to other traditional methods. In the present study, three methods requiring different degrees of comprehension, production, and linguistic skill were used to assess reading comprehension.

In addition to assessing reading comprehension as a “product” through informal assessments, it was of interest in the present study to assess the reading comprehension “processes” that took place as students read the text. Think-aloud procedures have been useful for understanding how people process text, particularly, what “good readers” do during reading (Pressley & Afflerbach, 1995). Encoding non-verbal thoughts into verbal form requires an additional task. According to Ericsson and Simon (1993), however, think-aloud instructions that ask subjects to verbalize their thoughts without requiring an explanation of their thought processes, only adds time-on-task without changing “the course or structure of the thought processes” (p. 106). Only a couple of studies to date

have empirically tested the issue of reactivity of thinking aloud while reading in L2 on reading comprehension (Bowles & Leow, 2005; Leow & Morgan-Short, 2004). When participants were instructed to “simply verbalize” their thoughts as they read the text, both studies found that thinking aloud had no effect on reading comprehension as measured by tasks with a multiple-choice format. The present study included both a think-aloud and a no-think-aloud condition to assess whether the think-aloud procedure had an effect on reading comprehension.

The aim of the present study was to provide evidence justifying the use of native language in classrooms servicing language-minority adults. The main hypothesis tested was whether students would be able to comprehend an L2 text better if they read it in L1 followed by L2 than if they read it twice only in L2. It was expected that reading a text in one’s native language would provide a scaffold in the form of background knowledge. The concurrent L1-L2 input provided would, in turn, allow readers to make connections between knowledge in their first language and the new information in the second language, thus facilitating learning and performance in the second language.

Two simultaneous hypotheses and alternative outcomes were entertained. If, on the one hand, students’ ability to successfully transfer native-language knowledge and skills when reading the L2 text was dependent upon an L2 threshold of linguistic competence below which L1 could not positively impact cognitive functioning in L2 as proposed by Cummins (1979, 2000), it was expected that students with lower L2 proficiency would perform poorly across comprehension assessments even though they had good CALP skills in L1. In contrast, it was expected that students with higher L2 proficiency would be able to successfully transfer native-language literacy skills and

knowledge when reading the L2 text and their comprehension would be not be hampered. Hence, they would perform well across comprehension tasks. However, if, at the same time, L1 input indeed served as a scaffold to facilitate more accomplished performance in L2, it was expected that students' performance on L2 reading comprehension tasks would be better with than without L1 input. Furthermore, the difference in performance between reading conditions (i.e., L1-L2 vs. L2-L2) was expected to be greater among students with lower L2 proficiency than among students with higher L2 proficiency because the former would need more language support from L1.

The present research study holds important implications for theory and practice. Theoretically, it adds to our understanding of cross-linguistic influences on reading comprehension among young adults as well as to our understanding of how transfer effects interact with L2 proficiency. From a practical standpoint, this research can help inform the role of instruction in facilitating transfer and how to adapt teaching methodologies and materials to accommodate the needs of this population.

Chapter 2: Literature Review

Several areas of research complement this study. The purpose of this review is to highlight the major theoretical and practical implications as they bear on the study. First, bilingual education programs will be reviewed to provide an understanding of the current role and use of the native language in classrooms. Next, major theoretical views on cross-linguistic transfer will be discussed, followed by a review of reading comprehension processes - with an emphasis placed on the interactive nature of reading. Major components of reading comprehension instruction will also be discussed as well as implications of cross-linguistic input for L2 reading comprehension instruction. Related to reading comprehension instruction is reading comprehension assessment. In particular, the review focuses on informal measures of assessment and the interaction between test type and reading comprehension performance as well as the implications for using think-aloud procedures on performance.

Bilingual Education

It is not possible to draw generalizations about classroom practices in bilingual programs because they vary between and within countries. In countries such as United States, Belgium, Bolivia, Canada, Nigeria, and Sweden, various types of bilingual programs have been implemented (Paulston, 1988). These programs vary in a number of factors including the ratio of L1 to L2 use, distribution schedules, functions allocated to the two languages, and the nature of the curriculum materials (see Baker & Jones, 1998, for a detailed description of bilingual programs).

In the New York City school system, the Department of Education (NYCDOE) asks parents to complete a Home Language Identification Survey when enrolling their child to determine the child's home language. When a language other than English is spoken at home, the child is administered the Language Assessment Battery-Revised (LAB-R) test to establish English proficiency. Children who score below proficiency on the LAB-R become eligible for state-mandated services for English Language Learners (ELLs). Based on parental choice and program availability, ELLs are then placed in a Transitional Bilingual Education (TBE), Dual Language (DL), or English as a Second Language (ESL) program. Continuing eligibility for ELL services for the following school year is determined by the student's performance on the New York State English as a Second Language Achievement Test (NYSESLAT), which is administered each spring to all K through 12th-grade ELLs. In 2008-2009, 70.7% of ELLs in New York City were enrolled in ESL programs; 19.3% in TBE programs; and 3.7% in DL programs (NYCDOE, 2009).

TBE programs provide academic instruction in both English and the native language. It includes ESL, English Language Arts (ELA), and Native Language Arts (NLA) components designed to develop literacy and language skills in English and the native language including reading, writing, listening, and oral communication, in addition to content area instruction in both languages to teach mathematics, science, and social studies. Enrichment classes such as Art, Music, Computer Lab, and Physical Education are taught in English. The ratio of instruction in the native language to instruction in the second language is established according to the Language Allocation Policy (LAP) set by the NYCDOE, and it is determined by the students' level of English proficiency. In the

beginning stages, the ratio of native-language (NL) instruction to English-language (EL) instruction is 60% (NL) – 40% (EL); in the intermediate level is 50% (NL) – 50% (EL); and in the advanced level is 25% (NL) – 75% (EL). The objective of this program is to have ELLs become literate and fluent in English within three years or less, after which the student transitions to an all-English class.

In DL programs, students of one home language and monolingual English students are grouped together and taught in both languages. Academic instruction is 50% in English and 50% in another language (e.g., Spanish and English or Chinese and English). The goal of this program is to develop and maintain literacy and language skills in the two languages. This is the only program that promotes *additive* bilingualism by building on what the student brings to the classroom.

In ESL programs all language arts and subject matter instruction is provided in English. Support in the native language may be available through pull-out programs, but these are limited. In pull-out programs, children are temporarily removed from their regular classroom for up to 60 minutes while they receive ESL instruction. This program, however, is the least recommended, although it is not clear why. Instead, the two recommended programs are self-contained and push-in (NYCDOE, 2008, p. 24). In self-contained ESL classes, all academic instruction during the school day is in English, but the core content areas are taught using ESL methodologies and materials stipulated by the New York State Education Department Office of Bilingual Education and Foreign Language Studies ([NYSED OBE-FLS]; 2009). For instance, the NYSED OBE-FLS provides ESL teachers with written specification of learning standards and performance indicators for students at predetermined levels of English proficiency, as well as sample

classroom tasks and instructional modifications, materials and resources, and assessment tools and techniques that are adequate to use at each level. In push-in ESL classes, academic instruction during the school day is in English, but the regular teacher works in conjunction with an ESL teacher so that instruction can be integrated into the regular classroom through the content areas using ESL standards-based methodologies.

Whatever the program, the objective is for ELLs to become fluent and literate in English (see NYCDOE, 2008, for a complete description of the programs offered in NYC public schools).

Educational services for ELLs are provided in accordance to Part 154 of the Regulations of the New York State Commissioner of Education (NYCDOE, 2008). CR Part 154 requires that schools provide bilingual education programs in grades K-8 when there are 15 or more ELLs of the same language in two contiguous grades, and in grades 9-12 when there are 20 or more ELLs in any single grade. Freestanding ESL programs are implemented in schools that have fewer than 20 students with limited English proficiency at the same grade level and who speak the same native language. Currently, programs are available in seven languages including Spanish, Haitian-Creole, Chinese, Russian, Yiddish, French, and Korean.

Research on the nature of language use in the classroom has shown that when languages are typologically similar, as it is the case between English and Spanish, L1 may contribute to the acquisition of L2. In their research review, the National Literacy Panel on Language-Minority Children and Youth (August & Shanahan, 2006) found no indication that bilingual instruction impedes academic achievement in either the L1 or the L2 just as predicted by Cummins' (1979, 2000) linguistic interdependence hypothesis,

and that language-minority students instructed in both L1 and L2 tend to perform better on L2 reading measures than students who are instructed only in L2.

In the United States, views differ regarding the value of the use of the first language for instruction. Proponents of programs such as TBE claim that teaching academic subjects through the stronger language will maximize the students' attention, will allow for increased time spent on conceptual learning, and will prevent teachers from having to teach both English and concepts at the same time (Dulay & Burt, 1980).

Opponents such as Rossell and Baker (as cited in Cummins, 2003), on the other hand, argue that bilingual education will promote social fragmentation in addition to being ineffective. The use of bilingual programs is a topic of increasing controversy in spite of the evidence that such programs do help children develop literacy skills (see Greene, 1997; Lucas & Katz, 1994; Slavin & Cheung, 2004, for meta-analyses comparing bilingual and ESL programs). The present study sought to add to this body of research by investigating the potential benefits of bilingual instruction on reading comprehension among ESL adult students.

Hypotheses About Cross-Linguistic Transfer

In their synthesis of research on the development of language and literacy skills of language-minority children, the National Literacy Panel on Language-Minority Children and Youth (August & Shanahan, 2006) found that many factors that are known to influence the development of L1 literacy skills among native-English speakers such as phonological processing skills, letter knowledge, word reading skill, and vocabulary knowledge, also influenced the development of L2 literacy among non-native-English-speaking children. But in addition, other factors unique to L2 learners exerted an

additional influence. Some the factors were: the age at which L2 skills were acquired, the degree of L1 and L2 oral proficiency and literacy, the extent to which the L1 and L2 orthography differed, and initial language of instruction. These findings demonstrate the complexities involved when studying specific literacy outcomes for L2-learners beyond those already inherent in the learning process of any given literacy skill. Research among older children and adolescents has shown that older learners who have had several years of uninterrupted schooling in L1, and who are proficient in L1, acquire L2 school language most efficiently, except for pronunciation (Collier, 1987, 1989).

In 1984, Alderson put forth the question, “*Is reading in a non-native language a reading problem or a language problem?*” This question has prompted much research on the transfer of reading skills from a native language to the learning of a foreign or second language. Alderson defined reading as a complex activity involving an interaction between a reader and a text. He suggested that poor reading in L2 is a function of either inadequate L2 proficiency or poor L1 reading ability, or both, whereas good reading in L2 is a result of good L1 reading ability or good L2 competence, or both. Nowadays, L2 reading is rarely viewed as strictly a language or a reading problem. Instead, many researchers agree that both L2-specific factors (e.g., lexical knowledge) and non-L2-specific factors (e.g., metalinguistic awareness) contribute to L2 reading; however, the relative importance of each is not clear.

Closely related to Alderson’s (1984) hypotheses are Cummins’ (1979, 2000) linguistic interdependence hypothesis and linguistic threshold hypothesis. According to Cummins’ linguistic interdependence hypothesis, the development of competence in L2 is partly a function of competence already developed in L1 at the time when intensive

exposure to L2 begins; that is, knowledge acquired in L1 can positively influence the development of skills in L2. In an attempt to account for the relationship between language proficiency and academic achievement, Cummins (2010) suggested that not all aspects of language proficiency could be incorporated into one global dimension of language proficiency. Instead, language proficiency may be defined along two intersecting continua based on the cognitive demands and contextual support inherent in tasks and activities. Along one dimension, tasks may be defined as cognitively demanding or undemanding, and along the other intersecting dimension, the contexts in which tasks are presented may be embedded or reduced. Hence, to the extent that an individual is able to negotiate meaning and learning will depend upon the cognitive demands of the task and the support provided.

In context-embedded activities, communication and comprehension are supported by external cues, whereas during context-reduced activities there is little support from external sources and successful negotiation of meaning will be highly dependent upon knowledge of the language and skills the individual possesses. The same distinction can be made along the cognitive dimension. Activities that are automatized require little cognitive effort, whereas activities that are not automatized require more cognitive effort. Two major constructs linked to these assumptions are BICS and CALP. On the one hand, BICS refers to *basic interpersonal communicative* skills and was intended to explain conversational aspects of language proficiency (i.e., skills used in context-embedded, cognitively-undemanding activities). CALP, on the other hand, refers to *cognitive academic language proficiency* skills and was intended to explain academic

aspects of language proficiency (i.e., skills used in context-reduced, cognitively-demanding activities).

Thus, he draws a distinction between BICS, which are language skills involved in the acquisition of conversational language needed for every day face-to-face social interactions (e.g., conversational fluency), and CALP, which are literacy-related language skills, on the premise that they follow different developmental patterns. In fact, BICS may only take a few years to develop whereas CALP continues to develop throughout many years of schooling. Research has shown that when second-language students are exposed to the L2 in an L2 context, BICS may take one to two years before word-level skills start catching-up to native-speakers' developmental levels, and at least four to five years before text-level skills start to catch-up, although the latter may take as much as 12 years (e.g., Collier, 1987, 1989; Cummins, 1980, 1981; Hakuta, Butler, & Witt, 2000).

Among bilingual children, CALP refers to those aspects of language proficiency that are not necessarily language specific and that are closely related to the development of literacy and cognitive skills in L1 and L2. Hence, L1 and L2 CALP are developmentally linked to a *common underlying proficiency* which makes possible the transfer of skills from one language to another. Cummins (1979) suggested that certain aspects learned in L1 such as reasoning and problem-solving skills, strategy use, and other higher-order cognitive processes do not need to be re-learned in L2 despite the obvious differences there may be between the L1 and L2 languages in terms of surface features. For instance, when students acquire reading skills in their native language, they are not just learning how to read strategically in the L1, but they can also use this knowledge when reading in a foreign or second language. Hence, with respect to

reading, we should expect that “the ability to extract meaning from printed text can be transferred easily from one language to another” (p. 234), as long as experiences in L1 have afforded the prerequisites for acquiring fluent reading skills and/or the development of strategic reading skills. Furthermore, we should also expect L2 CALP skills to develop more rapidly and L1 skills to transfer more easily among older L2 learners with better developed L1 CALP compared to children with undeveloped L1 CALP.

Research evidence for cross-language interdependence comes from studies that show positive and significant relationships among L1 and L2 analogous dimensions of cognitive and academic proficiency, and from studies showing transfer of knowledge and skills across languages (e.g., operational and conceptual knowledge/schemata and reading skills). When significant relationships between L1 and L2 skills are detected, it is inferred that these skills are part of an underlying cognitive apparatus, and are hence, interdependent. Various studies with Hispanic children in the U.S. have shown interdependence on standardized measures of academic language proficiency such as between the Spanish Comprehensive Test of Basic Skills (C.T.B.S) and the English Maculaitis Assessment Program - MAC (Ramirez, 1987), as well as between the Peabody Picture Vocabulary Test (PPVT) in Spanish and English (Hakuta & Diaz, 1985).

Transfer effects have also been reported for reading skills, awareness of cognate vocabulary, strategy use, and reading comprehension ability in studies with Spanish-English bilinguals. Goldman, Reyes, and Varnhagen (1984), for instance, investigated whether there was a relationship between L1 and L2 reading comprehension of fables among children Grades 4 to 6. They found that reading comprehension ability in both L1 and L2, as measured by a recall task, increased from Grade 4 to Grades 5 and 6.

Working with children in Grades 7 and 8, Nagy, McClure, and Mir (1997) found that L1 syntactic knowledge influenced guesses the students made about the meanings of unfamiliar words they encountered in L2. Hancin-Bhatt and Nagy (1994) found that students across Grades 4 through 6 successfully used their L1 morphological knowledge in L2 cognate recognition. In addition, Jiménez, García, and Pearson (1996) found that successful L2 readers in Grades 6 and 7 used their knowledge of L1 cognate vocabulary to enhance their comprehension of L2 texts. Carlisle, Beeman, Davis, and Spharim (1999) found that vocabulary knowledge in L2 contributed significantly to reading comprehension in L2, but in addition, L1 vocabulary knowledge also contributed significantly to reading comprehension in L2 among first- and second-graders. In a longitudinal study, Reese et al. (2000) found that performance on L1 measures of early literacy in kindergarten predicted reading performance in middle school. Specifically, children with higher levels of L1 proficiency in kindergarten, not only transitioned to an all-English class sooner, but also were better L2 readers in seventh-grade.

These studies provide ample evidence that some L1 and L2 literacy skills and sub-skills are related in important ways, especially when the L1 and L2 languages are typologically close such as Spanish-English. However, in an attempt to account for conflicting findings involving language proficiency in two languages and cognitive development, Cummins (1979) also suggested that there *may be* threshold levels of linguistic competence in L2 that must be attained in order for the L1 to positively influence cognitive and academic functioning in L2. In other words, it seems that to the extent that an individual with high L1 proficiency will be able to transfer literacy-related language skills and knowledge from the native language to the target language will be

hampered by L2 proficiency when L2 proficiency is below a threshold. In regards to cognitive functioning and development, when proficiency in both languages is at or below a *lower threshold* level, academic skills in both languages is low, which may have negative effects on cognitive development. In contrast, when proficiency is above the lower threshold, academic skills in both languages may be unbalanced but the level of bilingualism attained will not have negative effects on cognitive development.

Similar to Cummins' notion of a threshold, Clarke's (1979, 1980) short-circuit hypothesis asserts that some reading skills transfer from L1 to L2, but that this may vary as a function of the student's proficiency in L2. More specifically, he suggests that although some reading behaviors may be universal in that readers may approach texts in the same manner whether it is written in L1 or L2, there may be a *language competence ceiling* that will prevent good L1 readers from using effective reading behaviors when reading in L2. In that sense, low L2 proficiency creates a *linguistic ceiling* that "short-circuits" the individual's ability to transfer effective comprehension processes (e.g., use of prior knowledge and reading strategies) that are used during L1 reading to interpret the L2 text. Instead, the individual's "limited control over the second language will short-circuit the good reader's system, causing him/her to revert to poor reader strategies when confronted with a difficult or confusing task" (1980, p. 206).

The linguistic interdependence and threshold hypotheses are associated. When significant relationships between L1 and L2 cognitive/academic language proficiency skills are detected, it is inferred that these skills are interdependent. In addition, the fact that these skills are more strongly related at higher levels of L2 proficiency than at lower levels of L2 proficiency is interpreted as support for the threshold hypothesis. Also

research showing a failure to transfer conceptual knowledge and skills across languages at lower levels of L2 proficiency is interpreted as supporting this view. And so, to answer Alderson's (1984) question, it seems that poor *reading in a non-native language* may be due to both *a reading problem* and *a language problem* -- we can conclude it is a reading problem if proficient L1 readers who also have high L2 proficiency are poor L2 readers; and it well may be a language problem when proficient L1 readers are poor L2 readers and also have low L2 proficiency overall.

Lee and Schallert (1997), for instance, investigated whether insufficient L2 oral proficiency (i.e., lexical and grammatical knowledge) prevented students from effectively employing processes and strategies associated with accurate L1 reading comprehension. The students were 809 Korean children in 9th- and 10th-grade from across 40 schools in Korea, and who were learning English as a foreign language. The students read two narrative passages and two expository passages in both L1 and L2. For each passage, students answered five multiple-choice questions. They found that students with higher L2 proficiency performed better on the L2 reading comprehension tests. In addition, both L2 proficiency and L1 reading ability provided unique variance for predicting L2 reading scores, although L2 proficiency was a more significant predictor. In contrast, among students with lower L2 proficiency, correlations between L1 and L2 reading scores were low. The authors suggested that this may indicate that transfer of L1 strategic reading skills to processing L2 texts was short-circuited.

Clarke (1979, 1980) also investigated whether proficient L1 readers transferred reading skills to L2. The sample consisted of 21 Latino "low-level" ESL students. The students completed four cloze tests, two in each language. Overall, the L1 tests contained

1,903 words and 90 blanks, whereas the L2 tests had 602 words and 57 blanks. Based on their L1 cloze-test performance, students were classified as “good” or “poor” L1 readers. The average score was 60 and 45, respectively. Competence in L2 did not differ. Despite a 10-point overlap in the range of L2 scores between good and poor L1 readers, there was a significant positive correlation between L1 and L2 test performance when cloze tests were scored using the “exact” method. That is, good L1 readers performed better than poor L1 readers at providing the exact word replacement in both L1 and L2. However, when responses were analyzed for syntactic and semantic acceptability, L2 performance was equally poor between good and poor L1 readers.

Both studies provide some support for the threshold hypothesis, but they allow no direct conclusions as to the combined influence of L1 and L2 proficiency on L2 reading comprehension. In both studies, the authors assumed the students were proficient L1, but did not systematically test whether they were proficient in L1 or they were just good test-takers. Furthermore, although some students performed better than others on L2 tests, their overall level of L2 reading competence was equivalent. It was expected that the present study would provide further insight into the threshold hypothesis by exploring its existence within an ESL context with a group of young adults who were fluent and equally proficient in L1 but had distinctly different levels of L2 proficiency.

Although there are some obvious differences between the Spanish and English language systems, there are also many similarities between them. Therefore, in the present study, it was expected that transfer effects would be likely, especially among individuals who were proficient in L1 since they may be either more aware of these similarities or more likely to notice them when appropriate L1 input was offered, and use

these to their benefit. In addition, cross-linguistic transfer of knowledge and skills was expected to be particularly salient among students with higher L2 proficiency since their proficiency in L2 was expected to be above the lower threshold.

It is important to note that the theoretical framework for the linguistic threshold hypothesis remains incomplete. Cummins (2000) admits that it is not yet entirely clear whether specific thresholds are associated with specific cognitive outcomes. Furthermore, it has been difficult for the research community as a whole to provide a common operational definition of *language proficiency*; hence, there are no specifiable, pre-determined levels of proficiency matched to threshold levels. We can only speculate as to its approximation based on students' abilities at specific L2 proficiency levels on specific tasks. In the present study, cognitive academic language proficiency (CALP) levels were defined and measured with the Woodcock-Muñoz Language Survey-Revised ([WMLS-R]; Woodcock et al., 2005). Based on the definitions of each CALP dimension, CALP Level 2, which is defined as *very limited*, was taken as representing students within the *lower threshold*. According to Cummins (2000), a student's ability to understand and interact with an L2 text at the lower proficiency level is "restricted." Likewise, WMLS-R defines CALP 2 as an individual with very limited CALP who will find the language demands of the learning task extremely difficult. Students in the CALP 3 dimension, which is defined as *limited*, were taken as representing "higher" proficiency. Although these students are still expected to find the language demands of the learning task difficult to accomplish, they are significantly more proficient in L2 than students at CALP Level 2. Most college-level students enrolled in ESL classes have CALP 2 and 3 proficiency levels in L2.

Reading Comprehension Processes

Reading comprehension is not a single process, but rather a complex combination of cognitive and linguistic processes that entail more than the mere decoding or recognition of words. According to the National Reading Panel, comprehension involves “constructing meaning that is reasonable and accurate by connecting what has been read to what the reader already knows, and thinking about all of this information until it is understood” (NICHD, 2000, p. 30). Various reading models have been proposed, but most reading models emphasize the interaction of lower- and higher-level processes as the reader attempts to make sense of the text.

For instance, Rumelhart’s (1977) interactive model of reading provides an account of the various knowledge sources a reader uses to recognize and interpret text. Adapting Rumelhart’s model, Ehri (1998) provides a clear description of the various bottom-up and top-down processes that take place throughout the reading process. She identifies several knowledge sources including: orthographic/graphophonic, lexical, syntactic, semantic, pragmatic/contextual, schema/background knowledge, and metacognitive, each working in parallel as the most probable interpretation of the text is reached. For example, reader’s knowledge of the graphophonic system as well as their lexical knowledge allows them to decode unknown words and to read words by sight, but their perception of the words depends on the syntactic and semantic environment in which the words appear. Furthermore, the general context in which the text appears also impacts the interpretation and meaning the reader constructs. Hence, several knowledge sources place constraints on each other.

Stanovich (1980, 1984) also argues for an interactive model; however, he adds a

compensatory mechanism to account for individual differences in reading, particularly as a function of context effects on word recognition (i.e., reading fluency). He suggests that when processing at one level is deficient, information processing at another level will compensate irrespective of the level where the deficiency occurs. However, this compensation comes at the expense of the already-limited cognitive capacity available for the comprehension process, which may hinder comprehension further. One implication of this mechanism is that good L1 readers with poor L2 decoding skills may rely on contextual information when the context is adequately understood to facilitate word recognition.

One of the most influential models of text comprehension is Kintsch's construction-integration model of text comprehension (1988, 2004, 2005). Kintsch defines reading comprehension not as understanding per se, nor as conscious, deliberate problem-solving, but rather as a combination of both processes. In that sense, comprehension is the automatic construction of meaning via constraint satisfaction. The constraints include the perceptions about the input received that occurs without conscious effort, as well as conscious problem solving and planning (i.e., the purposeful use of cognitive strategies) whenever a conflict arises. In his construction-integration model, Kintsch describes text comprehension as involving processing at different levels (i.e., levels of text representation).

The first level of representation is the *surface-level* and refers to memory/processing of the actual words and phrases in the text. The second level is the *semantic-level* and refers to representations of the ideas expressed in the text. At this level of processing, word meanings are combined according to their connection strengths

to form idea units or propositions that are interrelated in a complex network called *microstructure* (i.e., the network of propositions that represents the meaning of the text). That is, once a given word is identified, lexical nodes and their associates are activated, forming propositions. These propositions, which may or may not be accurate, are directly related to the printed input and trigger the retrieval of associated knowledge (i.e., prior knowledge) stored in long-term memory, allowing for the elaboration and refinement of the original propositions and/or further activation of closely related knowledge propositions. These propositions may be related referentially (i.e., coreference/argument overlap in which two or more propositions refer to same concept), causally, or logically. Furthermore, these ideas and interrelationships become organized into higher-order units corresponding to the global topics in the text and their interrelationships and lead to the formation of the *macrostructure* or mental representation of the text. The micro- and macro-structure together form the *textbase*, which is the semantic underpinning (i.e., explicit meaning) of the text.

According to this model, even when the reader has formed an adequate textbase, comprehension is shallow. In order to reach a deeper understanding, the reader must construct a *situation model*. This level of processing represents the integration of information provided in the text with information from the reader's prior knowledge. The situation model is restricted not only by the individual's prior knowledge regarding the topic, but also by their interests, beliefs, and their particular goals for reading the text. In a nutshell, text processing progresses hierarchically through these levels of processing by first understanding word meanings, connecting these to form propositions, connecting larger parts of the text, integrating the parts to form a comprehensive and meaningful

whole, and finally, integrating the information with that in our present schemata through accommodation and assimilation, all of which is aided by inferences made at each level.

An implication that follows from these models is that for learners with low levels of L2 competence, reading comprehension in L2 will be hampered by a deficiency of lower-level skills, but readers can compensate by calling on higher-level processes such as activating prior knowledge. According to Cummins' (1979, 2000) and Clarke's (1980) hypotheses, however, we should expect low-proficiency L2 learners to have very limited access to such prior knowledge and other higher-order reading skills (i.e., strategies) acquired during the development of L1 literacy-related and language skills until they have attained a certain level of proficiency in L2. In the meantime, the characteristic interactive nature of the reading process in good L1 reading will change to mostly bottom-up processing in L2. This in turn may lead to the formation of an incomplete textbase and an inadequate situation model. As a consequence, the meaning of the text will not be fully understood.

Taking these propositions into account, a question of interest in the present study was whether exposure to print in one language could provide the necessary scaffold to facilitate reading performance in another language. It is possible that concurrent exposure to print in the native- and second-language may facilitate connections among vocabulary and content prior knowledge, which will in turn aid making inferences aimed at constructing a mental representation of the text irrespective of the level of L2 proficiency.

Reading Comprehension Instruction

The ultimate goal of reading is comprehension. In the 1970's, Markman found that both young and mature readers failed to monitor their comprehension and did not recognize when they did not understand what they read (as cited by NICHD, 2000). Since then, an increasing amount of research has been dedicated to the identification of strategies that students can be taught to enhance reading comprehension (NICHD, p 4-39). A search of the literature on reading comprehension instruction revealed that much of the research has been devoted to the teaching of cognitive strategies alone or in combination including: comprehension monitoring, reciprocal teaching, cooperative learning, graphic organizers, question answering, question generation, and summarization. In general, findings indicate that reading comprehension can be improved by teaching students how to interact with text strategically.

In its research review regarding the development of reading comprehension skill and comprehension instruction, the National Reading Panel (NICHD, 2000) found that when students were taught to use cognitive strategies successfully, they performed better on reading comprehension tasks including recall, answering questions, generating questions, and summarizing texts. Explicit cognitive strategy instruction has been successful across various grade levels and ability levels, as well as with English language learners. However, not only is reading comprehension a complex process that requires an intentional and thoughtful interaction between the reader and the text, but also vocabulary plays a crucial role. Hence, comprehending the meaning of texts is often difficult for language-minority students because they often do not have the extensive vocabulary that native speakers have.

For instance, Hadley (1993) found that the vocabulary of adults at the beginning stages of learning English only consisted of about 2,000 to 7,000 words compared to 10,000 to 100,000 words among fluent English speakers. Various vocabulary aids can be used for teaching vocabulary, including inserting glosses in the text, and using dictionaries and glossaries before, during or after reading the text. Other vocabulary instruction methods consist of teaching students how to analyze and build word parts such as roots, prefixes and suffixes, and how to use the word origin clues to determine vocabulary definitions. Another common method is to teach students how to derive word meanings by using context clues in the text (see Nation, 2005, for a discussion regarding teaching and learning vocabulary).

There is no doubt that this gap in vocabulary, among other skills, is detrimental to academic success, and the importance of the essential components of effective reading instruction identified by the National Reading Panel (NRP), which include vocabulary, have not gone unnoticed (NICHD, 2000). The New York City Department of Education uses a Balanced Literacy approach to teach reading that stresses each of the essential components identified by the NRP. In kindergarten, instruction is targeted at the development of phonological and phonemic awareness, print awareness, letter knowledge and phonics, fluency, and comprehension strategies. In first grade, background knowledge and vocabulary development through direct means are added to the repertoire and continue until middle school. Much less is known about reading instruction in the upper grades and beyond, and the potential benefit that native-language use may have on the development of literacy-related skills in L2 among older learners. In the review of the literature, only a handful of studies were identified that specifically investigated how

the use of the students' native language helped older students understand L2 input. The present study addressed this need.

Cross-Linguistic Input and Reading Comprehension

Research on cross-linguistic influences on text-level skills is scant, especially so for investigating reading comprehension among language-minority adults. However, the available research suggests that native-language use in reading materials may serve as an aid by providing the reader with background knowledge of various types (Alderson, 2000), including formal schemata (i.e., how texts are organized in terms of structural and linguistic features), and content schemata including knowledge of the world (i.e., knowledge which may or may not be relevant to the content of the text), subject-matter knowledge (i.e., knowledge directly relevant to the content of the text), and cultural knowledge (i.e., influences posed by personal history and unique experiences).

Friesen and Jared (2007) investigated whether exposure to a text in one language would facilitate reading performance in another language. Specifically, they examined whether vocabulary encountered in a text written in L1 would facilitate subsequent reading of the same words in a text written in L2 as measured by reading times and eye gazes and fixations. Undergraduate students who were native speakers of English (L1) and were learning French (L2) as a second language classified as more- and less-skilled bilinguals read five pairs of passages and answered five multiple-choice questions. The passages were combined in multiple ways so that the students would read either an identical text twice in the same language (i.e., L1-L1 or L2-L2), the same story translated in two different languages (i.e., L1-L2 or L2-L1), or two unrelated stories in different languages.

The authors expected that if there were transfer effects, the second text should be read as quickly when preceded by a translation or an identical text, although the transfer effects may be greater for identical passages, and the second text should be read slower when preceded by an unrelated text. The results showed that more-skilled bilinguals read the texts faster than less-skilled bilinguals and that L1 passages were read faster than L2 passages. In addition, passages were read faster when they were preceded by either an identical passage or by a text translation than when they were preceded by an unrelated passage. Lastly, it was found that the effects were greater in the L1-L2 condition than in the L2-L1 condition. The authors interpreted the results as suggesting that students transferred knowledge from L1 to reading L2 translations of the texts; that mental representations of a text are not language-specific, and that meaning can transfer independently of surface form (i.e., language of text). Performance on the comprehension questions was not discussed. Instead, scores on the multiple-choice questions were used to classify students as more- and less-skilled bilinguals.

The study above is informative about the potential benefits of L1 input to L2 reading in terms of transfer effects of mental representations of the meaning of a text from L1 to L2; however, it fails to answer the primary question of interest in this study of whether ELLs transfer knowledge acquired from reading a text in one language to the reading of the same text in another language to aid reading comprehension as opposed to reading ability based on decoding speed. This is not a novel idea. In 1949, Sparkman argued adamantly that it is unreasonable to ignore the student's native language when teaching a new language, at least during or after adolescence. He explained,

The labels of meaning are purely arbitrary to begin with. *Cat* might just as well have been *tac*. In fact a completely different set of labels may be

used and taught to children just across an international boundary line; but meaning is essentially the same anywhere in the world. Since it was labeled in the beginning [while learning one's native language], why can't a new label be pasted over the old one? This is what really happens while learning a second language during or after adolescence. The concepts already exist – they are the tools of thought ... – and each is tagged with its proper native-language word label. The fundamental problem is to retag the learner's concepts. (p. 356)

Bernard (1950) suggests that, in general, there are two possible ways to acquire the meanings of a new symbol (i.e., a foreign word). One way is by *directly* associating the symbol with the object. Another way is by *indirectly* associating the symbol with an old symbol for which the learner already has assigned meaning (i.e., the native word). Through this latter method, he argued, is the way most language is learned because “whether we like it or not, whether we do it voluntarily or involuntarily, consciously or unconsciously, *we must at the initial stage of learning at first always translate into our own language* before the new symbol acquires meaning” (p. 92).

Acquiring meaning is, therefore, making an association to something that symbolizes that meaning. Once the meaning is obtained, the indirect link will no longer be needed. If there is frequent association of the foreign symbol and its meaning, subsequently, the learner will be able to bind the direct link of the foreign symbol to its meaning. Based on this idea, Bernard (1950) proposed the parallel-translation or *bilingual reading method* as the method to be followed to enable learners of a second language make appropriate use of the L1. It should be noted that for this method to be effectively applied, the learner must have a good foundation in the L1.

Parent and Belasco (1970) were among the first to test Bernard's claim. They examined the effectiveness of the parallel-column bilingual method as a *pedagogical*

device for developing reading comprehension. The sample consisted of 127 undergraduate students enrolled in an intermediate French course. The Modern Language Association Cooperative Classroom French Reading Test (MLA) was administered to determine L2 reading ability. The students were then ranked based on their performance and randomly assigned to one of three experimental groups.

In the first experimental condition, the students were exposed to all-French reading materials. In the second experimental condition, the students used parallel-column French-English reading materials. In the third experimental condition, the students used parallel-column French-English reading materials in addition to French audio-tapes of the reading materials. Once a week, the students were given 60 minutes to read the materials silently and 15 minutes to complete 10 multiple-choice questions in French. The readings were excerpts from a novel the students were reading in class. During the last week of the experiment, the students completed an attitude inventory and an alternate form of the MLA test. The intervention lasted a period of 10 weeks.

The results revealed that both the silent bilingual reading group and the bilingual reading group who also heard the French tapes outperformed the all-French reading group on the weekly tests. The difference in performance between the silent bilingual reading group and the bilingual reading group plus French audio was not significant. Overall, there were no significant differences in the performance of the groups on the MLA post-test.

Additional analyses were performed according to reading ability. Students scoring above 1 *SD* on the MLA pre-test were classified as high-ability ($n = 23$), and students scoring below 1 *SD* were classified as low-ability ($n = 24$). This time, there

were no significant differences in the performance of the experimental groups among high-ability students on the weekly tests. Among low-ability students, however, the silent bilingual reading group and the bilingual reading group who also heard the French tapes outperformed the all-French reading group on the weekly tests.

The results suggested that the use of L1 as a pedagogical device in the form of parallel-bilingual readings is effective for increasing L2 comprehension. Although, the performance of the silent bilingual reading group was similar to that of the bilingual reading group plus French audio, the authors pointed out that approximately 40% of the students who had access to the French audio did not use it. Therefore, it is reasonable to conclude that the advantage found for the bilingual reading group who also heard the French tapes was more a function of the bilingual materials and not because of access to the French audio. In regard to the non-significant difference among the high-ability students, we can speculate that their L2 knowledge had reached a threshold for which exposure to L1 no longer provided an added advantage.

Although the authors made an effort to utilize one of the most powerful designs, namely, a randomized control group design, some procedural elements were unclear. For instance, with the exception that it was a novel, we know nothing about the nature of the text such as its length and readability in order to assess whether one hour was an appropriate length of time for the students to read the texts each week. We can also wonder whether the same results would be found if we were to use other reading materials with different structures. In addition, the students read the texts silently so we know little of the processes involved when reading the bilingual materials in order to try

to understand *why* these were helpful. These questions deserve further exploration and were addressed in the present study.

Jarvis and Jensen (1982) also examined the extent to which parallel translations would facilitate reading comprehension in another language. The sample consisted of 57 undergraduate students: 39 students in the first or second semester of Russian (i.e., beginning level) and 18 students in the third semester and beyond (i.e., advanced level). The students read eight passages. The passages were dialogs from a film script each about 100 words in length. Four passages were “easy” and four passages were “difficult.” At each difficulty level, the passage was accompanied by one of four reading aids: occasional side glosses, standard parallel translations, word-for-word parallel translations, or standard plus word-for-word parallel translations. Standard parallel translations consisted of commercially-available translations of the film and word-for-word parallel translations consisted of word-for-word translations written by the researchers (D. Jarvis, personal communication, November 11, 2008). No explanation was given as to how specifically these two differed in content.

The order in which the passages were presented was randomized by a computer. All the students read all the passages. The students had five minutes to read each passage and nine minutes to complete a seven-item multiple-choice cloze test on the content of the passage and a seven-item multiple-choice English-to-Russian translation test focusing on syntactic elements found in the passage. Lastly, they were asked which reading mode they preferred.

The results of a 4 (aid type) x 4 (task type) x 2 (proficiency level) ANOVA at the $p < .10$ level of significance showed that the translation modes resulted in higher scores

for both beginning- and advanced-level students across all four tasks, with the exception of advanced-level students on the cloze tasks for “easy” content. On the cloze tasks for easy content, advanced-level students performed better when occasional side glosses were available than when word-for-word parallel translations were available. In an exit interview, it was found that the majority of the students preferred the translation modes over occasional side glosses.

Although informative, the study had several limitations. First, the passages were very short and plenty of time was allotted to read them. Students could have read them more than once without controls in place. In addition, no information was given about how the passages were rated for difficulty level. Also, no descriptions were given about the students, except that they were learning Russian as a foreign language. It could be that their knowledge of the Russian language did not differ to a great extent and all the students were similarly proficient in the second language. Learning a language as a foreign language differs from learning a language as a second language in that in the latter case, the learner is usually exposed to the language in a variety of settings outside of the classroom, whereas in the former case the students are usually exposed to the language only in the classroom. Most pressing is the limitation that no detailed explanation of the conditions or procedures was given so we cannot know with certainty how the translation modes differed from the occasional side glosses. Another difference between this study and the others previously discussed in this review is that the researchers used dialogues from one film, not narratives or expositions, and it is not clear whether the students had been previously exposed to the film.

A modification to the parallel-column translation method was studied by Lasisi, Falodun, and Onyehalu (1988). They used a bilingual repeated-reading paradigm to investigate whether exposure to culturally unfamiliar stories in one language would facilitate reading in another language. The participants were 300 seventh-grade Nigerian students selected at random from five schools. The Slosson's Oral Reading Test (SORT) was administered to ensure that the children had a "reasonable level of reading ability" in English. Students who could not read the words on the SORT were eliminated from the study.

The students were randomly assigned to one of six conditions. In each condition, students read two culturally unfamiliar passages, one literal and one interpretative. Literal passages were described as simple plot-based prose passages, and interpretative passages were described as theme-based prose passages. The passages, which were originally written in English (L2), were translated into Yoruba (L1). In the first condition, students read the passages in L2 once, and answered 10 multiple-choice questions in L2: 5 literal and 5 interpretative. In the second condition, students read the passages in L1 once, and answered the multiple-choice questions in L1. In the third condition, students read the passages in L1 followed by the same passages in L2, and answered the multiple-choice questions in L2. In the fourth condition, students read the passages in L2 followed by the same passages in L1, and answered the multiple-choice questions in L1. In the fifth condition, students read the passages in L2 followed by the same passages in L1, and answered the multiple-choice questions in L2. In the last condition, students read the passages in L1 followed by the same passages in L2, and answered the multiple-choice questions in L1.

The analyses revealed only one statistically significant finding. Scores obtained by students who read the literal passages in L1 and answered the questions in L1 were significantly higher than the scores for students who read the literal passages in L2 and answered the questions in L2. This finding may not be surprising given that you would expect students to do better when reading materials in their first language and being tested in the first language compared to reading and testing in a second language for which they have “poor language arts teaching and supervision” (p. 35).

Of interest in this study is the finding that no significant differences were found on L2 performance when children read the passages in both languages, either in L1-L2 or L2-L1, compared to when they read the passage only in L2. However, a comparison of the mean performance of each experimental group revealed that when students were tested in L2, they tended to perform better in the L1-L2 condition than in the L2-L1 or L2 only conditions. In addition, the performance of students who read the passages in L2-L1 and were tested in L1 did not differ from the performance of those who read the passages in L1-L2 and were tested in L2 regardless of the type of text. These findings suggest that in addition to language use, language sequence might have an effect on L2 comprehension.

In sum, these results suggest that L1 input may facilitate L2 reading comprehension of culturally unfamiliar material and that when students read texts in two languages, reading the L1 version of the text first may be more beneficial than reading it second. However, just as with the study previously discussed, this study had a few limitations. For instance, L2 proficiency was based on a single measure of reading ability that required students to read a list of words, and no information was provided regarding

the students' L1 and L2 language abilities or the criteria used for inclusion in the study. In addition, no information was provided about the nature of the texts such as length and readability, nor was information given about how much time was allotted or how long the students took to read the texts.

Taken together, despite their limitations, the studies reviewed provide support for the hypothesis that L1 input aids L2 reading comprehension. As students encounter comprehensible texts in L1, they become familiar with its structure in addition to building background knowledge. Specifically, the mental representations that are formed while reading a text in L1 are not language-specific and become available when encountering the same text in L2 to aid comprehension. The results of the studies above suggest that the use of L1 as a pedagogical device in the form of translations during parallel-bilingual readings or during repeated-readings is effective for increasing L2 comprehension, and that the effects might be stronger for students with lower English proficiency whereas high-ability students may have reached a threshold for which exposure to L1 no longer provides an added advantage. With respect to reading, however, even when students have comprehended what they read, their ability to demonstrate what they have understood can be compromised by the level of linguistic knowledge and production required to complete various reading comprehension assessments. This topic will be discussed next.

Reading Comprehension Assessment

Assessing reading comprehension is not straight-forward because the processes used by readers to create meaning are invisible and the results from various assessment tasks provide only a partial picture of how much comprehension has taken place. Several

methods for measuring reading comprehension have been devised and there is disagreement in the field regarding what the best approach is for measuring reading comprehension. Bachman (1990) suggests that test type can positively or negatively affect test performance.

Multiple-choice tests are the most common. Bachman (1990) calls these *selected* response tasks because the test-taker is provided with retrieval cues that can facilitate recall rather than having to generate the recall. Another criticism is that this type of test only taps selected portions of the text (Fuchs, Fuchs, & Maxwell, 1988). An alternative to multiple-choice is a short-answer test in which the test-taker is asked to provide a brief response to a question. This task only requires a limited production response (Bachman & Palmer, 1996).

Free-recall tests, on the other hand, require an extended production response (Bachman & Palmer, 1996). Here, test-takers are asked to read a text and then write or say everything they can remember from the text. This task is often thought to provide a purer measure of comprehension because it does not intervene between the reader and the text. In addition, this type of task allows the test administrator to have a sense of how students process the text in terms of quantity, quality, and organization of information (Lipson & Wixson, 2009). Hayes (1989) considers the recall protocol analysis to be the most powerful tool for tracking psychological processes; and it has been advocated over traditional test methods such as multiple-choice, cloze tests, and open-ended questions because it is thought to “provide a purer measure of comprehension, uncomplicated by linguistic performance and tester interference” (Bernhardt, 1991, p. 200). An obvious problem encountered by ELLs is that students may understand the text but are unable to

completely express the ideas verbally or in writing due to limited linguistic skill or limited experience with this type of task. In addition, ability to recall information is also heavily influenced by memory (Fuchs, Fuchs, & Maxwell, 1988).

Cloze tests are useful as well as easy to prepare and score, but their validity is also controversial. These tests are usually constructed by deleting every *nth* word. A variation is a rational-deletion cloze, in which specific words that demonstrate certain knowledge of interest to the researcher are selected for deletion. What an individual cloze test measures depends on the words that are deleted (Alderson, 2000). For instance, by deleting mainly content words you may be testing overall understanding of the text. On the other hand, by deleting function words you may be testing grammatical knowledge. You can also provide multiple-choices, which may be inserted after the gap for the omitted word or you can show the list of omitted words at the end of the text. This latter method is known as matching cloze tasks. Critics of the cloze task argue that it measures textual redundancy instead of comprehension, but correlations among cloze scores and other measures of comprehension have been typically found to be over .80 (Fuchs, Fuchs, & Maxwell, 1988).

According to Bachman (1990), in *constructed* response tasks such as open-ended questions, cloze tasks, and recall protocols, the test-taker recalls the appropriate text information and produces the response. Because retrieval cues are not provided, recall becomes more difficult in these tasks. On the other hand, they may provide a more authentic measure of comprehension. Other assessment measures involve having students complete an ordering task in which they have to order a set of sentences or

paragraphs to reproduce the correct order of the text; answer true or false items, or complete editing tests in which the student is asked to correct errors in a passage.

Whatever the task, the fact is that students' ability to demonstrate comprehension may be affected by testing method, and students who are ELLs may be especially vulnerable to test type influences. For example, Shohamy (1984) found that testing method had a differential effect on reading comprehension when comparing performance on multiple-choice versus open-ended questions. In her study, 12th-grade Israeli students with low levels of English (L2) proficiency had difficulty completing both tasks, and their performance was generally very poor. In contrast, students with higher L2 proficiency were not affected as much as students with lower L2 proficiency. However, it is not always the case that advanced-level students are not affected by test type. Wolf (1993) examined whether college-level students enrolled in a fourth- and eight-semester Spanish (L2) course would perform differently on or across different assessment tasks. She found that both advanced- and less-advanced students performed significantly better on multiple-choice tests than on open-ended questions or a rational-deletion cloze test.

The available research suggests that all reading comprehension assessments are not necessarily equivalent; that measuring reading comprehension with the use of one method is inadequate; and that there is simply no one best method for assessing reading comprehension. Therefore, it is important to use multiple measures when testing reading comprehension. This caveat was taken into account in the present study. Three methods were used to assess reading comprehension including a free-recall, open-ended questions, and a cloze task. These measures required different degrees of comprehension,

production, and linguistic skill and provided a more complete assessment of comprehension than could have been obtained with one measure alone.

An Inside Look Into Reading Comprehension

In addition to assessing reading comprehension as a “product” through informal assessments, the present study analyzed the reading comprehension “processes” that took place as students read the text. Think-aloud (TA) procedures have been useful for understanding the underlying cognitive, reasoning, and affective processes that take place during performance of a task (e.g., while reading a text). TA studies have been critical for understanding the nature of good reading. Pressley and Afflerbach (1995) reviewed 38 TA studies published between 1964 and 1993, and found that the activities reported in verbal protocols could be classified as of three main types: constructing meaning, monitoring, and evaluating. They further divided the strategies to identify what “good readers” do before, during, and after reading. When analyzing the verbal protocols in the present study, reading activities were carefully analyzed in search of strategies based on information discussed by Pressley and Afflerbach (1995) happen during reading (see pp. 34-36, 43-44, 46-48, 51-53, 55-57, 62-73, for a list of strategies that occur during reading). Based on a complete analysis of the transcriptions, all the preliminary strategies identified could be classified within those used for interpreting information and monitoring comprehension (see pp. 55-57, 62-73). These were then cross-referenced with strategies representing conscious constructive responses to text (see p. 105).

After referring to Pressley and Afflerbach (1995), a search of the literature on the use of TA for identifying reading strategies while reading in L2 was conducted. While reviewing the studies, one of the main goals was to seek out commonalities to further

refine the labels for the strategies identified and their definitions. It was found that Hock Seng and Hashim's (2006) taxonomy was the most suitable for the research question in the present study. Of similar interest, they told students to use whatever language they wanted as they completed the TA, which allowed the authors to analyze the extent to which students used both L1 and L2 during reading. In this way, they were able to identify how L1 directly impacted performance in L2.

The authors identified a total of 18 strategies, which were categorized as "reader-based" or "text-based" in L1 or L2, according to the language used in the TA. Text-based strategies were thought to reflect resolutions at the word-level, whereas reader-based strategies reflected resolutions at the idea-level. The authors were able to successfully identify these processes. Therefore, strategies in the present study were defined and classified as text-based (i.e., bottom-up) or reader-based (i.e., top-down) in L1 and L2 following their taxonomy as appropriate. For example, not all strategies were found across, and some definitions were altered for a more stringent coding. At the end, 12 strategies were identified. These are discussed in detail in the Methods section.

Encoding non-verbal thoughts into verbal form, however, requires an additional task that critics of the TA methodology fear alters the normal reading processes. According to Ericsson and Simon (1993), ordinary thinking processes do not change much under certain TA instructions. They classified concurrent verbalizations (i.e., thoughts while performing a task) into three levels. Level 1 verbalization refers to reports of thoughts as required in a particular task (e.g., saying out-loud the steps taken to change an oil filter while performing the task). Level 2 verbalization refers to explanations of thought content (e.g., making overt a thought about difficulties you

encounter when changing oil filters as it occurs). Level 3 verbalization refers to interpretative explanations of thought processes across a time-span (e.g., explaining the steps taken to change an oil filter and the implications at each step). Ericsson and Simon found that TA instructions that asked subjects to verbalize their thoughts without requiring an explanation of their thought processes (i.e., verbalizations at Levels 1 and 2), only added time-on-task without changing “the course or structure of the thought processes” (p. 106).

Although the TA procedure has been used extensively, only a couple of studies to date have empirically tested the issue of reactivity of thinking aloud while reading in L2 on reading comprehension. Just a few years ago, Leow and Morgan-Short (2004) conducted the first study of this kind. The participants were 77 college students who were attending a first-year Spanish (L2) language program. Students read a magazine article (384 words) that contained 17 target words of the impersonal imperative linguistic form in L2 (i.e., *haga, ponga*). Some students were required to think-aloud and others not. They compared students’ performance on comprehension, intake, and production of the impersonal imperative linguistic form. The comprehension task required a mix of short and multiple-choice answers. The questions were presented in English (L1), and responses were required in both L1 and L2. No significant differences in performance were found between groups.

Bowles and Leow (2005) conducted a similar study, but with students enrolled in a fifth-semester Spanish (L2) course. Forty-five students were randomly assigned to one of two TA groups (metalinguistic or non-metalinguistic) or to a control group. Students read a magazine article (861 words) that contained 28 target words of the pluperfect

subjunctive syntactic form in L2 (i.e., *hubiera nacido, hubiera sido*). They compared students' performance on comprehension, intake, and production of the pluperfect subjunctive syntactic form, as well as on the amount of time taken to complete the task. Comprehension was assessed with a 10-item multiple-choice task with questions presented in English (L1). No significant differences in performance were found between the TA and control groups. However, the non-metalinguistic group (i.e., Level 2 verbalization) performed significantly better than the metalinguistic group (i.e., Level 3 verbalization). In addition, both TA groups spent significantly more time processing the text compared to the control group.

These studies indicate that when subjects are instructed to “simply verbalize” their thoughts as they read the text, thinking aloud appears to have no effect on reading comprehension as measured by tasks with a multiple-choice format. The present study differed from these two studies in important ways. In the present study, reading comprehension was assessed using three tasks that differed in format, none of which involved multiple-choice; each task was administered in L2 not in L1; and completion of the tasks was required in L2 not in L1. In addition, students in the present study were learning L2 in an L2 context and had higher levels of L2 proficiency. Hence, it was important to include a no think-aloud procedure control group in the present study.

A TA technique similar to Olshavsky's (1977) was followed for building the TA instructions and the data analysis of the verbal protocols. For example, the participants were told that the purpose of the study was to investigate “how readers understand short stories” as a whole and not to test the reading processes of him/her personally. Likewise, participants were informed that they would “see a red star at the end of the sentence to

remind” them to stop and talk-out their thoughts. Lastly, participants’ verbalizations were transcribed and matched with the corresponding sentence in the text in order to identify strategies.

Summary

Adult Spanish-speakers are the largest percentage of language-minority adults in the U.S. and the largest growing language-minority population. Developing reading skill in L2 has proven to be one of the most challenging literacy tasks for this population. According to Cummins’ (1979, 2000) linguistic interdependence hypothesis, the development of competence in L2 is partly a function of competence already developed in L1 at the time when intensive exposure to L2 begins. Furthermore, Cummins’ linguistic threshold hypothesis suggests that there may be threshold levels of linguistic competence in L2 that must be attained in order for the L1 to positively influence cognitive and academic functioning in L2.

Reading comprehension entails a complex interaction of lower- and higher-level processes and knowledge sources as the reader attempts to make sense of the text (Ehri, 1998; Rumelhart, 1977; Stanovich, 1980, 1984). In order to reach deep understanding, the reader must construct a situation model of text representation by integrating the information in the text with information from his/her prior knowledge (Kintsch, 1988, 2004, 2005). According to Cummins’ (1979, 2000) linguistic interdependence and threshold hypotheses, with respect to reading, we should expect second-language learners to be able to easily transfer their ability to extract meaning from printed text from L1 to L2. However, the individual’s ability to transfer literacy-related language skills and knowledge from the native language to the target language will be short-circuited if L2

proficiency is below a threshold. This in turn will lead to the formation of an inadequate situation model and, consequently, poor L2 reading performance.

Several research studies provide support for the hypothesis that L1 input may serve as a scaffold for increasing L2 reading comprehension, and that the effects might be stronger for students with lower L2 proficiency, whereas students with higher L2 proficiency may have reached a threshold for which exposure to L1 no longer provides an added advantage (Friesen & Jared, 2007; Jarvis & Jensen, 1982; Lasisi, Falodun, & Onyehalu, 1988; Parent & Belasco, 1970). The purpose of the present study was to investigate the effect of native-language input and second-language proficiency on L2 reading comprehension among language-minority adults whose native language was Spanish. The primary question of interest was whether adults who were learning English as a second language would be able to comprehend printed text more effectively if they read the text first in L1 and then in L2, than if they read the text twice only in L2. It was expected that reading a text in one's native language would provide a scaffold in the form of background knowledge that would allow readers to make connections between their first language and the new information in the second language, thus facilitating learning and performance in the second language.

Two concurrent outcomes were possible. If, on the one hand, students' ability to successfully transfer native-language knowledge and skills when reading the L2 text was dependent upon an L2 threshold of linguistic competence below which L1 could not positively impact cognitive functioning in L2 as proposed by Cummins (1979, 2000), it was expected that students with lower L2 proficiency would perform poorly across comprehension assessments even though they had good CALP skills in L1. In contrast, it

was expected that students with higher L2 proficiency would be able to successfully transfer native-language literacy skills and knowledge when reading the L2 text and their comprehension would not be hampered. Hence, they would perform well across comprehension tasks. However, if, at the same time, L1 input indeed served as a scaffold to facilitate more accomplished performance in L2, it was expected that students' performance on L2 reading comprehension tasks would be better with than without L1 input. Furthermore, differences in performance across reading conditions (i.e., L1-L2 vs. L2-L2) were suspected would be greater among students with lower L2 proficiency than among students with higher L2 proficiency since they needed the most language support.

It was expected that the present research study would increase our understanding of cross-linguistic influences on reading comprehension among young adults, and also help inform the role of instruction in facilitating learning for this population.

Chapter 3: Rationale and Hypotheses

Preliminary Study

A pilot study was conducted in the spring of 2008 to examine whether L1 input would facilitate L2 reading comprehension. One of the goals of the pilot study was to account for the limitations encountered in previous research. For instance, L2 proficiency was often gauged based on the number of semesters the student had been studying the L2, their scores on a few multiple-choice questions, or the student's ability to read a list of words. The present study measured more aspects of L2 and L1 oral language proficiency and lexical development including expressive vocabulary, reasoning using lexical knowledge, reading ability, and applied language proficiency using a psychometrically-sound standardized test. In addition, this study examined online reading processes as the students read a text aloud through the use of a think-aloud procedure. Finally, reading comprehension was assessed through multiple tasks.

The main goal of the study was to investigate the effects of cross-linguistic input and L2 proficiency on L2 reading comprehension among Spanish-speaking adults learning English as a second language. The main hypothesis was that adults with adequately developed L1 language and reading skills who are learning English as a second language would comprehend printed text more effectively if they read the text first in L1 and then in L2, than if they read the text twice only in L2. Although children of limited English proficiency may not be more fluent in their native language than in English, the reverse may be true for young adults with limited English proficiency who

have adequately developed native-language skills. These adults may benefit from interventions that make increased use of their L1.

It was expected that reading a text in one's native language would provide a scaffold in the form of background knowledge that would allow readers to make connections between their first language and the information repeated in the second language, thus, facilitating better learning in the second language. That is, learners would use what they know in L1 (e.g., vocabulary and content knowledge) to facilitate their understanding of the L2 input. Although it may not be common for adults who are learning a second language to read a text in one language followed by a reading of the same text in another language, in the context of research, this repeated-reading paradigm can shed light on the nature of cross-linguistic influences in forming text representations when background knowledge is limited. For instance, students in bilingual programs in New York City public schools may have access to storybooks written in both Spanish and English languages, but teachers may read the storybooks in only one language during class. Students can read the storybooks in the other language if they choose, but during their free reading time (L. Kaplan, personal communication, May 14, 2010).

Twenty undergraduate students (16 females and 4 males) fluent in Spanish according to the Woodcock-Muñoz Language Survey-Revised ([WMLS-R]; Woodcock, Muñoz-Sandoval, Ruef, & Alvarado, 2005) participated in the pilot study. The participants were enrolled in remedial ESL courses at a local community college in New York City. All the students were born outside of the United States and were between 17.8 to 37.9 years of age at the time of the study ($M = 23.0$ years, $SD = 5.86$). Ten students were classified as having "lower" English proficiency (i.e., L2 CALP 2) and 10

students were classified as having “higher” English proficiency (i.e., L2 CALP 3) based on their cluster scores on the English version of tests 1 through 4 of the WMLS-R. Within each L2 CALP level, students were randomly assigned to one of two reading conditions (i.e., L1-L2 and L2-L2), forming four experimental groups. Students in the L1-L2 reading condition read a text in Spanish followed by an equivalent version in English, whereas students in the L2-L2 reading condition read the text twice in English. Students were asked to “think aloud” after reading each sentence when reading the text for the second time. Reading comprehension was assessed via a recall task, eight open-ended comprehension questions, and a rational-deletion cloze task, in that order. All testing was done individually in a closed study room within the college’s library. Each student was seen twice within a four-week period. The Principal Investigator administered all assessments. None of the tasks had a time limit. On average, the first meeting lasted approximately 55 minutes and the second meeting lasted approximately 40 minutes.

Three ANOVAs were conducted with two between-subjects factors, L2 proficiency level (higher *vs.* lower) and language of text (Spanish/English *vs.* English/English) using the raw scores on the comprehension measures as the dependent variables. The analyses revealed a statistically significant interaction between L2 proficiency and language of text on the recall and Q&A tasks. Among students with lower L2 proficiency, those who read the text in both languages provided more complete retellings and answered the comprehension questions more accurately than students who read the text only in English. On the cloze task, a statistically significant main effect for L2 proficiency was found. Students with higher L2 proficiency performed significantly

better on the cloze task than students with lower L2 proficiency. Inspection of the think-aloud protocols revealed that students accessed both languages when reading the L2 text to interpret the information and to monitor their comprehension regardless of the reading condition to which they were assigned or their level of L2 proficiency.

The results of the pilot study provided partial support for the main hypothesis. It appeared that for students with lower English proficiency, reading the text in both languages significantly helped the students in comprehending the text, at least as measured by the recall and Q&A tasks. However, for students with higher English proficiency, the language of the text tended to have less impact. Given the statistical limitations posed by the small sample size, the findings were only preliminary and had to be considered with caution. The dissertation study, therefore, consisted of an expansion and modification of the pilot study. It was an expansion in that it involved a replication of the methods and procedures employed in the pilot study but with a larger sample. In addition, it was modified to include a control condition in which students were not required to do a think-aloud while reading the text.

Dissertation Study

The aim of the dissertation study was to investigate further the effects of cross-linguistic input and L2 proficiency on L2 reading comprehension among Spanish-speaking adults learning English as a second language. The research questions and hypotheses from the pilot study were retained and included the following:

- Will language-minority adults comprehend printed text better if they read it in L1 followed by L2 than if they read the text twice only in L2? If so, how does L1 input facilitate the comprehension of L2 input? The hypothesis was that

students would be able to comprehend the text better if they read it in L1 followed by L2 than if they read it twice only in L2. It was expected that students would use the L1 input as a scaffold by making connections between the information presented in L1 and the information presented in L2 to facilitate the understanding of unknown vocabulary words appearing in the L2 version of the text as well as the overall ideas presented in the text. It was anticipated that the think-aloud protocols would provide direct evidence for this hypothesis, whereas scores on the comprehension tasks would provide indirect evidence as to the benefits of L1 input.

- Will the contribution of L1 input vary as a function of L2 proficiency? That is, will the facilitative effect of L1 input on L2 reading comprehension depend upon the level of L2 proficiency of the reader? Two simultaneous hypotheses and alternative outcomes were entertained. If, on the one hand, students' ability to successfully transfer native-language knowledge and skills when reading the L2 text was dependent upon an L2 threshold of linguistic competence below which L1 could not positively impact cognitive functioning in L2 as proposed by Cummins (1979, 2000), it was expected that students with lower L2 proficiency would perform poorly across comprehension assessments even though they had good CALP skills in L1. In contrast, it was expected that students with higher L2 proficiency would be able to successfully transfer native-language literacy skills and knowledge when reading the L2 text and their comprehension would not be hampered. Hence, they would perform well across comprehension tasks. However, if, at

the same time, L1 input indeed served as a scaffold to facilitate more accomplished performance in L2, it was expected that students' performance on L2 reading comprehension tasks would be better with than without L1 input. Furthermore, differences in performance across reading conditions (i.e., L1-L2 vs. L2-L2) were suspected would be greater among students with lower L2 proficiency than among students with higher L2 proficiency since they needed the most language support. In the present study, cognitive academic language proficiency (CALP) levels were defined and measured with the Woodcock-Muñoz Language Survey-Revised ([WMLSR]; Woodcock et al., 2005). Based on the definitions of each CALP dimension, CALP Level 2, which is defined as *very limited*, was taken as representing students within the *lower threshold*. Students in the CALP 3 dimension, which is defined as *limited*, were taken as representing "higher" proficiency.

- Will performance across reading comprehension assessment tasks differ? The hypothesis was that students with both higher and lower L2 proficiency would perform better on comprehension tasks that required language recognition and limited language production skills (i.e., short-answer questions) compared to tasks that required more extensive language production (i.e., retelling of the text) because of their limited oral language proficiency in English.

The following research question was introduced:

- What effect does the think-aloud procedure have on comprehension? There were three possible outcomes. On the one hand, it was thought that having to interrupt the reading process to report one's thoughts would hinder

comprehension because of the cognitive burden that the think-aloud would add to an already challenging task. On the other hand, having to stop and comment on what was being read would increase the student's awareness and comprehension by requiring the student to actively interact with, think about, and rehearse the text. A third possibility was that the TA procedure would have no effect; that thinking aloud would merely provide a window for observing processes that normally occur as text is read and comprehended.

Chapter 4: Method

Participants

The participants were 80 Spanish-speaking students enrolled in remedial ESL courses at a community college from the City University of New York (CUNY). The sample included students in the pilot study tested in 2008 combined with students in the extended study tested in 2009. The Principal Investigator visited the classrooms offering the targeted ESL courses during the spring 2008 through the summer 2009 semesters to recruit participants. Student participants were required to sign a written informed consent form, which was distributed in Spanish and English (see Appendix A). Signed consent forms were collected from 168 students. Students were compensated \$20 for agreeing to participate and were allowed to withdraw their participation from the study at any time without consequence. Those who consented to participate were screened for L1 and L2 proficiency using the Woodcock-Muñoz Language Survey-Revised ([WMLS-R]; Woodcock et al., 2005).

Students had to meet the following criteria for inclusion in the study: obtain a CALP (cognitive-academic language proficiency) score of at least 3.5 (*fluent*) in L1 and a CALP score of either 2 (*very limited*) or 3 (*limited*) in L2 as measured by the WMLS-R (Woodcock et al., 2005). Students scoring below CALP 2 in L2 were excluded from the study because it was expected that a student below level 2, defined as having negligible L2 proficiency, would find the language demands of the learning task impossible to manage. Students who scored at level 2 on L2 CALP were classified as having “lower” English proficiency, and students who scored at level 3 were classified as having

“higher” English proficiency. Out of the 168 students that consented to participate, 107 (64%) students completed the first meeting. Of these, 19 (18%) students were excluded because they scored below CALP 2 in L2 and 88 (82%) students met the criteria for inclusion in the study. Among those who met the criteria, 80 (91%) students attended the second meeting.

All the students were born outside of the United States and came from a variety of countries including Dominican Republic, Ecuador, Mexico, Peru, Honduras, Puerto Rico, Guatemala, and Panama. There were 54 females and 26 males. The participants’ mean chronological age at the time of testing was 22.9 years. They had completed up to one year of studies at the community college. The students had been living in the U.S. for an average of approximately 6.3 years.

Materials

Using a 2 x 2 x 2 experimental design, the students were randomly assigned to the L1-L2 and L2-L2 reading conditions within each L2 CALP/Proficiency level (higher vs. lower) and Think-Aloud (present [TA] vs. absent [No TA]) conditions, yielding eight experimental groups. The students met individually with the Principal Investigator on two occasions. In the first meeting, the students were screened for L1 and L2 proficiency using the WMLS-R (Woodcock et al., 2005), and in the second meeting they completed the reading and comprehension tasks. None of the tasks had a time limit. The first meeting lasted, on average, approximately 45 minutes and the second meeting lasted approximately 35 minutes. The following instruments were used:

Woodcock-Muñoz Language Survey-Revised (WMLS-R, Form A; Woodcock et al., 2005). This is a norm-referenced measure of reading, writing, listening, and

comprehension available in Spanish and English. This was considered an appropriate measure given its favorable psychometric properties, availability in both languages, facility in scoring and interpretation, and usefulness in the design as it measures CALP. The WMLS-R battery consists of seven tests measuring 11 academic language clusters. The tests in the *Broad English Ability* and *Broad Spanish Ability* academic language cluster were used as a screening tool to assess L2 and L1 proficiency. Broad language proficiency is a measure based on a cluster score on Tests 1 through 4. The WMLS-R evaluates performance according to six CALP competence levels in addition to other derived scores. The WMLS-R Comprehensive Manual (Alvarado, Ruef, & Schrank, 2005) provides the following descriptions for each CALP dimension:

Level 1: *Negligible*. When compared with others of the same age or grade, an individual at level 1 demonstrates negligible cognitive-academic language proficiency. If provided with monolingual instruction at the subject's chronological age or corresponding grade level, it is expected that a student at level 1 will find the language demands of the learning task impossible to manage.

Level 2: *Very Limited*. When compared with others of the same age or grade, an individual at level 2 demonstrates very limited cognitive-academic language proficiency. If provided with monolingual instruction at the subject's chronological age or corresponding grade level, it is expected that a student at level 2 will find the language demands of the learning task extremely difficult.

Level 3: *Limited*. When compared with others of the same age or grade, an individual at level 3 demonstrates limited cognitive-academic language proficiency. If provided with monolingual instruction at the subject's chronological age or corresponding grade level, it is expected that a student at level 3 will find the language demands of the learning task difficult.

Level 4: *Fluent*. When compared with others of the same age or grade, an individual at level 4 demonstrates fluent cognitive-academic language proficiency. If provided with monolingual

instruction at the subject's chronological age or corresponding grade level, it is expected that a student at level 4 will find the language demands of the learning task manageable.

Level 5: *Advanced*. When compared with others of the same age or grade, an individual at level 5 demonstrates advanced cognitive-academic language proficiency. If provided with monolingual instruction at the subject's chronological age or corresponding grade level, it is expected that a student at level 5 will find the language demands of the learning task very easy.

Level 6: *Very Advanced*. When compared with others of the same age or grade, an individual at level 6 demonstrates very advanced cognitive-academic language proficiency. If provided with monolingual instruction at the subject's chronological age or corresponding grade level, it is expected that a student at level 6 will find the language demands of the learning task extremely easy.

Each test is administered by establishing basal and ceiling levels for each participant. The participants completed eight individually administered tests, four in Spanish and four in English.

Test 1: Picture Vocabulary (Alvarado et al., 2005). This test measured aspects of oral language, including language development and lexical knowledge. Students were required to say the name of pictured objects until they were unable to identify six consecutive pictured objects. The items became increasingly difficult as testing progressed. Administration of the L2 test began at item 7, which corresponded to items in the Kindergarten to Grade 2 ability range. The highest level administered was Grade 7-Average Adult. The test had a total of 59 items. Administration of the L1 test began at item 21, which corresponded to items in the Grades 3-4 ability range. The highest level administered was the University Students-Above Average Adult. The test

had 58 items. According to the publisher, this test has a median reliability of .91 in the age range of 5 to 19 years and .95 in the adult range.

Test 2: Verbal Analogies (Alvarado et al., 2005). This test measured aspects of oral language, including the ability to reason using lexical knowledge. Students listened to three words of an analogy and were asked to complete the analogy by verbally producing an appropriate fourth word. The relationships among the words became increasingly difficult as the test progressed. Testing ceased when the student was unable to complete six consecutive analogies. Administration of the L2 test began at item 1, which corresponded to items in the Preschool to Kindergarten ability range. The highest level administered was College-Above Average Adult. The test had a total of 35 items. Administration of the L1 test began at item 9, which corresponded to items in the Grades 4-5 ability range. The highest level administered was the University Students-Above Average Adult. The test had 40 items. According to the publisher, this test has a median reliability of .90 in the age range of 5 to 19 years and .94 in the adult range.

Test 3: Letter-Word Identification (Alvarado et al., 2005). This test measured aspects of reading ability, including letter and word identification skills. Students were required to read as fluently as possible a list of words until they were unable to read six consecutive words. The words became increasingly difficult as the test progressed. Administration of the L2 test began at item 24, which corresponded to items in the mid second-grade ability level. The highest level administered was College-Above Average Adult. The test had a total of 75 items. Administration of the L1 test began at item 29, which corresponded to items in the fifth-grade ability level. The highest level administered was the University Students-Above Average Adult. The test had 75 items.

According to the publisher, this test has a median reliability of .97 in the age range of 5 to 19 years and .98 in the adult range.

Test 4: Dictation (Alvarado et al., 2005). This test measured aspects of applied language proficiency, including ability to respond in writing to questions pertaining to letter forms, spelling, punctuation, capitalization, bound morpheme agreement, and word usage. For example, students were asked to provide, in writing, the missing letter of a word, spell a word or proper name, provide the plural of a word, provide the contraction of two words, or add punctuation marks. Most of the words the students were asked to spell were given within the context of a sentence. Administration of the test stopped when the student was unable to provide appropriate written responses to six consecutive items. Administration of the L2 test began at item 17, which corresponded to items in the third-grade ability level. The highest level administered was College-Above Average Adult. The test had a total of 62 items. Administration of the L1 test began at item 17, which corresponded to items in the Grades 4-6 ability range. The highest level administered was the University Students-Above Average Adult. The test had 59 items. According to the publisher, this test has a median reliability of .94 in the age range of 5 to 19 years and .96 in the adult range.

Demographics and Language Background Interview. This was an oral interview used to obtain information regarding individual characteristics such as place of birth, years of schooling in Spanish and English, years living in the United States, primary language spoken in everyday interactions, etc. The participants were not required to provide written responses (see Appendix B for a complete list of questions).

Reading Passage. An expository text adapted from *The Kim Marshall Series: Reading Book 1* (Marshall, 2000) was used in the study. Several researchers have argued that expository texts are more difficult to process than narrative texts because expository texts are often more unfamiliar in terms of structure and content, they are more unpredictable, and they require more attention to lower-level processes, leaving fewer resources available for higher-order text integration (Zabrucky & Moore, 1999). However, given that most texts encountered by these students in their L2 classes are expository texts, an informative expository text unfamiliar to the participants was used.

The text was entitled *Danger Under Water* (see Appendix C). The original version, written in English, consisted of over 500 words and had a 7.0 grade level readability. The text was edited to include 298 words and a 5.1 grade level readability based on the Harris-Jacobson Readability Formula (Harris & Sipoy, 1980). The text was then translated into Spanish by the Principal Investigator and revised for precision and lack of bias by two graduate students, one who was residing in Puerto Rico and one in New York who was studying to be a Spanish teacher. The Spanish version was 336 words long and had a Fernando Huerta Reading Ease score, which is a readability score analogous to the Flesch-Kincaid score, of 83 (“easy”). The Flesch-Kincaid score of the English version was 70 (“fairly easy”). The text was presented on a computer screen one sentence at a time as a Microsoft PowerPoint presentation. The student was required to press the spacebar to advance the text to the next sentence.

Although the Spanish text was longer than the English text, the Spanish text provided no additional information, and each contained equivalent number of statements. Both texts consisted of 17 sentences plus the title. The average sentence length for the

English text was 16.5 words, whereas the average sentence length for the Spanish text was 18.6 words. The disparity in word count across texts was due to language features of the Spanish language system not present in English. For instance, unlike English, Spanish nouns, adjectives, and articles have masculine and feminine markers. In addition, unlike English, Spanish uses double negative, and it has no use of auxiliary for negation. Inspection of the texts revealed that many of the additional 38 words in the Spanish text constituted prepositions and grammatical markers indicating number and gender agreement. In other instances, translating into Spanish required many words to preserve the meaning embedded in the English text. For example, the first two clauses in one sentence were made up of 10 words and read: *Although the submarine had sunk, some air remained trapped inside....* The Spanish equivalent read: *A pesar de que el submarino se había hundido, un poco de aire permaneció atrapado adentro....*, and had 16 words.

Think-Aloud vs. No Think-Aloud Protocol Manipulation. A think-aloud (TA) technique similar to Olshavsky's (1977) was followed for the TA condition in order to identify strategies used to comprehend the text and to investigate the effects, if any, of the TA on reading comprehension tests. Students were asked to read the text twice. During the first reading, students read the text aloud in either L1 or L2, sentence by sentence without interruption. During the second reading, students read the text aloud again, sentence by sentence, but this time the text was presented in L2. In addition, they were required to "think aloud" at the prompt, which consisted of a red star at the end of each sentence. Instructions were delivered following a script so that the process of introducing

the task was uniform for all participants. Students were not “taught” or given practice on how to think aloud. Instead, the students heard the following instructions:

Now you are going to read the same story aloud to me a second time around, but I will ask you to read it in a special way this time. I’ll explain how. Stop me if something doesn’t make sense and I’ll explain. I would like you to think aloud after you read each sentence. That is, I want you to tell me everything that comes to mind after you read each sentence. For example, maybe you think about what will happen next. Maybe you think of a different way to explain what is happening. Maybe you don’t understand so you talk about what is confusing or what words you don’t know. I want you to tell me every thought that comes to mind as you read the story even if it doesn’t seem related to the story. This will help me understand how you are making sense of the story and how it relates to your ideas and experiences. You will see a red star at the end of each sentence to remind you to stop and say everything that you are thinking about at the moment. This will tell me whether and how you are understanding the story. You may do this in your language of preference. Try to understand the story as best you can, because I’ll be asking you some questions when you finish. Do you have any questions? Are you ready? Remember to read the story aloud to me and think aloud after you read each sentence. If you come to some hard words, I won’t be able to help you, so please read them the best you can. Let’s begin.

If the student remained silent for three seconds after reading the sentence, he/she was prompted to think aloud by asking, “*What are you thinking?*” If the student read on without stopping to report his/her thoughts, he/she was prompted to think aloud by saying, “*Remember to tell me everything that comes to mind after you read each sentence.*” Students were told to use whatever language they preferred to verbalize their thoughts as not to confound language production skills with the extent and quality of the verbalizations (see Appendix D for a complete script of the instructions given to the participants).

The decision to use a think-aloud procedure was made with the expectation that it would expose any transfer of knowledge from the first reading, either L1 or L2, as well as provide a window on comprehension processes. The participants’ verbal reports were

audio-taped, transcribed, and coded to identify strategies used. In the No TA condition, all procedures remained the same with the exception that students were not required to think aloud at the prompt. Instead, they read the text aloud without interruption.

Reading Comprehension Tasks. As mentioned before, assessing learners' reading comprehension is not straight-forward because the processes by which readers create meaning are invisible and the results from each assessment task only provide a partial picture of reading comprehension. Therefore, students were asked to complete multiple tasks. The tasks were administered in the order described below (see Appendix D for a detailed explanation of the instructions given to the participants in each task).

Recall task. Reading comprehension was first measured by means of a recall task in which the participants were asked to, "*Tell me everything you remember from the story you just read. Pretend that you are telling the story to someone who has never read it or heard it.*" According to Bernhardt (1991), this task has the potential to provide the best measure of comprehension; however, this may not necessarily be the case with this population because of their limited oral language abilities. The recall task was scored based on the number of idea units recalled. Each correct idea was analyzed for completeness and given a maximum score of 1. The maximum total score was 24 (see scoring sheet in Appendix E). In addition to hand-notations, the retellings were audio-taped and transcribed to allow for greater precision in scoring. Although it has been argued that individuals should be allowed to reply in any language so as not to confound responses on the measure with production skills in the target language (e.g., Brisbois, 1995; Friesen & Jared, 2007; Goldman et al., 1984), the participants were asked

to complete the recall task in English because the goal of the study was to examine the extent to which L1 input affected L2 output.

To establish the reliability of the scoring key, an independent rater was given a predetermined list of idea units. The rater was trained on how to interpret and score each idea unit and was asked to score 10% of the cases using the list. The rater's scores were then compared with the researcher's scores in order to determine inter-rater reliability. The percentage of agreement between the researcher and the independent rater was 91%.

Q&A task. The second reading comprehension assessment consisted of eight open-ended comprehension questions that required participants to provide a verbal short-answer. The questions could be answered with information presented in the text. The questions appeared one at a time on the computer screen using Microsoft PowerPoint. The student was required to press the spacebar to advance through the questions. As part of the instructions, they were asked to provide their answers orally in English. Partial credit was given for correct but incomplete responses. For example, one of the questions was, "*Why was the submarine dangerous?*" The expected response was that the submarine was dangerous because it was carrying torpedoes filled with mercury. If the participant only responded that the submarine was dangerous because "it had poison" or "it had mercury" or "it had torpedoes", only half of a point was credited. If the participant responded that the submarine had "poisoned torpedoes" and could not name the poison as mercury, only three-quarters of a point were credited.

Students were not allowed to refer back to the text while answering the questions. The answers were audio-taped and transcribed to allow for greater precision in scoring. The maximum possible score was 9 (see Appendix F for a complete list of the

questions asked and the scoring key). To establish the reliability of the scoring key, an independent rater was given an answer key with examples of possible answers and their scores for each question. The rater was trained on how to analyze and score each question for completeness and was asked to score 10% of the cases using the answer key. The rater's scores were then compared with the researcher's scores in order to determine inter-rater reliability. The percentage of agreement between the researcher and the independent rater was 94%.

Cloze task. The cloze task was the last task administered. In this task, the participants were given the English version of the text with 43 words omitted. For this task, the text was printed in 12-point Times New Roman type font double-spaced on an 8.5x11 paper with omitted words replaced by underlined blanks of equal length so as not to provide clues regarding the length of the missing word. For example, one of the sentences was: "When the _____ (government) in Florida heard about this, it was _____ (worried) because the wreck was near a _____ (public) beach." The omitted words were chosen pseudo-randomly following Greene's (2001) technique so that the number and placement of the omissions assessed integration and understanding of content as opposed to syntactic knowledge. With this technique, every seventh word was deleted, but when the seventh word was not a content word the closest content word was deleted instead. Of the 43 omissions, 17 words (30.5%) were the actual seventh word; another 17 words were the preceding word; 8 words (18.6%) were the following word; and in one instance, the word omitted was the third word after the seventh word since the two preceding and the two following words were prepositions. The answers for 15 of the gaps could be inferred from the sentences elsewhere in the text.

Credit was given only when the student provided the exact word that appeared in the text to complete the sentence. Students were asked to write the answer in the blank space, and to answer in English. The maximum possible score was 43 (see cloze task and scoring key in Appendix G). The *Spearman-Brown prophecy coefficient* for split-half reliability was .81.

Exit Interview. At the end of the reading and assessment session, students were asked nine brief questions in order to learn about their perceptions regarding the difficulty level of the text for both Spanish and English versions, their level of background knowledge on the topic, difficulties with vocabulary words, and their perceptions in other aspects of the tasks. Some of the questions were: *Was the story difficult or easy to understand? Why? Were you familiar with this story? Where there any words you didn't understand? If you had to rank the tasks from hard to easy, which task was easiest? Hardest?* The questions were asked in the language preferred by the student, and the student was allowed to answer in the language preferred. Students provided oral responses which were written down by the researcher (see complete list of the questions asked in Appendix H).

Procedure

Students who provided written informed consent were invited to participate. Participants were seen individually on two occasions during off-class hours in the college's library. On average, the meetings lasted approximately 40 minutes. The Principal Investigator conducted all testing sessions. During the first meeting, students completed the WMLS-R (Woodcock et al., 2005), first in L2 and then in L1, and the Demographics and Language Background Interview. The reason why the WMLS-R was

first administered in L2 to all students was to avoid confounds due to fatigue or transfer from L1 to L2 on items that were repeated on both measures. This session served as the screening session. Students who met the selection criteria were invited to the experimental session. The second meeting took place at a later day within a four-week period.

During the second meeting, the students read an expository text twice. Based on their cluster scores on the L2 WMLS-R (Woodcock et al., 2005), the students were classified as having lower (i.e., CALP 2) or higher (i.e., CALP 3) English proficiency. Students were then randomly assigned to the L1-L2 or L2-L2 reading conditions within each L2 CALP level and TA conditions. Students in the TA condition completed a think-aloud procedure, whereas students in the No TA condition did not complete a think-aloud procedure. Those in the L1-L2 reading condition read the text in Spanish first and then in English (i.e., Spanish/English reading condition), whereas those in the L2-L2 reading condition read the text twice in English (i.e., English/English reading condition). Hence, there were four groups in each of the TA conditions for a total of eight groups with 10 participants in each group.

As depicted in Figure 1, Group 1 consisted of students with lower English proficiency assigned to the Spanish/English and TA conditions; Group 2 consisted of students with lower English proficiency assigned to the English/English and TA conditions; Group 3 consisted of students with lower English proficiency assigned to the Spanish/English and No TA conditions; Group 4 consisted of students with lower English proficiency assigned to the English/English and No TA conditions. Groups 5 through 8

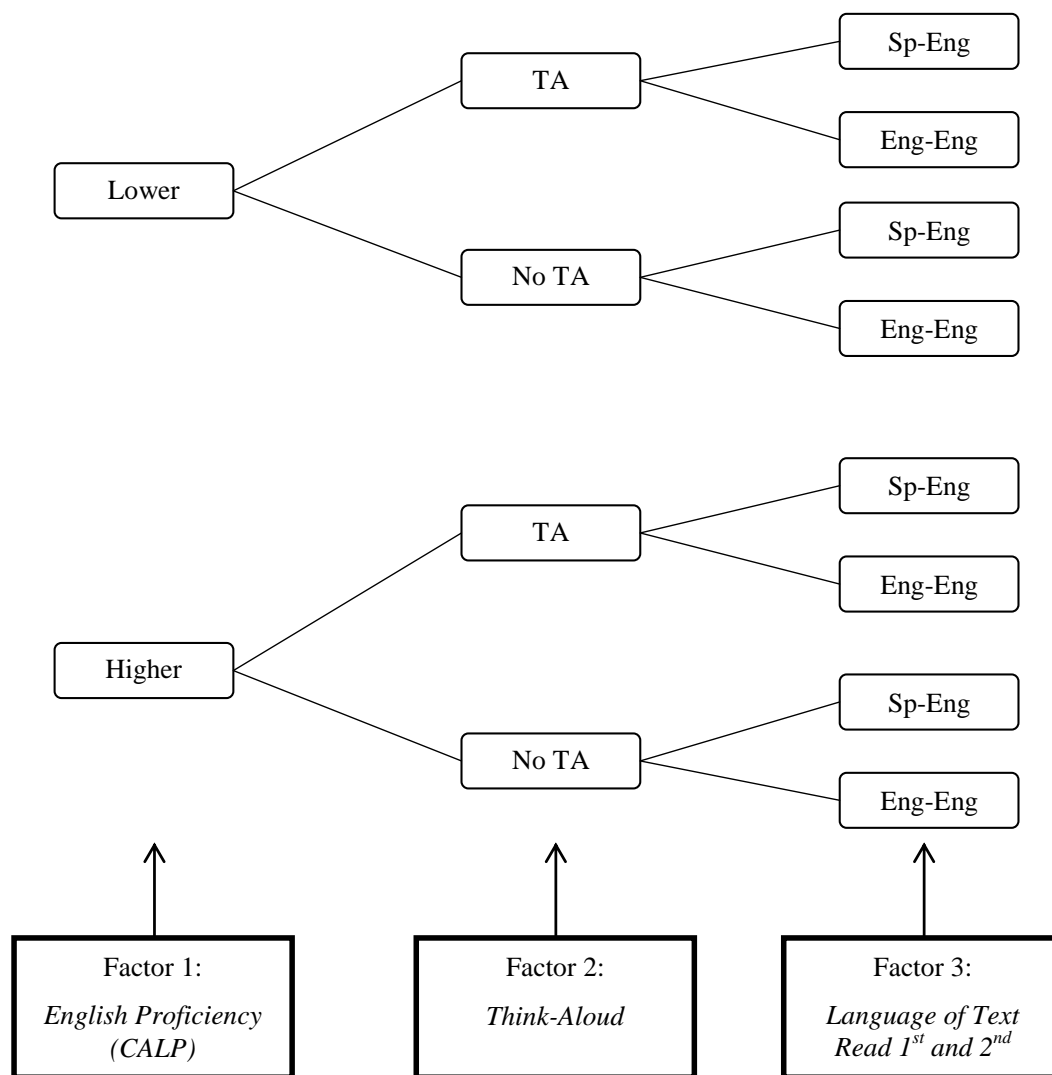


Figure 1. Study Design.

mirrored Groups 1 through 4 with the exception that they consisted of students with higher English proficiency.

Students were asked to read the text aloud on both occasions. Students in the TA condition were also asked to “think aloud” after reading each sentence during the second reading of the text, which was always in English. The purpose of the TA manipulation was to investigate whether the think-aloud had an effect on reading comprehension. A red star appeared at the end of each sentence to remind the student to stop and report on his/her thinking. Students in the No TA condition were not asked to think aloud after reading each sentence. Instead, they read the text twice without interruption and were not shown the red star at the end of each sentence. The text was presented on a computer using a Microsoft PowerPoint slide show. Students were only able to see one sentence at a time. They were instructed to press the spacebar on the computer keyboard in order to advance the slide to the next sentence whenever they were ready to proceed. The students were able to read the focal sentence as many times as they wished before pressing the spacebar. After the participants finished reading the text twice, they were asked to complete the recall task, answer eight open-ended questions that required an oral short-answer, and complete the cloze task, in that order. Lastly, the students completed the exit interview. None of the tasks had a time limit.

Design and Statistical Analyses

The study was conducted employing a 2 x 2 x 2 experimental design. There were three independent variables with two levels each: L2 language proficiency level (higher vs. lower), think-aloud (present vs. absent), and language of text (Spanish/English vs. English/English). The Statistical Package for Social Sciences (SPSS 17) program was

used to conduct tests of the hypotheses. The nominal level of alpha was .05. According to Cohen's (1988) power analysis, a sample size of 80, achieved in this study, would yield a small to moderate effect size of $\pm.33$ at $\alpha = .05$ with 80% power.

Before conducting statistical analyses to test the hypotheses, preliminary analyses were conducted to ensure that the sample was comparable in a variety of language measures tested in Spanish, and that it differed in a variety of language measures tested in English. An independent sample *t*-test was computed comparing lower English proficiency and higher English proficiency groups. The independent variable was L2 proficiency level (higher vs. lower). The dependent variables were the raw scores on the eight tests as well as the cluster scores on the WMLS-R (Woodcock et al., 2005) English and Spanish forms.

Additional preliminary analyses were conducted to ensure that the students recruited in 2008 were comparable to the students recruited in 2009. To do this, two independent sample *t*-tests were conducted to assess whether the samples differed on a number of demographic characteristics collected during the initial interview and on the L1 and L2 proficiency measures. One *t*-test was computed comparing the 2008 and 2009 samples of students with lower English proficiency. A second *t*-test was computed comparing the 2008 and 2009 samples of students with higher English proficiency. The independent variable was year of recruitment (2008 vs. 2009). The dependent variables were: age, years living in the U.S., years of schooling, number of ESL classes taken, percent of daily use of L2, and the raw score on the eight tests as well as the cluster scores on the WMLS-R (Woodcock et al., 2005) English and Spanish forms. Chi-square tests were computed on gender, whether L2 was learned in the birth country, whether L2

was used in the birth country, whether they attended school in the U.S., and whether they had repeated an ESL class.

To rule out the potential effect of term in which the student was recruited on reading comprehension performance, a MANOVA was computed to investigate whether there were significant differences in performance on the reading comprehension measures as a function of various factors. The independent variables were: term (Spring 08 *vs.* Spring 09 *vs.* Summer 09 *vs.* Fall 09), think-aloud (present *vs.* absent), L2 proficiency level (higher *vs.* lower), and language of text (Spanish/English *vs.* English/English). The raw scores on the reading comprehension measures were used as the dependent variables.

As a further precaution, a MANOVA was computed to verify that the eight experimental groups only differed in English proficiency and not in Spanish proficiency at the onset of the study. The independent variables were: L2 proficiency level (higher *vs.* lower), think-aloud (present *vs.* absent), and language of text (Spanish/English *vs.* English/English). The dependent variable was the cluster score on the WMLS-R (Woodcock et al., 2005) for both English and Spanish forms.

To test the hypotheses set out in the study, a series of ANOVAs were computed to investigate whether there were significant differences in performance on the reading comprehension measures as a function of think aloud and language of text at each level of English language proficiency. Separate analyses were conducted on the two proficiency levels since they were expected to differ in their reading comprehension performance and, therefore, there was little interest in comparing their scores directly. Instead, the primary interest was in the impact of the language of text and the think-aloud procedure on reading comprehension within each L2 proficiency level. The independent

variables were: think-aloud (present *vs.* absent), and language of text (Spanish/English *vs.* English/English). The dependent variables were the raw scores on the three reading comprehension measures. The maximum score on the recall, Q&A, and cloze tasks was 24, 9, and 43, respectively. To compare performance across English proficiency levels and language of text conditions, *Cohen's d* effect sizes were calculated using the formula $d = M1 - M2/\sigma_{\text{pooled}}$.

In addition, a 2 x 3 ANOVA was computed to investigate whether students performed differently as a function of test type at each level of L2 proficiency. The independent variables were: L2 proficiency (higher *vs.* lower), and test type (Recall *vs.* Q&A *vs.* Cloze). The latter was a repeated measure. The dependent variable was the percent of correct responses on the three reading comprehension measures.

In order to assess relationships between performance on the reading comprehension measures and students' L2 and L1 proficiency on each of the component tests of the WMLS-R (Woodcock et al., 2005), Pearson product-moment correlation coefficients were computed. This was followed by standard multiple regression analyses between the WMLS-R L1 and L2 measures as independent variables and reading comprehension as the dependent variable to investigate the relative contributions, unique and shared, of L2 and L1 proficiency to reading comprehension. In each regression, the WMLS-R L2 cluster score was entered first and the WMLS-R L1 cluster score was entered second using the *enter* (default) method for each outcome variable: recall, Q&A, and cloze. All correlational and regression analyses were conducted in two ways: for the total sample and by language of text reading condition.

Qualitative analyses were conducted on the think-aloud protocols to investigate online reading processes and strategies used while reading the text. Students' comments while reading the text were transcribed and coded according to a rubric. ANOVA analyses were computed comparing the proportion of strategies reported by group when appropriate, using L2 proficiency level (higher vs. lower) and language of text (Spanish/English vs. English/English) as independent variables. Lastly, comments on the exit interview were reviewed and analyzed for patterns in responses.

Chapter 5: Results

Characteristics of Participants

The purpose of this research study was to investigate the effect of L1 input and L2 proficiency on L2 reading comprehension among language-minority adults. The main research question was whether adults who are learning English as a second language would be able to comprehend printed text more effectively if they read the text first in L1 and then in L2, than if they read the text twice only in L2. The participants were 80 (54 females and 26 males) Spanish-speaking students enrolled in remedial ESL courses at a CUNY community college.

All the students were born outside of the United States. Forty-nine students were from the Dominican Republic, 10 students were from Ecuador, 8 students were from Mexico, 6 students were from Peru, 3 students were from Honduras, 2 students were from Puerto Rico, and the remaining 2 students were from Guatemala and Panama. The participants were between 16.9 and 39.6 years of age at the time of testing ($M = 22.9$ years, $SD = 5.46$). The students were in their first to third semester of undergraduate studies in the U.S. Two students had completed two years of college in their native countries and three had completed a bachelor's degree. Seventeen students (21%) were repeating the ESL 094 course. Forty-three students (54%) reported that they learned "basic" English in their birth countries but did not speak English outside of the context of the classroom. Forty-four students (55%) had been living in the U.S. for five years or less, while the others had been living in the U.S. "on and off" for up to 25 years ($M = 6.3$ years, $SD = 5.00$). Forty-seven students (59%) reported speaking or writing in English

about 50% of the time or less during everyday interactions. Twenty-four students (30%) reported having taken 1 ESL course, 44 students (55%) reported having taken 2 ESL courses, 10 students reported having taken 3 ESL courses, and 2 students reported having taken 5 or 6 ESL courses.

L1 and L2 Language Proficiency

Students were screened for L1 and L2 proficiency using the WMLS-R (Woodcock et al., 2005). Students had to meet the following criteria for inclusion in the study: obtain a CALP score of at least 3.5 (*fluent*) in L1 and a CALP score of 2 (*very limited*) or 3 (*limited*) in L2 as measured by the WMLS-R. Students who scored at level 2 on L2 CALP were classified as having lower English proficiency ($n = 40$), and students who scored at level 3 were classified as having higher English proficiency ($n = 40$).

The average L2 raw cluster score across the tests by students with lower English proficiency (i.e., CALP 2) was 135, which equaled a grade equivalent of approximately fourth-grade ($range = 3.1$ to 4.8). On average, tasks below the third-grade level would have been quite easy for these students, whereas tasks above the grade 5.5 level would have been quite difficult. The average grade equivalent for word reading (i.e., Test 3: Letter-Word Identification) was 9.0 ($range = 5.0$ to 18.0).

The average L2 raw cluster score across the tests by students with higher English proficiency (i.e., CALP 3) was 154, which equaled a grade equivalent of approximately sixth-grade ($range = 4.9$ to 7.1). On average, tasks below the fifth-grade level would have been quite easy for these students, whereas tasks above the grade 7.8 level would have been quite difficult. The average grade equivalent for word reading was 12.7 ($range = 8.0$ to 18.0).

As expected, students' L1 CALP levels varied more, ranging from 3.5 to 5.0. The average L1 raw cluster score by students with lower English proficiency was 181, which equaled a grade equivalent of approximately 12th grade. The average L1 raw cluster score by students with higher English proficiency was 184, which equaled a grade equivalent of approximately 13th grade. The average grade equivalent for word reading (i.e., Test 3: Letter-Word Identification) was >18.0 (*range* = 11.0 to >18.0). These values were very similar, suggesting that the two groups did not differ in L1 proficiency.

An independent sample *t*-test was conducted to confirm that L2 CALP 2 and L2 CALP 3 groups differed in L2 proficiency, but not in L1 proficiency. The independent variable was L2 proficiency level (higher *vs.* lower). The dependent variables were the raw scores on each of the eight tests on the WMLS-R (Woodcock et al., 2005), as well as the cluster scores on the WMLS-R for both the L1 and the L2 forms. Mean performance is shown in Table 1. Results confirmed that the groups had equivalent proficiency levels in L1, and that they did differ significantly on all aspects of L2 proficiency.

Time of Recruitment Effects

Independent sample *t*-tests were conducted to ascertain that there were no entry-level differences resulting from time of recruitment, and that the samples were equivalent. One set of *t*-tests was computed comparing the 2008 and 2009 samples of students with lower English proficiency on demographic factors and language skills. Another set of *t*-tests was computed comparing the 2008 and 2009 samples of students with higher English proficiency on demographic factors and language skills. The independent variable was year of recruitment (2008 *vs.* 2009). The dependent variables were: age, years living in the U.S., years of schooling, number of ESL classes taken,

Table 1

Mean Performance (and Standard Deviation) on the Component Tests of the Woodcock-Muñoz Language Survey-Revised by Level of English Proficiency

	<i>Lower Proficiency</i>	<i>Higher Proficiency</i>	<i>t stat</i>
<i>WMLS-R Screening Measure</i>			
<i>English Form</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	
Expressive Vocabulary (max. 59)	20.87 (3.38)	27.65 (3.66)	8.60***
Analogies (max. 35)	15.55 (1.60)	18.25 (2.44)	5.86***
Word Reading (max. 75)	59.83 (4.58)	64.92 (2.88)	5.97***
Dictation (max. 62)	39.10 (3.21)	43.48 (3.85)	5.52***
L2 Cluster Score (max. 231)	135.35 (6.30)	154.30 (5.76)	14.04***
<i>Spanish Form</i>			
Expressive Vocabulary (max. 58)	44.50 (2.89)	44.95 (3.24)	.66ns
Analogies (max. 40)	27.65 (4.06)	28.73 (4.33)	1.15ns
Word Reading (max. 75)	66.75 (3.80)	67.88 (4.19)	1.26ns
Dictation (max. 59)	41.83 (4.17)	42.77 (3.77)	1.07ns
L1 Cluster Score (max. 232)	180.73 (10.29)	184.33 (11.45)	1.48ns

N = 80. *** $p < .001$. ns = not significant.

percent of daily use of L2, and the raw score on each of the eight tests on the WMLS-R (Woodcock et al., 2005), as well as the cluster scores on the WMLS-R for both L1 and L2 forms. Chi-square tests were computed to compare groups on percentage of students who learned L2 in the birth country, percentage of students who used L2 in the birth country, percentage of students that attended school in U.S., and percentage of students that had repeated an ESL class.

Demographic characteristics of the participants with lower English proficiency are shown in Table 2. It is evident that students in the two samples were quite similar on many of the characteristics. Results of the Pearson chi-square tests and *t*-tests revealed only one statistically significant difference. On average, students recruited in 2009 reported using L2 more often than students recruited in 2008. Students did not differ on any other demographic characteristic. Mean performance on the WMLS-R (Woodcock et al., 2005) language tests among participants with lower English proficiency is shown in Table 3. Results indicated that students recruited in 2008 had equivalent language proficiency levels compared to students recruited in 2009 in both L1 and L2 on all tests and overall CALP.

Demographic characteristics of the participants with higher English proficiency are shown in Table 4. It is apparent that these students were also quite similar on many of the characteristics. Here too, results of the Pearson chi-square tests and *t*-tests revealed only one statistically significant difference. On average, students recruited in 2009 also reported using L2 more often than students recruited in 2008. In addition, a greater proportion of students recruited in 2009 were males than students recruited in the 2008. Students did not differ on any other demographic characteristic.

Table 2

Demographic Characteristics of Students With Lower English Proficiency by Year of Recruitment

	Lower English Proficiency			
	2008	2009	Overall	
	Sample	Sample	Sample	
<u>Demographics</u>	<u>N</u>	<u>N</u>	<u>N</u>	
Sample size	10	30	40	
Gender	3M / 7F	7M / 23F	10M / 30F	
Ethnicity				
Dominican Republic	9	22	31	
Ecuador	0	3	3	
Mexico	1	2	3	
Peru	0	2	2	
Guatemala	0	1	1	
	<u>Mean (SD)</u>	<u>Mean (SD)</u>	<u>Mean (SD)</u>	<u>t stat</u>
Age	22.02 (5.16)	21.49 (3.36)	21.62 (3.82)	.38ns
Years living in U.S.	4.20 (2.29)	4.60 (2.61)	4.50 (2.51)	.43ns
Years of schooling	12.95 (1.19)	12.77 (.74)	12.81 (.86)	.58ns
ESL classes taken	1.80 (.63)	1.97 (.99)	1.92 (.92)	.49ns
% daily L2 use in the U.S.	40%	60%	50%	2.82**
	<u>%</u>	<u>%</u>	<u>%</u>	<u>X² stat</u>
L2 learned in birth country	50%	63%	60%	.56ns
L2 use in birth Country	40%	13%	20%	3.33ns
Non-college schooling in U.S.	80%	63%	68%	.95ns
% repeating ESL class	20%	20%	20%	.00ns

** $p < .01$. ns = not significant.

Table 3

Mean Performance (and Standard Deviation) of Students With Lower English Proficiency on the Component Tests of the Woodcock-Muñoz Language Survey-Revised by Year of Recruitment

	Lower English Proficiency		
	2008	2009	<i>t</i> stat
	Sample	Sample	
<i>WMLS-R Screening Measure</i>			
<u>English Form</u>	<u>Mean (SD)</u>	<u>Mean (SD)</u>	
Expressive Vocabulary (max. 59)	19.90 (3.14)	21.20 (3.44)	1.06 <i>ns</i>
Analogies (max. 35)	15.00 (2.00)	15.73 (1.44)	1.26 <i>ns</i>
Word Reading (max. 75)	61.10 (5.57)	59.40 (4.22)	1.02 <i>ns</i>
Dictation (max. 62)	37.90 (2.64)	39.50 (3.32)	1.38 <i>ns</i>
L2 Cluster Score (max. 231)	133.90 (7.22)	135.83 (6.02)	.84 <i>ns</i>
<u>Spanish Form</u>			
Expressive Vocabulary (max. 58)	43.50 (3.21)	44.83 (2.76)	1.27 <i>ns</i>
Analogies (max. 40)	28.20 (4.21)	27.47 (4.07)	.49 <i>ns</i>
Word Reading (max. 75)	68.60 (4.45)	66.13 (3.42)	1.83 <i>ns</i>
Dictation (max. 59)	40.70 (5.93)	42.20 (3.45)	.76 <i>ns</i>
L1 Cluster Score (max. 232)	181.00 (12.88)	180.63 (9.53)	.10 <i>ns</i>

N = 40 (10 students in 2008 and 30 students in 2009). *ns* = not significant.

Table 4

Demographic Characteristics of Students With Higher English Proficiency by Year of Recruitment

	Higher English Proficiency			
	2008	2009	Overall	
	Sample	Sample	Sample	
<i>Demographics</i>	<u>N</u>	<u>N</u>	<u>N</u>	
Sample size	10	30	40	
Gender	1M / 9F	15M / 15F	16M / 24F	
Ethnicity				
Dominican Republic	4	14	18	
Ecuador	3	4	7	
Mexico	0	5	5	
Peru	1	3	4	
Honduras	1	2	3	
Puerto Rico	1	1	2	
Panama	0	1	1	
	<u>Mean (SD)</u>	<u>Mean (SD)</u>	<u>Mean (SD)</u>	<u>t stat</u>
Age	24.06 (6.59)	24.29 (6.59)	24.24 (6.50)	.10ns
Years living in U.S.	7.40 (5.00)	8.33 (6.54)	8.10 (6.14)	.41ns
Years of schooling	12.75 (.49)	12.80 (1.01)	12.79 (.91)	.15ns
ESL classes taken	2.30 (.68)	1.77 (.82)	1.90 (.81)	1.86ns
% daily L2 use in the U.S.	50%	60%	60%	2.12*
	<u>%</u>	<u>%</u>	<u>%</u>	<u>X² stat</u>
L2 learned in birth country	50%	47%	48%	.03ns
L2 use in birth Country	30%	17%	20%	.83ns
Non-college schooling in U.S.	80%	80%	80%	.00ns
% repeating ESL class	40%	17%	23%	2.34ns

* $p < .05$. ns = not significant.

Mean performance on the WMLS-R (Woodcock et al., 2005) language tests among participants with higher English proficiency is shown in Table 5. Results revealed that, on average, students recruited in 2008 performed significantly better on the L1 word reading test than students recruited in 2009. Nevertheless, they had comparable scores on all other L1 tests and on the overall cluster score determining L1 CALP, confirming that both samples were fluent in Spanish. Most importantly, the samples had equivalent proficiency levels in all L2 tests and overall L2 CALP. Taken together, the results confirm that, for the most part, the 2008 and 2009 samples were equivalent on demographic characteristics and cognitive academic language abilities in L1 and L2.

In addition, a MANOVA was computed to determine whether performance on the reading comprehension measures varied as a function of term in which the student was recruited. The independent variables were: term (Spring 08 vs. Spring 09 vs. Summer 09 vs. Fall 09), think-aloud (present vs. absent), L2 proficiency level (higher vs. lower), and language of text (Spanish/English vs. English/English). The dependent variables were the raw scores on the reading comprehension measures. The multivariate test statistic Wilks' lambda for the term factor allowing for heterogeneity of variances was non-significant, $\Lambda = .83$, $F(9, 132) = 1.15$, $p > .05$. Based on these analyses, it was determined that performance on reading comprehension measures by term groups did not differ. Therefore, the variable term was excluded from further analyses and the 2008 and 2009 samples were treated as one.

Comparison of Experimental Groups Prior to Treatment

A MANOVA was computed to verify that the eight experimental groups only differed in L2 proficiency at the onset of the study. No other differences were expected

Table 5

Mean Performance (and Standard Deviation) of Students With Higher English Proficiency on the Component Tests of the Woodcock-Muñoz Language Survey-Revised by Year of Recruitment

	Higher English Proficiency		<i>t</i> stat
	2008	2009	
	Sample	Sample	
<i>WMLS-R Screening Measure</i>			
<i>English Form</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	
Expressive Vocabulary (max. 59)	26.90 (3.90)	27.90 (3.61)	.74 <i>ns</i>
Analogies (max. 35)	17.90 (2.85)	18.37 (2.33)	.52 <i>ns</i>
Word Reading (max. 75)	65.20 (3.68)	64.83 (2.63)	.35 <i>ns</i>
Dictation (max. 62)	43.90 (4.28)	43.33 (3.76)	.40 <i>ns</i>
L2 Cluster Score (max. 231)	153.90 (4.86)	154.43 (6.10)	.25 <i>ns</i>
<i>Spanish Form</i>			
Expressive Vocabulary (max. 58)	44.60 (3.31)	45.07 (3.27)	.39 <i>ns</i>
Analogies (max. 40)	27.80 (3.43)	29.03 (4.60)	.78 <i>ns</i>
Word Reading (max. 75)	71.30 (1.49)	66.73 (4.18)	5.08**
Dictation (max. 59)	43.00 (3.62)	42.70 (3.87)	.22 <i>ns</i>
L1 Cluster Score (max. 232)	186.70 (8.11)	183.53 (12.38)	.75 <i>ns</i>

N = 40 (10 students in 2008 and 30 students in 2009). ** *p* < .01. *ns* = not significant.

because participants were randomly assigned to these groups. The independent variables were: L2 proficiency level (higher *vs.* lower), think-aloud (present *vs.* absent), and language of text (Spanish/English *vs.* English/English). The dependent variable was the cluster score on the WMLS-R (Woodcock et al., 2005) for both the Spanish and the English form. As evident in Table 6, there was a significant main effect of L2 language proficiency. However, none of the other main effects or interactions was statistically significant. This verifies that the experimental groups did not differ in their language skills except for L2 proficiency prior to treatment.

Effects on Reading Comprehension Among Students With Lower L2 Proficiency

The question of primary interest was whether language of text would influence students' reading comprehension. Three ANOVAs were computed with two between-subjects factors: think-aloud (present *vs.* absent), and language of text (Spanish/English *vs.* English/English) for the sample of students with lower L2 proficiency. Raw scores on the three reading comprehension measures served as the dependent variables. Levene's tests of homogeneity of variance on each of the reading comprehension measures were not statistically significant ($p > .05$), indicating that the assumption of homogeneity of variance was tenable for these data. Table 7 shows the mean raw scores and standard deviations for each of the experimental groups on the reading comprehension measures. It is important to note that there were no floor or ceiling effects limiting the observation of performance differences on these tests. ANOVA test statistics and partial-eta squared estimates are shown in Table 8.

Table 6

Results of MANOVA for the Component Tests of the Woodcock-Muñoz Language Survey-Revised as a Function of English Proficiency (CALP), Think-Aloud Procedure (TA), and Language of Text (L) for Students Prior to Treatment

<i>WMLS-R English Form</i>			
Source	<i>df</i>	<i>MS</i>	<i>F</i>
CALP	1	7182.05	188.64**
TA	1	2.45	.06ns
L	1	31.25	.82ns
CALP x TA	1	4.05	.11ns
CALP x L	1	.05	.00ns
TA x L	1	48.05	1.26ns
CALP x TA x L	1	14.45	.38ns
Error	72	38.07	

<i>WMLS-R Spanish Form</i>			
Source	<i>df</i>	<i>MS</i>	<i>F</i>
CALP	1	259.20	2.12ns
TA	1	7.20	.06ns
L	1	.05	.00ns
CALP x TA	1	22.05	.18ns
CALP x L	1	80.00	.65ns
TA x L	1	320.00	2.61ns
CALP x TA x L	1	2.45	.02ns
Error	72	122.38	

N = 80. ** *p* < .01. *ns* = not significant.

Table 7

Mean Performance (and Standard Deviation) on the Three Measures of Reading Comprehension by the Four Experimental Groups on Each Think-Aloud Condition and Overall

<i>Comprehension Measure</i>	Think-Aloud			
	Lower English Proficiency		Higher English Proficiency	
	S/E	E/E	S/E	E/E
	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>
Recall	13.75 (3.51)	9.85 (1.51)	13.20 (3.75)	12.75 (4.05)
Q&A	6.90 (1.07)	5.00 (1.01)	6.85 (1.40)	6.65 (1.23)
Cloze	26.80 (6.12)	27.20 (2.25)	31.40 (4.55)	33.20 (3.71)
<i>Comprehension Measure</i>	No Think-Aloud			
	Lower English Proficiency		Higher English Proficiency	
	S/E	E/E	S/E	E/E
	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>
Recall	12.00 (2.25)	10.25 (2.67)	13.15 (3.55)	13.35 (3.38)
Q&A	6.23 (1.30)	5.95 (.97)	6.93 (1.56)	6.83 (.87)
Cloze	29.30 (4.45)	30.50 (3.95)	32.60 (4.30)	32.70 (4.37)
<i>Comprehension Measure</i>	Overall			
	Lower English Proficiency		Higher English Proficiency	
	S/E	E/E	S/E	E/E
	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>	<i>Mean (SD)</i>
Recall	12.88 (3.01)	10.05 (2.12)	13.18 (3.56)	13.05 (3.65)
Q&A	6.56 (1.21)	5.48 (1.08)	6.89 (1.44)	6.74 (1.04)
Cloze	28.05 (5.37)	28.85 (3.56)	32.00 (4.35)	32.95 (3.95)

Note. Recall maximum score = 24. Q&A maximum score = 9. Cloze maximum score = 43.

S/E = Spanish/English; E/E = English/English. *N* = 10 students in each of the eight TA groups.

Table 8

Results of ANOVAs and Partial Eta-Squared Estimates for the Three Measures of Reading Comprehension as a Function of Think-Aloud Procedure (TA) and Language of Text (L) for Students With Lower English Proficiency

Comprehension Measure and ANOVA Tests	<i>df</i>	<i>MS</i>	<i>F</i>	Partial η^2
<i>Recall</i>				
L	1	79.81	11.91**	.25
TA	1	4.56	.68 <i>ns</i>	.02
L x TA	1	11.56	1.72 <i>ns</i>	.05
Error	36	6.70		
<i>Q&A</i>				
L	1	11.83	9.86**	.22
TA	1	.19	.16 <i>ns</i>	.00
L x TA	1	6.60	5.50*	.13
Error	36	1.20		
<i>Cloze</i>				
L	1	6.40	.33 <i>ns</i>	.01
TA	1	84.10	4.31*	.11
L x TA	1	1.60	.08 <i>ns</i>	.00
Error	36	19.49		

N = 40. ** *p* < .01. * *p* < .05. *ns* = not significant.

Performance on the recall task. The first assessment measure used to examine the effect of L1 input on L2 reading comprehension was a recall task. The hypothesis was that students who read the text in both languages would perform better than students who read the text only in English. Students' oral retellings were scored based on the completeness of the idea units reported in English. The maximum score was 24. The independent variables were think-aloud (present *vs.* absent) and language of text (Spanish/English *vs.* English/English). Mean performance is reported in Table 7.

The 2 (TA *vs.* No TA) x 2 (Spanish/English text *vs.* English/English text) ANOVA revealed a statistically significant main effect for language of text. Neither the main effect for TA nor the interaction between TA and language of text was statistically significant (see Table 8). Language of text accounted for 25% of the variance on the recall scores. Mean performance and confidence intervals on the recall task are shown in Figure 2. On average, students who read the text in both languages had significantly more accurate retellings than students who read the text only in English. Mean performance was comparable across TA and No TA conditions.

Performance on Q&A task. The study was also intended to examine the effect of L1 input on L2 reading comprehension as measured by a Q&A task. The hypothesis was that students who read the text in both languages would perform better than students who read the text only in English. Students responded orally in short answers. The maximum score was 9. The independent variables were think-aloud (present *vs.* absent) and language of text (Spanish/English *vs.* English/English). Mean performance is reported in Table 7.

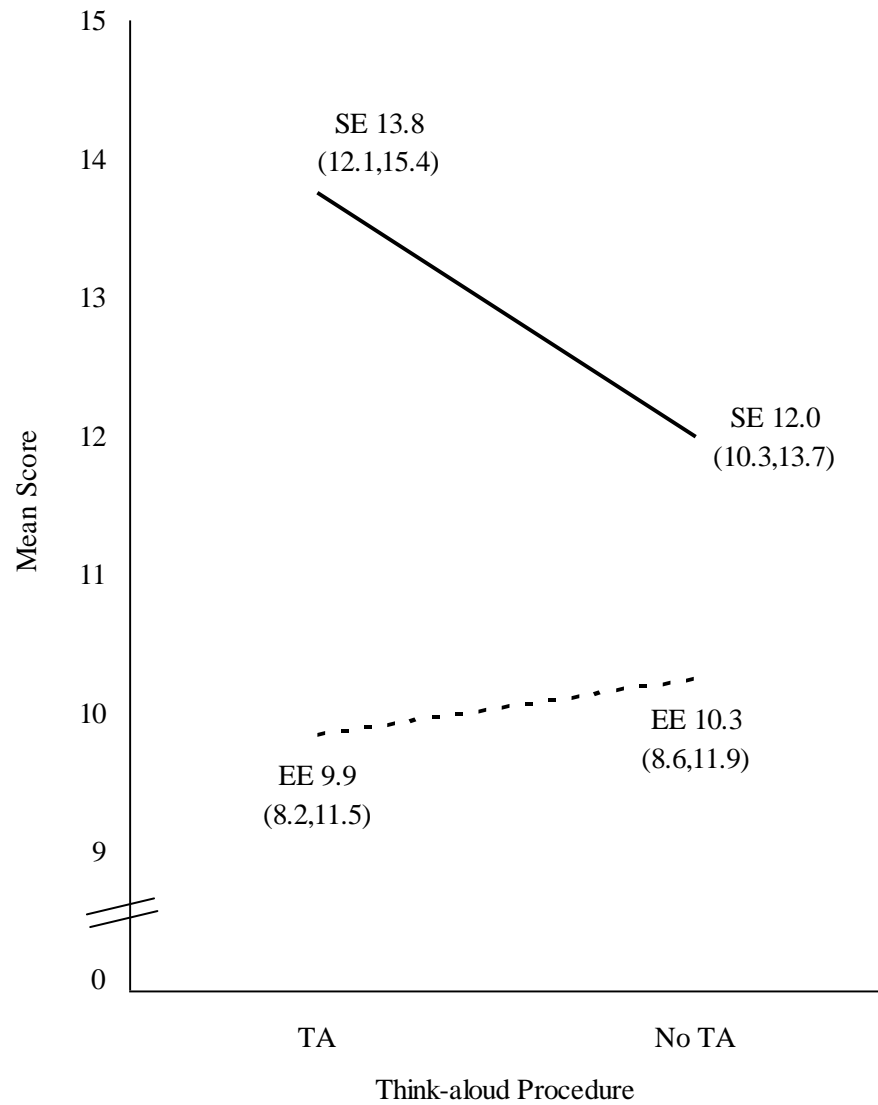


Figure 2. Mean Scores (and 95% Confidence Intervals) on the Recall Comprehension Test by Students With Lower English Proficiency as a Function of Think-Aloud and Language of the Text.

The 2 (TA *vs.* No TA) x 2 (Spanish/English text [SE] *vs.* English/English text [EE]) ANOVA revealed a statistically significant main effect for language of text. The main effect for TA was not significant. The interaction effect between think-aloud and language of text was statistically significant (see Table 8). Mean performance and confidence intervals on the Q&A task are shown in Figure 3.

Post-hoc comparisons using the Tukey HSD test indicated that the mean performance for students who read the text in both languages and performed the think aloud was significantly greater than the mean performance for students who read the text only in English and performed the think aloud. Mean performance between other pairs of means (i.e., No TA EE *vs.* No TA SE, TA EE *vs.* No TA EE, and TA SE *vs.* No TA SE) did not differ significantly from each other.

Language of text accounted for 22% of the variance on the Q&A scores.

Performance on the cloze task. Finally, the effect of L1 input on L2 reading comprehension was examined using a cloze task. The hypothesis was that students who read the text in both languages would perform better than students who read the text only in English. The maximum score was 43. The independent variables were think-aloud (present *vs.* absent) and language of text (Spanish/English *vs.* English/English). Mean performance is reported in Table 7.

According to the 2 (TA *vs.* No TA) x 2 (Spanish/English text *vs.* English/English text) ANOVA results, the main effect for language of text was not statistically significant, but the main effect for TA was statistically significant. The interaction effect was not statistically significant (see Table 8). TA accounted for 11% of the variance on the cloze scores. Mean performance and confidence intervals on the cloze task are shown

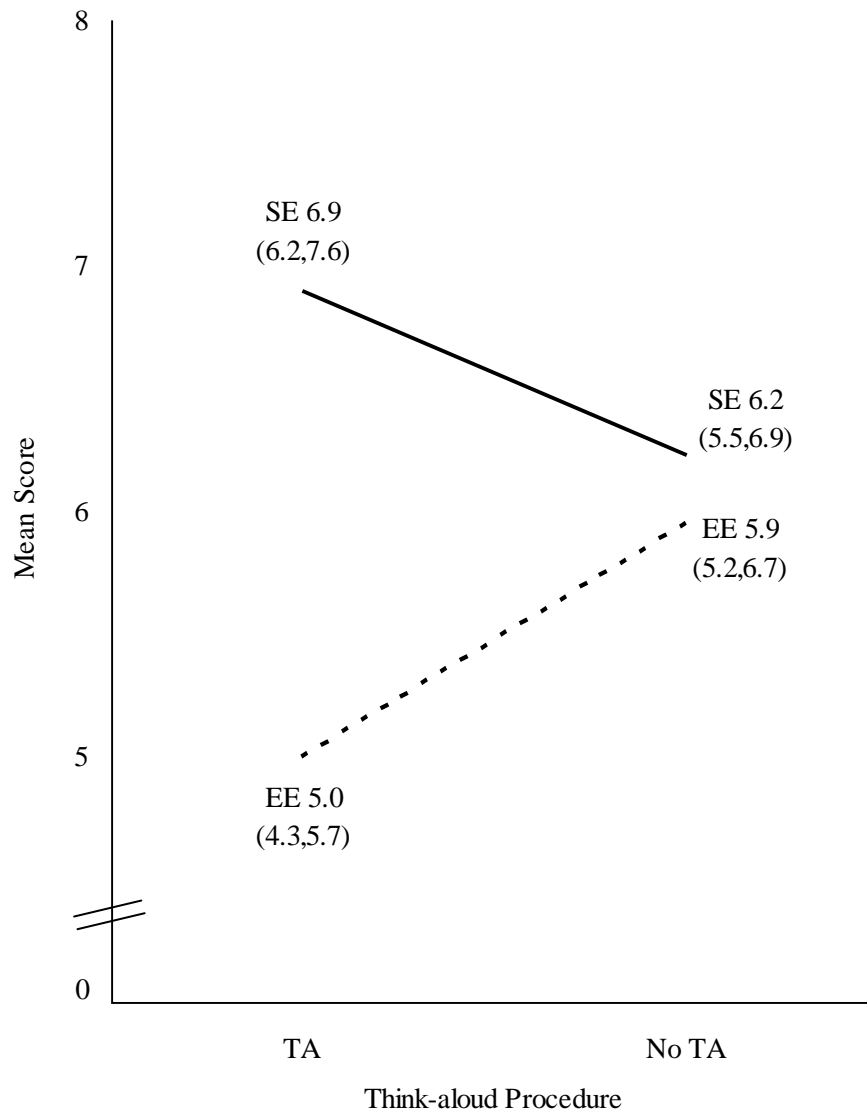


Figure 3. Mean Scores (and 95% Confidence Intervals) on the Q&A Comprehension Test by Students With Lower English Proficiency as a Function of Think-Aloud and Language of the Text.

in Figure 4. On average, students who did not perform the think-aloud procedure ($M = 29.9$, $SD = 4.14$) scored significantly better on the cloze task compared to students who did perform the think-aloud procedure ($M = 27.0$, $SD = 4.49$). Mean performance was comparable across Spanish/English and English/English reading conditions.

Effects on Reading Comprehension Among Students With Higher L2 Proficiency

The question of primary interest was whether language of text would influence students' reading comprehension. Three ANOVAs were computed with two between-subjects factors: think-aloud (present *vs.* absent), and language of text (Spanish/English *vs.* English/English) for the sample of students with higher L2 proficiency. Raw scores on the three reading comprehension measures served as the dependent variables. Levene's tests of homogeneity of variance on each of the reading comprehension measures were not statistically significant ($p > .05$), indicating that the assumption of homogeneity of variance was tenable for these data. Table 7 shows the mean raw scores and standard deviations for each of the experimental groups on the reading comprehension measures. ANOVA test statistics and partial-eta squared estimates are shown in Table 9.

Performance on the recall task. As with the sample of students with lower L2 proficiency, it was expected that students who read the text in both languages would perform better than students who read the text only in English. The independent variables were think-aloud (present *vs.* absent) and language of text (Spanish/English *vs.* English/English). Mean performance is reported in Table 7.

The 2 (TA *vs.* No TA) x 2 (Spanish/English text *vs.* English/English text) ANOVA did not reveal a statistically significant main effect for language of text or for

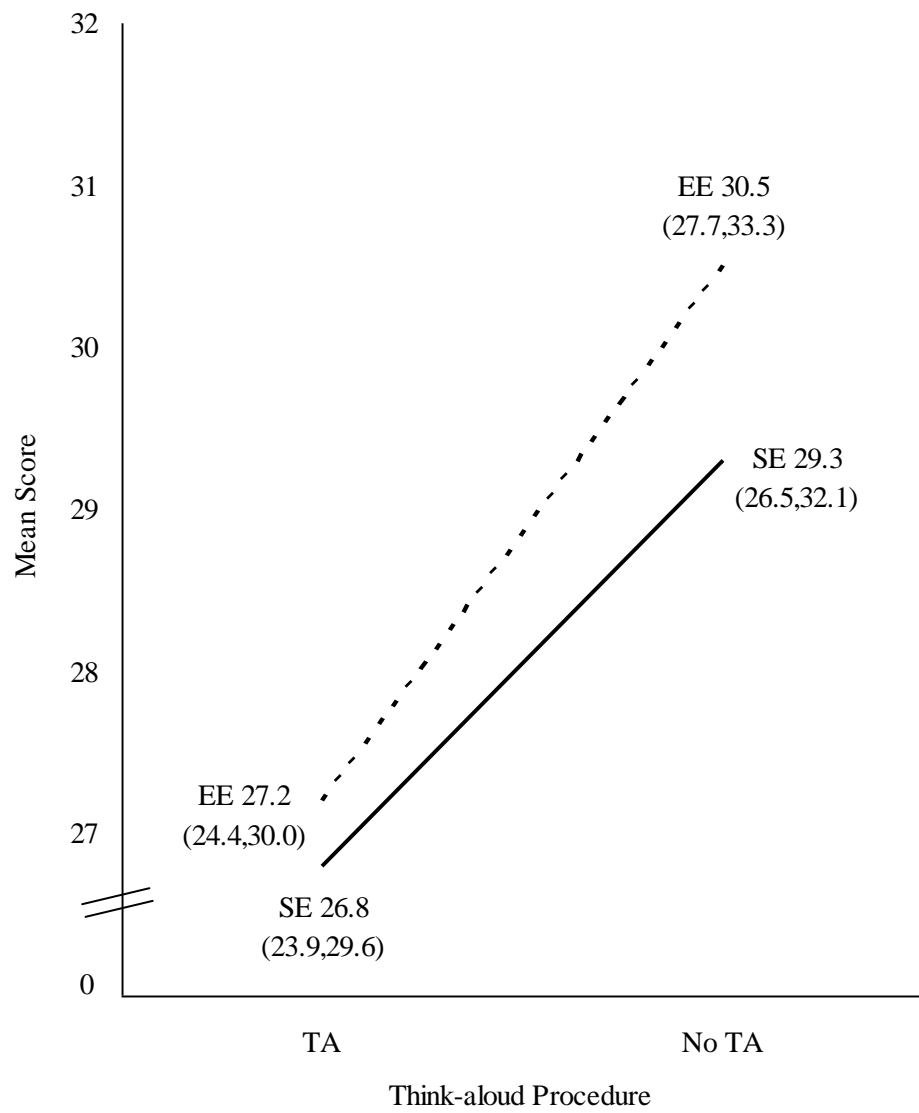


Figure 4. Mean Scores (and 95% Confidence Intervals) on the Cloze Comprehension Test by Students With Lower English Proficiency as a Function of Think-Aloud and Language of the Text.

Table 9

Results of ANOVAs and Partial Eta-Squared Estimates for the Three Measures of Reading Comprehension as a Function of Think-Aloud Procedure (TA) and Language of Text (L) for Students With Higher English Proficiency

Comprehension Measure and ANOVA Tests	<i>df</i>	<i>MS</i>	<i>F</i>	Partial η^2
<u><i>Recall</i></u>				
L	1	.16	.01 <i>ns</i>	.00
TA	1	.76	.06 <i>ns</i>	.00
L x TA	1	1.06	.08 <i>ns</i>	.00
Error	36	13.63		
<u><i>Q&A</i></u>				
L	1	.23	.13 <i>ns</i>	.00
TA	1	.16	.09 <i>ns</i>	.00
L x TA	1	.03	.01 <i>ns</i>	.00
Error	36	1.67		
<u><i>Cloze</i></u>				
L	1	9.03	.50 <i>ns</i>	.01
TA	1	1.23	.07 <i>ns</i>	.00
L x TA	1	7.23	.40 <i>ns</i>	.01
Error	36	18.01		

N = 40. *ns* = not significant.

TA. The interaction effect was also not statistically significant (see Table 9). These findings indicate that reading the text in both Spanish and English did not provide the added advantage over reading the text twice in English for students with higher L2 proficiency as it did for students with lower L2 proficiency.

Performance on Q&A task. The hypothesis was that students who read the text in both languages would perform better than students who read the text only in English. The independent variables were think-aloud (present vs. absent) and language of text (Spanish/English vs. English/English). Mean performance is reported in Table 7. The 2 (TA vs. No TA) x 2 (Spanish/English text vs. English/English text) ANOVA did not reveal a statistically significant main effect for language of text or for TA. The interaction effect was also not statistically significant (see Table 9). On average, students who read the text in both languages performed similarly to students who read the text twice in English, suggesting that L1 input did not provide an added advantage over reading the text twice in L2 for students with higher L2 proficiency.

Performance on the cloze task. The hypothesis was that students who read the text in both languages would perform better than students who read the text only in English. The independent variables were think-aloud (present vs. absent) and language of text (Spanish/English vs. English/English). Mean performance is reported in Table 7. According to the 2 (TA vs. No TA) x 2 (Spanish/English text vs. English/English text) ANOVA results, neither the main effect for language of text nor the main effect for TA was statistically significant. The interaction effect was also not statistically significant (see Table 9). On average, mean performance on the cloze task was comparable across Spanish/English and English/English reading conditions, suggesting that L1 input did not

provide an added advantage over reading the text twice in L2 for students with higher L2 proficiency.

Comparison of Effect Sizes Between English Language Proficiency Groups

The study was intended to examine whether level of English proficiency and language of text would influence students' reading comprehension. In order to compare the magnitude of effect of language of text on the two L2 proficiency groups, effect sizes were calculated for each of the comprehension measures. The effect size was determined by subtracting the SE group mean from the EE group mean and dividing the difference by the pooled standard deviation for each reading comprehension measure obtained by each L2 proficiency level. Means and standard deviations used to calculate effect sizes appear in Table 7. Effect sizes are considered to be small when $d = .20$, medium when $d = .50$, and large when $d = .80$ (Cohen, 1988).

Research on the effect of language of text on reading comprehension has reported a wide range of effect sizes. For instance, Jarvis and Jensen (1982) reported an effect size of .39 on passages with easy content and .43 on passages of difficult content when comparing performance on cloze tests for students who read L2 texts with occasional side glosses compared to students who read L2 texts with parallel translations, favoring the groups who read the text in both L1 and L2. Parent and Belasco (1970) reported an effect size of .91 when comparing average performance on weekly multiple-choice tests when reading passages only in L2 versus L1-L2, and an effect size of 1.08 when comparing performance when reading passages only in L2 versus listening to L2 recordings of the passages along with the text in L1-L2. In both instances, students performed better when parallel text translations were available. More generally, in a review of intervention

studies, Edmonds et al. (2009) reported that when compared to remedial reading instruction (i.e., instruction as it occurred typically in school) effects of reading interventions targeted at increasing comprehension (e.g., strategy training) among older struggling readers showed a weighted average effect size of 1.23, in favor of strategy training. The average effect size for the effect of different interventions on reading comprehension (i.e., word study, fluency, multi-component, and comprehension) was .89, also favoring specific interventions over typical reading instruction. These results indicate that a variety of interventions targeted at increasing reading comprehension have been quite effective, including providing students with bilingual reading materials.

In the present study, it was found that students with lower L2 proficiency who read the text in both languages scored significantly higher than students in the English-only condition on the recall task, $d = 1.10$. That is, students who read the text in both languages scored, on average, more than one standard deviation higher than students who read the text only in English on the recall task. This is quite a large effect. In contrast, among students with higher L2 proficiency, the language of the text had little impact on their recall of the text, $d = .03$. On the Q&A task, it was found that in the think-aloud condition, students with lower L2 proficiency who read the text in both languages scored significantly higher than students in the English-only condition, $d = 1.82$. In contrast, the language of the text had little impact among students with higher L2 proficiency in the TA condition, $d = .15$. Thus, both the recall and Q&A measures revealed that L1 input provided a substantial benefit on reading comprehension for students with lower English proficiency. However, among students with higher English proficiency, the benefit of reading the text in both languages was slight.

On the cloze task, it was found that the performance of students with lower L2 proficiency who read the text in both languages was similar to the performance of students in the English-only condition, $d = -.18$. Likewise, mean performance among students with higher L2 proficiency was comparable across Spanish/English and English/English reading conditions, $d = -.23$. In fact, effect sizes were negative, indicating that students were slightly better able to recall the deleted words when they had read the text twice in English than when they had read the English version only after reading the Spanish version.

Comparison of Performance Across Comprehension Measures

In addition to the main research question of whether L1 input had an effect on L2 reading comprehension, another question of interest was whether students would perform differently across the different reading comprehension tasks, specifically, whether students' performance would improve as the cognitive demands of the task decreased. The task considered the most cognitively demanding was the recall task because unprompted, extended language production was required. In contrast, the other two tasks were considered less demanding because they provided prompts in the form of questions or the text itself targeting the recall of specific information and words.

In order to compare performance across tasks, the raw scores were converted to percentages. A 2 x 3 repeated measures ANOVA was computed to investigate whether students performed differently as a function of test type. The between-subjects factor was L2 proficiency (higher *vs.* lower), and the within-subjects factor was test type (Recall *vs.* Q&A *vs.* Cloze). The dependent variable was the percent of correct responses on each reading comprehension measure.

The ANOVA revealed a statistically significant main effect for L2 language proficiency, $F(1, 78) = 12.56, p < .001$. Students with higher L2 proficiency outperformed students with lower L2 proficiency. In addition, the main effect for test type was significant, $F(2, 77) = 152.11, p < .001$. Bonferroni pairwise comparisons indicated that performance on the recall task was significantly poorer compared to performance on the Q&A and cloze tasks, whereas performance on the Q&A task did not differ significantly from performance on the cloze task.

Figure 5 shows the mean percent score on each reading comprehension measure by language of the text for students with lower English proficiency. On average, students who read the text twice in English recalled about 42% of the ideas expressed in the text, answered about 61% of questions accurately, and completed about 67% of the blanks on the cloze task correctly. In contrast, students who read the text in both languages recalled about 54% of the ideas expressed in the text, answered about 73% of questions accurately, and completed about 65% of the blanks on the cloze task correctly. Figure 6 shows the mean percent score on each reading comprehension measure by language of the text for students with higher English proficiency. On average, students recalled about 55% of the ideas expressed in the text, and correctly completed about 76% of items on the Q&A and cloze tasks.

Taken together, results show that different assessment tasks provided a different picture as to the level of understanding students attained. Overall, performance on the recall task was significantly lower among students with both lower and higher L2 proficiency. It was not surprising that performance on the Q&A task was higher in

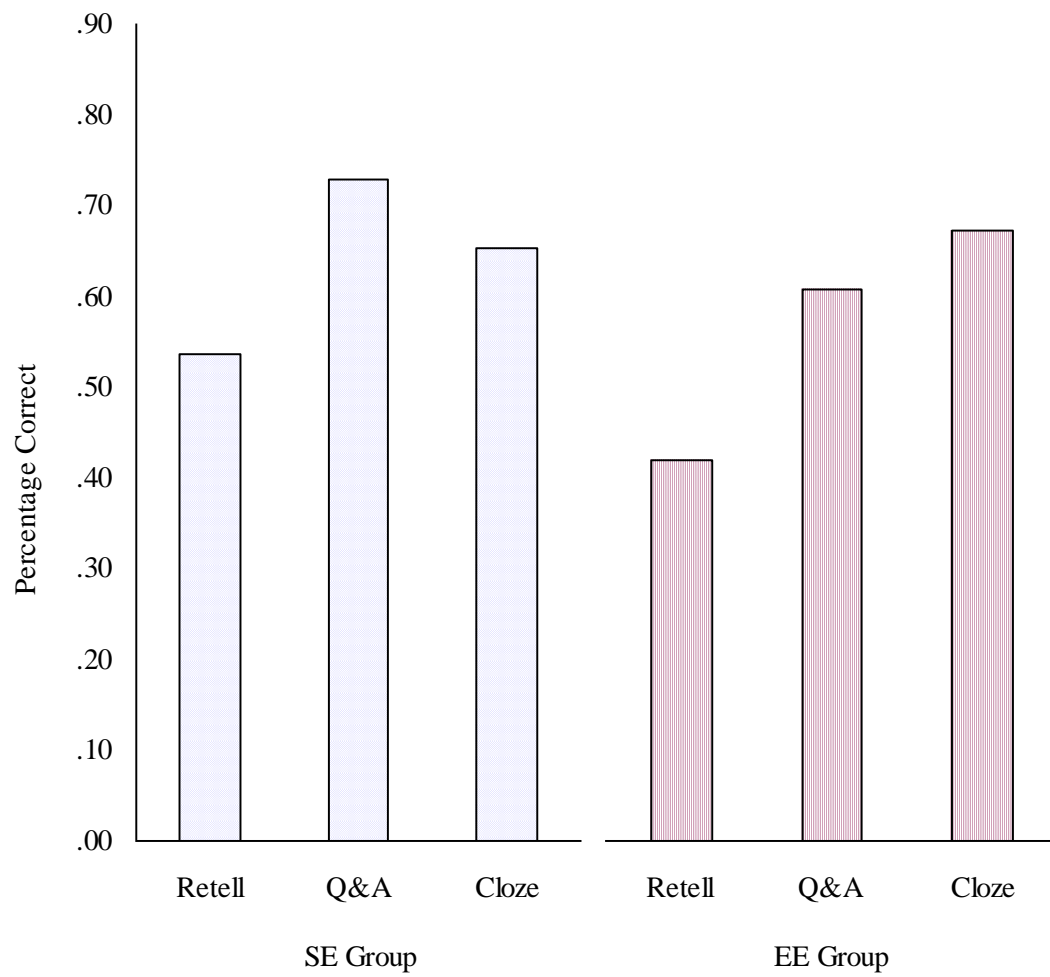


Figure 5. Mean Percent Scores on the Three Reading Comprehension Measures by Students With Lower English Proficiency Who Read the Text in Spanish and English (SE) or Twice in English (EE).

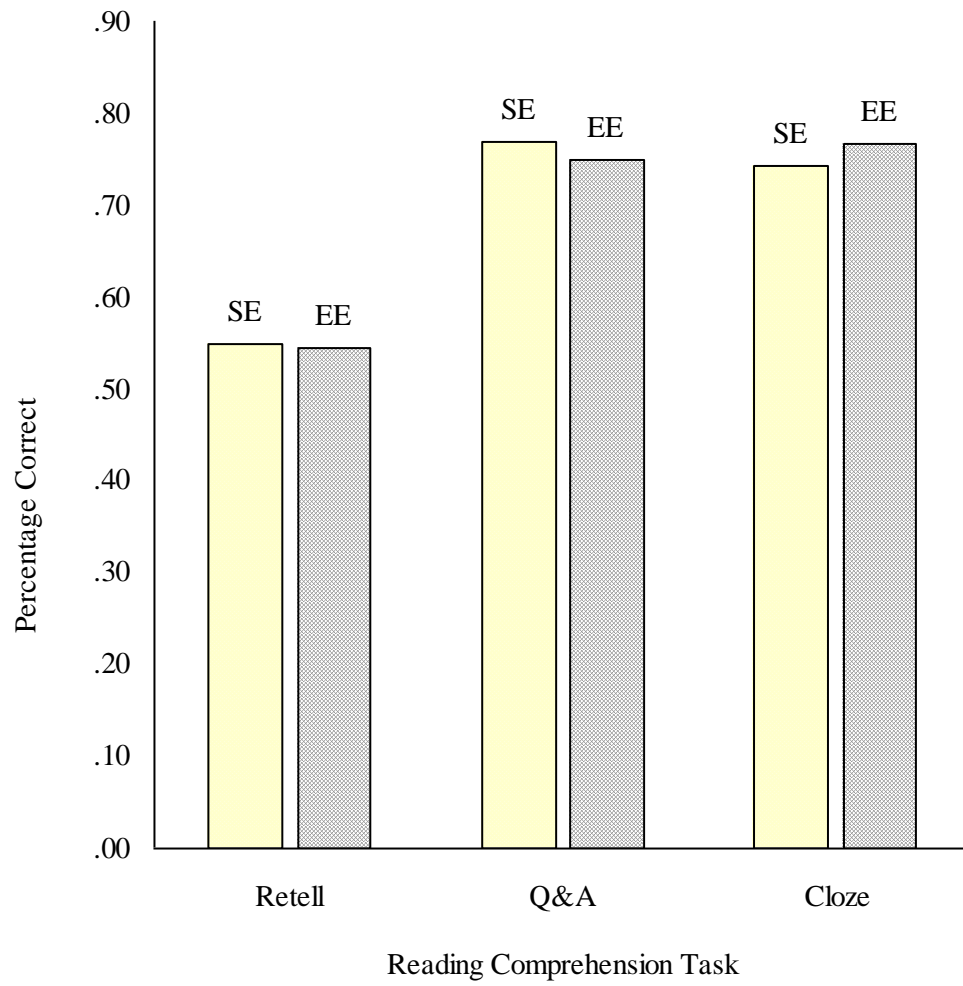


Figure 6. Mean Percent Scores on the Three Reading Comprehension Measures by Students With Higher English Proficiency Who Read the Text in Spanish and English (SE) or Twice in English (EE).

comparison with the recall task given that the Q&A task demanded less in terms of language production and memory.

Correlations and Regression Analyses

In order to assess relationships between performance on the reading comprehension measures and students' L2 and L1 proficiency, Pearson product-moment correlation coefficients were computed. Given the pronounced effects of language of text on some of the reading comprehension measures, *post hoc* exploratory correlational analyses were conducted in two ways. First, correlations were computed for the total sample without regard to the effects of language of text. These results are presented in Table 10. The second analysis explored whether the relationships between students' L2 and L1 proficiency and performance on the reading comprehension measures changed as a function of the language of text read. The results for the correlations qualified by language of text are presented in Table 11.

The results for the uncontrolled correlations, which are presented in Table 10, revealed statistically significant, positive correlations among the scores on the various L2 and L1 WMLS-R (Woodcock et al., 2005) measures of language proficiency. The component tests, which were the same in L1 and L2, were the following: 1. Picture Vocabulary; 2. Verbal Analogies; 3. Letter-Word Identification; and 4. Dictation. The within-language L2 correlations were significant, ranging from .22 to .48. Similar to those shown for the L2 tests, within-language correlations among the L1 tests were small to moderate, ranging from .21 to .50. These results were expected. Interestingly, cross-language correlations for analogous skills, with the exception of Picture Vocabulary, were also positive and statistically significant. Performance on the Verbal Analogies test in L1 and L2 correlated .40; Word Identification in L2 correlated .31 with its counterpart

Table 10

Correlations Between the Woodcock-Muñoz Language Survey-Revised Pretest Measures in English and Spanish and the Three Measures of Reading Comprehension

	L2-T1	L2-T2	L2-T3	L2-T4	L2 WC	L1-T1	L1-T2	L1-T3	L1-T4	L1 WC	Retell	Q&A	Cloze
L2-T1	1												
L2-T2	.48**	1											
L2-T3	.38**	.22*	1										
L2-T4	.38**	.22*	.24*	1									
L2 WC	.79**	.62**	.71**	.64**	1								
L1-T1	.01	.18	.06	.17	.13	1							
L1-T2	-.01	.40**	-.06	.22*	.14	.50**	1						
L1-T3	-.02	.07	.31**	.14	.19	.26*	.21	1					
L1-T4	-.16	.13	-.01	.32**	.07	.45**	.42**	.33**	1				
L1 WC	-.07	.28*	.10	.30**	.18	.73**	.75**	.64**	.77**	1			
Retell	.27*	.33**	.21	.21	.35**	.29**	.31**	.14	.11	.29**	1		
Q&A	.33**	.31**	.09	.31**	.36**	.27*	.31**	.00	.29*	.30**	.64**	1	
Cloze	.47**	.42**	.22*	.33**	.51**	.31**	.38**	.09	.10	.30**	.59**	.56**	1

N = 80. * p < .05. ** p < .01.

L2-T1 = Picture Vocabulary test WMLS-R, English
 L2-T2 = Verbal Analogies test WMLS-R, English
 L2-T3 = Word Identification test WMLS-R, English
 L2-T4 = Dictation test WMLS-R, English
 L2 WC = Cluster score on the WMLS-R, English

L1-T1 = Picture Vocabulary test WMLS-R, Spanish
 L1-T2 = Verbal Analogies test WMLS-R, Spanish
 L1-T3 = Word Identification test WMLS-R, Spanish
 L1-T4 = Dictation test WMLS-R, Spanish
 L1 WC = Cluster score on the WMLS-R, Spanish

Table 11

Correlations Between the Woodcock-Muñoz Language Survey-Revised Pretest Measures in English and Spanish and the Three Measures of Reading Comprehension as a Function of Language of Text

	L2-T1	L2-T2	L2-T3	L2-T4	L2 WC	L1-T1	L1-T2	L1-T3	L1-T4	L1 WC	Retell	Q&A	Cloze
L2-T1	1	.37*	.32*	.44**	.79**	-.04	-.05	.07	-.10	-.04	.30	.46**	.42**
L2-T2	.61**	1	.07	.41**	.57**	.26	.56**	.28	.28	.51**	.44**	.46**	.52**
L2-T3	.42**	.39*	1	.29	.65**	.15	-.22	.34*	-.13	.03	.20	.10	.19
L2-T4	.09	.15	.20	1	.77**	.26	.28	.23	.41**	.43**	.33*	.39*	.41**
L2 WC	.79*	.69**	.76**	.51**	1	.20	.13	.32*	.13	.28	.43**	.49**	.52**
L1-T1	.04	.12	-.01	.09	.07	1	.23	.30	.31	.61**	.18	-.03	.19
L1-T2	.03	.29	.09	.17	.17	.73**	1	.25	.42**	.72**	.22	.17	.25
L1-T3	-.10	-.20	.29	.03	.05	.23	.19	1	.24	.66**	.08	-.05	.06
L1-T4	-.23	-.10	.11	.22	.00	.61**	.45**	.43**	1	.75**	.08	.36*	.15
L1 WC	-.09	.04	.16	.17	.10	.83**	.79**	.62**	.81**	1	.20	.19	.24
Retell	.25	.14	.23	.08	.27	.43**	.48**	.18	.12	.40*	1	.54**	.71**
Q&A	.23	.11	.08	.23	.24	.57**	.53**	.01	.20	.42**	.69**	1	.59**
Cloze	.51**	.41**	.25	.28	.52**	.39*	.47**	.12	.08	.35*	.58**	.61**	1

Note. Correlations above the diagonal display results for the group who read the text twice in English. Correlations below the diagonal display results for the group who read the text in Spanish and English. *N* = 40/per group. * *p* < .05. ** *p* < .01.

.2-T1 = Picture Vocabulary test WMLS-R, English
 .2-T2 = Verbal Analogies test WMLS-R, English
 .2-T3 = Word Identification test WMLS-R, English
 .2-T4 = Dictation test WMLS-R, English
 .2 WC = Cluster score on the WMLS-R, English

L1-T1 = Picture Vocabulary test WMLS-R, Spanish
 L1-T2 = Verbal Analogies test WMLS-R, Spanish
 L1-T3 = Word Identification test WMLS-R, Spanish
 L1-T4 = Dictation test WMLS-R, Spanish
 L1 WC = Cluster score on the WMLS-R, Spanish

in L1; and scores on the Dictation test in L2 correlated .32 with Dictation scores in L1. Although modest, these correlations provided further support for research indicative of cross-linguistic relationships in first- and second-language domains related to literacy.

Table 10 also shows that correlations between reading comprehension measures were positive and statistically significant. The correlations were .64 between the recall task and the Q&A task; .59 between the recall task and the cloze task; and .56 between the Q&A task and the cloze task. Although the correlations were strong, they were not high enough to suggest they were measuring the same capability. It is possible that the tasks may have assessed somewhat different aspects of reading comprehension because they placed different demands on the various knowledge sources and skills necessary for reading with comprehension.

Given the pronounced effects of language of text on some of the reading comprehension measures in the ANOVAs, it was important to investigate whether the relationships between students' proficiency on the WMLS-R (Woodcock et al., 2005) L1 and L2 tests and performance on the reading comprehension measures differed as a function of the language of text read. Perhaps it was possible that L1 proficiency might exert a greater impact when the text was read in both L1 and L2 than when it was read only in L2. Table 11 displays correlations for the group who read the text twice in English (EE) above the diagonal, and those for the group who read the text in Spanish and English (SE) below the diagonal. Correlations between reading comprehension measures for both reading groups were comparable, ranging from .58 to .69 within the SE group and from .54 to .71 within the EE group. These correlations were sufficiently strong, but not high enough as to suggest they were measuring the same capability.

Correlations between the language proficiency skills and measures of reading comprehension differed by reading group. It is evident in Table 11 that more L1 skills were significantly related to performance on the reading comprehension measures for the SE group compared to the EE group (i.e., 6 vs. 1 significant correlations). In contrast, several more L2 skills were significantly related to performance on the reading comprehension measures for the EE group than for the SE group (i.e., 8 vs. 2 significant correlations).

In order to determine whether proficiency in both L1 and L2 were central in explaining performance on the various reading comprehension measures, regression analyses were conducted. As mentioned before, pronounced effects for language of text were found on some of the reading comprehension measures in the ANOVAs; therefore, the regression analyses were conducted for the sample as a whole and repeated for each reading group separately. To investigate the relative contributions, unique and shared, of WMLS-R L1 and L2 language proficiency to reading comprehension measures, standard multiple regression analyses were conducted. Although it would have been preferable to compare the relative contribution of individual WMLS-R component tests to reading comprehension overall and by reading group, it was not appropriate to perform such analyses due to the potential of a *multiple testing problem* (i.e., Type I error) given the limitation imposed by the sample size. Instead, the WMLS-R L2 cluster score was entered first and the WMLS-R L1 cluster score was entered second as independent variables. For each of the analyses, the following major assumptions about the data were met: a) there was no multicollinearity or singularity; b) there were no outliers; c) the residuals were normally distributed about the predicted scores; d) the residuals had a

straight-line relationship with the predicted scores; and e) the variances of the residuals about the predicted scores were homoscedastic. Table 12 shows the test statistics for the WMLS-R L1 and L2 proficiency measures. Their total, shared, and unique contributions are illustrated in Figure 7 for the overall sample and in Figure 8 as a function of language of text.

When comparing the total amount of variance explained by the WMLS-R L1 and L2 proficiency tests across the various reading comprehension measures, is evident in Table 12 that in all but one case the total amount of explained variance was greater when the groups were considered separately by language of text read than when they were combined (i.e., explained variance ranged from 19% to 36% separately versus 18% to 31% combined). Therefore, it was reasonable to analyze the contributions of L1 and L2 language competencies separately by reading group. As revealed in Table 12, among students who read the text in both languages (SE), skills related to language proficiency in L1 reliably explained a significant amount of variance in scores on all reading comprehension measures, whereas L2 skills only explained a significant amount of variance on cloze performance. In contrast, among students who read the text only in English (EE), L1 skills did not contribute significantly to performance (see Figure 8).

Within the SE group, L1 and L2 proficiency in combination accounted for 21% of the variance on the recall task, 22% of the variance on the Q&A task, and 36% of the variance on the cloze task. Between these two, however, only L1 proficiency made a statistically significant contribution to performance on the recall and Q&A tasks as indicated by the squared semi-partial correlations (i.e., unique variance = 14% and 16%, respectively). On the other hand, both L1 and L2 proficiency accounted for a statistically

Table 12

Variance Explained by Woodcock-Muñoz Language Survey-Revised (WM) Measures in English (L2) and Spanish (L1) on the Three Measures of Reading Comprehension for the Overall Sample and as a Function of Language of Text

<i>Predictor</i>	<u>Recall</u>			<u>Overall</u>			<u>Cloze</u>		
	R^2	$R^2 +$	t	R^2	$R^2 +$	t	R^2	$R^2 +$	t
WM L2	.12	.09	2.92**	.13	.10	3.03**	.26	.22	4.88***
WM L1	.09	.06	2.23*	.09	.06	2.33*	.09	.05	2.21*
<i>Total</i>	.18			.19			.31		
<i>Shared</i>	.03			.03			.04		
<u>English-English</u>									
WM L2	.18	.15	2.63**	.24	.21	3.15**	.27	.22	3.41**
WM L1	.04	.01	.56ns	.03	.00	.37ns	.06	.01	.69ns
<i>Total</i>	.19			.24			.28		
<i>Shared</i>	.03			.03			.05		
<u>Spanish-English</u>									
WM L2	.07	.05	1.56ns	.06	.04	1.40ns	.27	.24	3.71**
WM L1	.16	.14	2.54*	.18	.16	2.74**	.12	.09	2.29*
<i>Total</i>	.21			.22			.36		
<i>Shared</i>	.02			.02			.03		

$N = 80$ (40/per language group). * $p < .05$. ** $p < .01$. *** $p < .001$. ns = not significant.

R^2 = amount of variance as indicated by the squared zero-order correlation coefficient.

$R^2 +$ = unique contribution to variance as indicated by the squared semi-partial correlation coefficient.

t = significance level for the regression coefficient at 77df (overall) and 37df (by language of text).

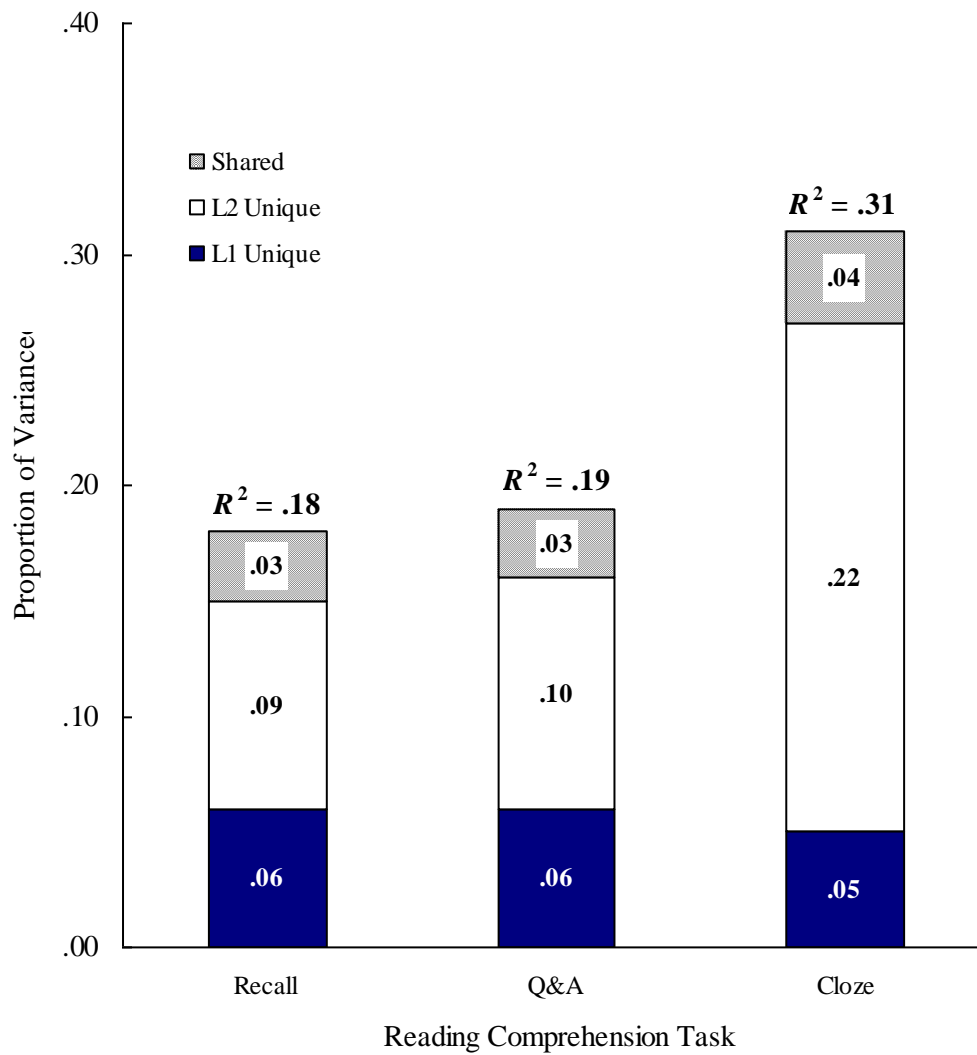


Figure 7. Decomposition of Variance Accounted for by Woodcock-Muñoz Language Survey-Revised Composite Scores in the Spanish (L1) and English (L2) Forms on the Three Measures of Reading Comprehension.

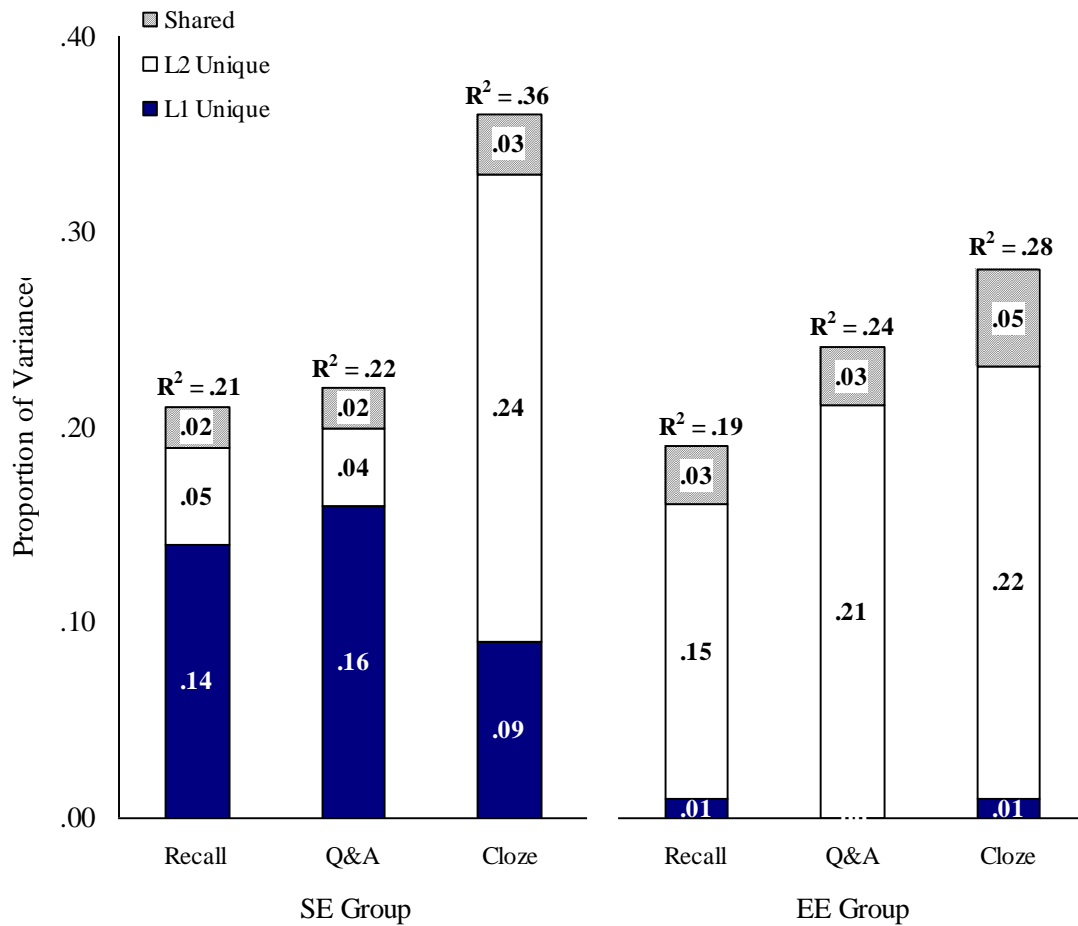


Figure 8. Decomposition of Variance Accounted for by Woodcock-Muñoz Language Survey-Revised Composite Scores in the Spanish (L1) and English (L2) Forms on the Three Measures of Reading Comprehension by Students Who Read the Text in Spanish and English (SE) or Twice in English (EE).

significant 33% of unique variance on the cloze task. Although between these two, L2 proficiency seemed more important.

Within the EE group, L1 and L2 proficiency in combination accounted for 19% of the variance on the recall task, 24% of the variance on the Q&A task, and 28% of the variance on the cloze task. Between these two, however, only L2 proficiency made a statistically significant, unique contribution on each reading comprehension measure. As indicated by the squared semi-partial correlations, L2 proficiency accounted for 15% of unique variance on the recall task, 21% of unique variance on the Q&A task, and 22% of unique variance on the cloze task.

The bivariate correlations in Table 11 suggest that for the SE group, L1 verbal analogical reasoning provided the greatest contribution on performance on the recall task, whereas for the EE group it was L2 verbal analogical reasoning. For the SE group, performance on the Q&A task was related the most to L1 expressive vocabulary, whereas for the EE group, L2 expressive vocabulary had the strongest relationship to performance on this task. Lastly, performance on the cloze task was related the most with L2 expressive vocabulary for the SE group, whereas L2 verbal analogical reasoning had the strongest relationship with performance on the cloze for the EE group.

Taken together, these findings suggest that when students were given L1 input, they drew more heavily from their L1 knowledge than their L2 knowledge to help them comprehend the text. However, when they received only L2 input, their L2 skills contributed more to their performance than their L1 skills. Performance across tasks seemed to be influenced the most by expressive vocabulary and verbal analogical reasoning, albeit in their native language or in their second language.

Online Reading Processes as Revealed by the Think-Aloud

A qualitative analysis of the strategies employed by the students and revealed through the think-aloud was conducted to investigate how L1 input facilitated the comprehension of L2 input. This analysis was exploratory. We expected that students' verbal protocols would provide a window as to how the L1 input provided explicit help to understand the L2 text. In addition, another goal was to examine the extent to which students used L1 while reading in L2, and whether text processing differed between groups as revealed by the reading strategies used. Furthermore, an investigation was made as to whether particular strategies were related in important ways to reading comprehension. The participants' verbal reports were transcribed, matched with the corresponding sentence in the text, and carefully analyzed in order to identify strategies used while reading the text. Some comments were coded as revealing more than one strategy.

To establish the reliability of the coding, an independent rater was given a predetermined list of strategies along with the definition and an example for each strategy. The rater was trained on how to interpret the list of strategies and identify them in the protocols. After training, the rater was asked to code the strategies in 10% of the protocols using the list. The rater's coding was then compared with the researcher's coding in order to determine interrater reliability. The percentage of agreement between the researcher and the independent rater was 86%. Discrepancies in coding were discussed and resolved. The following 12 strategies for identifying and interpreting information and monitoring comprehension were distinguished (see Pressley &

Afflerbach, 1995, for a discussion of research-based strategies students use before, during, and after reading):

1. *Elaborative inference-making*: providing descriptive details, making generalizations or connections about characters, actions, objects, or events in the text, and/or adding information not presented in the text in response to a particular piece of information in the focal sentence. (E.g., “The government was worried because maybe it [the torpedoes filled with mercury] could be damaging for the people on the beach.” – sentence 12)
2. *Summarizing*: recounting information and main ideas across multiple sentences (i.e., at least 3 sentences), or parts of the text.
3. *Paraphrasing*: restating the message of the focal sentence using own words or language very close to that in the text into more familiar terms. (E.g., “It is a dangerous toxic that can kill or make people crazy.” – sentence 13)
4. *Questioning word*: questioning the meaning of a difficult or unknown word. (E.g., “What does *drifted* mean? – sentence 9)
5. *Guessing word meaning*: self-interrogation and attempt to infer the probable meaning of a word not understood or its pronunciation. (E.g., “Drifted must be like *a la deriva* or *arrastrarse*.” – sentence 17)
6. *Questioning/Reacting to text*: reacting to the structure or information in the text or posing questions regarding the ideas being conveyed in the text. (E.g., “They should have not left it there.”)
7. *Visualizing*: creating a mental image of the events, characters, or situations in the text. (E.g., “I’m trying to picture in my mind what’s happening.”)

8. *Evaluating comprehension*: assessing understanding of what is being read and/or demonstrating awareness or comprehension failure or success. (E.g., “I’m not sure, but this is what I understand.”)
9. *Reread sentence*: reading the focal sentence over again after the first complete reading to clarify meaning or resolve a difficulty in comprehension.
10. *Reading on*: continuing to read the following sentence without stopping to report thoughts on the focal sentence.
11. *Translating*: restating the focal sentence directly or word-for-word in L1.
12. *Code-switching*: simultaneous use of both languages for the purpose of keeping the thought alive or flowing. (E.g., “Good idea *por que así pueden investigar.*” – sentence 3)

Table 13 shows the frequency of each strategy identified. Strategies were classified as text-based (i.e., bottom-up) or reader-based (i.e., top-down) following Hock Seng and Hashim’s (2006) taxonomy. In addition, each strategy was classified as L1 or L2, according to the language the student used to express his/her thoughts. For example, if a student said, “*No lo debieron dejar allí?*”, which in English means “*They should have not left it there?*”, this was coded as: L1 – reacting to text. Students’ comments regarding how the L1 input provided explicit help to understand the L2 text were almost non-existent; therefore, this research question was not pursued further.

The number of strategies produced by each experimental group ranged from 227 to 258. Overall, the strategies most often used were paraphrasing (24%), inferencing (20%), and reacting to/questioning the text (18%). In order to determine whether

Table 13

Number and Total Percent of Strategies Verbalized by Type for the Four Experimental Groups During the Think-Aloud Procedure

	Lower English Proficiency						Higher English Proficiency					
	Spanish/English			English/English			Spanish/English			English/English		
	<u>L1</u>	<u>L2</u>	<u>%</u>	<u>L1</u>	<u>L2</u>	<u>%</u>	<u>L1</u>	<u>L2</u>	<u>%</u>	<u>L1</u>	<u>L2</u>	<u>%</u>
<i><u>Text-based Strategies</u></i>												
Summarizing	1	0	.4%	1	0	.4%	0	0		0	0	
Paraphrasing	30	26	24%	39	46	36%	14	26	16%	41	10	22%
Questioning word	10	9	8%	21	5	11%	13	8	8%	7	8	7%
Rereading		4	2%		2	1%		10	4%		18	8%
Subtotal	41	39	34%	61	53	48%	27	44	28%	48	36	37%
<i><u>Reader-based Strategies</u></i>												
Guessing word meaning	3	4	3%	6	3	4%	8	1	3%	6		3%
Inferencing	28	17	19%	11	34	19%	20	32	20%	17	29	20%
Reacting to text	26	10	15%	12	20	14%	19	55	29%	3	20	10%
Evaluating comprehension	2	0	1%	0	0		1	0	.4%	0	0	
Visualizing	1	0	.4%	0	0		0	0		1	0	.4%
Reading on		37	16%		15	6%		20	8%		12	5%
Translating	12		5%	10		4%	15		6%	38		17%
Code-switching	13		6%	11		5%	16		6%	17		7%
Subtotal	85	68	66%	50	72	52%	79	108	72%	82	61	63%
Total Strategies	126	107	100%	111	125	100%	106	152	100%	130	97	100%

Note. L1 – used L1 in verbalization during the think-aloud. L2 – used L2 in verbalization during the think-aloud.

N = 40.

groups differed in the number of reader- and text-based strategies as well as in the number of paraphrases, three ANOVAs were performed with L2 proficiency level (lower vs. higher) and language of text (Spanish/English vs. English/English) as the independent variables. The dependent variable was the total number of text-based strategies, the total number of reader-based strategies, or the total number of paraphrases.

The results revealed a statistically significant main effect for language of text on the use of text-based strategies, $F(1, 36) = 4.99, p < .05$; reader-based strategies, $F(1, 36) = 6.43, p < .05$; and paraphrases, $F(1, 36) = 5.94, p < .05$. A main effect for L2 proficiency was not detected on the three measures of strategy use: $F(1, 36) = 1.02, p > .05$; $F(1, 36) = 3.43, p > .05$; $F(1, 36) = .02, p > .05$, respectively. The interactions between L2 proficiency and language of text were not statistically significant: $F(1, 36) = .10, p > .05$; $F(1, 36) = .17, p > .05$; $F(1, 36) = .23, p > .05$, respectively. On average, students who read the text in both languages employed significantly more reader-based strategies and fewer text-based strategies, and paraphrased what they read significantly less often than students who read the text only in English.

As mentioned before, of particular interest was the extent to which students used L1 across groups. As shown in Table 13, all four experimental groups used L1 in addition to L2 to express their thoughts during the think-aloud procedure. Overall, there was a total of 473 instances of strategy use where students used L1, accounting for 49.6% of all strategies identified in the protocols. The strategy most often used in L1 was also paraphrasing (26%), followed by inferencing (16.1%), which was on a par with translating (15.9%), and next, reacting to text (13%). Inspection of Table 13 shows that across groups, students who read the text only in L2 used translation and L1 paraphrases,

which closely resembled translations, more often than students who read the text in both languages. In contrast, students who read the text in both languages went beyond paraphrasing the text and often interacted with the text by constructing elaborative inferences.

A series of ANOVAs were computed with L2 proficiency level (lower vs. higher) and language of text (Spanish/English vs. English/English) as the independent variables to test if the groups differed statistically in the frequency in which they used L1 on text-based strategies, reader-based strategies, the sum of translations and L1 paraphrases, as well as on the total number of inferences made in L1 while interacting with the text. No statistically significant differences were found: $F(1, 36) = 1.36, p > .05$ (text-based strategies); $F(1, 36) = .59, p > .05$ (reader-based strategies); $F(1, 36) = 2.03, p > .05$ (translations and L1 paraphrases); and $F(1, 36) = .01, p > .05$ (inferences).

An analysis of the bivariate correlations between strategies used and scores on the three reading comprehension measures revealed a tendency for scores on the recall task to increase as students provided more paraphrases ($r = .28, p = .08$). To the extent that students grappled with unknown vocabulary words, scores on the Q&A task tended to decrease ($r = -.27, p = .09$). In addition, it appeared that the more that students read on and failed to stop to report their thoughts, the lower their scores were on the cloze task ($r = -.35, p = .03$). Lastly, reacting to the text in L1 appeared to be detrimental for performance on all three reading comprehension measures: $r = -.33, p = .04$ on the recall task; $r = -.28, p = .08$ on the Q&A task; and $r = -.46, p = .003$ on the cloze task. This last finding suggests that the strategies “reacting to or questioning the text” and “reading on” added nothing to the comprehension process and/or created confusion instead. Some of

these relationships were not surprising, and the exit interview helped provide possible explanations for these findings.

Exit Interview

At the end of the second meeting after completing the reading comprehension tasks, students were asked nine questions in order to learn about their perceptions regarding the difficulty level of the text, their level of background knowledge on the topic, and other aspects of the tasks. This was an oral interview and students were asked to provide brief answers in their language of preference. The responses were analyzed for patterns and tallied. The results are presented in Table 14.

The exit interview revealed that none of the students had heard the story before. As expected, students were able to understand the L1 version of the text quite easily. On the other hand, whereas some students found the L2 text easy to understand, others found it more difficult. This range in difficulty was also expected because of their varied degree of L2 proficiency as revealed by the WMLS-R (Woodcock et al., 2005). Therefore, it was anticipated that most students would have encountered unknown vocabulary words. When students were asked which words were most difficult to understand, most of the words they did not understand were key words essential to understanding the overall meaning of the text. Students who read the text in both languages found the L1 version “helpful” to understanding the L2 text at both the word- and phrase-level with the exception of two students who commented that reading the text in both languages “didn’t make a difference” in understanding the text. Likewise, students who read the text twice in L2 found it “helpful” to understanding the text as the second reading served as “review” with the exception of two students who also

Table 14

Responses to the Questions on the Exit Interview (N = 80)

Question 1	<i>Was the text difficult or easy to understand?</i>					
	<i>L1 text</i>	<u>N</u>	<u>%</u>	<i>L2 text</i>	<u>N</u>	<u>%</u>
	Easy	40	100%	Easy	62	77.5%
	Medium	0	-	Medium	15	18.8%
	Hard	0	-	Hard	3	3.8%
Question 2	<i>Were you familiar with the story?</i>					
	Yes	0	-			
	No	80	100%			
Question 3	<i>Were there any words you didn't understand?</i>					
	Yes	63	78.8%			
	No	17	21.3%			
	<i>If yes, which words?</i>					
	Drifted	23	33.3%	Torpedoes	5	7.2%
	Wreck	22	31.9%	Amazed	3	4.3%
	Sunk	6	8.7%	Seep	3	4.3%
	Divers	5	7.2%	Mercury	2	2.9%
Question 4	<i>Which comprehension task was the easiest? Why?</i>					
	Recall	28	35.0%			
	Q&A	26	32.5%			
	Cloze	26	32.5%			
	<i>Which comprehension task was the hardest? Why?</i>					
	Recall	34	42.5%			
	Q&A	6	7.5%			
	Cloze	40	50.0%			
Question 7	<i>How would you have done should you have recalled the story in Spanish?</i>					
	Better	44	55.0%			
	Same	30	37.5%			
	Worse	6	7.5%			
Question 8	<i>How would you have done should you have answered the questions in Spanish?</i>					
	Better	35	43.8%			
	Same	43	53.8%			
	Worse	2	2.5%			
Question 9	<i>How would you have done should you have completed the cloze in Spanish?</i>					
	Better	37	46.3%			
	Same	29	36.3%			
	Worse	14	17.5%			

commented that reading the text twice “didn’t make a difference” in understanding the text “better.”

Hence, the negative relationship between students grappling with unknown vocabulary words and scores on the Q&A task found through the correlation analysis could be explained by the extent to which key words in the text were difficult or unknown to the students. For instance, from the list of unknown words identified by students, the word *divers* appeared in 3 of the 8 questions. The words *drifted* and *sink*, which is a variation of sank, also appeared in the questions, whereas *torpedoes* and *mercury* were part of the answer to two questions. Accordingly, there was only one question free of any bias, *who might get hurt?* This question, nonetheless, required a very specific response.

As mentioned before, there was also a moderate positive relationship between students paraphrasing and scores on the recall task. The recall task allowed more flexibility in oral expression compared to the Q&A and cloze tasks, and it was influenced by memory. One possible explanation is that because of the students’ limited oral language abilities, paraphrasing not only allowed students to interpret the text into more familiar terms, but it also served as extra practice “reading” the text, further helping recall. Although the margin was small, the recall task was also perceived to be the easiest task. The most common explanation as to why students thought the recall task was the easiest was that “you use your own words” and “you have more choice with words.” Overall, students had the lowest scores on the recall task when comparing performance across tasks. Thus there was a discrepancy between students’ perception and their performance.

The cloze task, on the other hand, was perceived to be the hardest task. The comments students often made as to why they thought the cloze task had been the hardest was that the cloze task “required a lot of memory” and “specific words.” Overall, students had similar scores on the cloze and Q&A tasks when comparing performance across tasks; hence here too a discrepancy was evident between students’ perception as to what they felt was the “hardest” task and their performance. Not surprisingly, about half of the students felt that they would have performed better in the cloze and recall tasks if they had been allowed to complete the tasks in L1.

Chapter 6: Discussion

The purpose of this study was to investigate the effect of native-language input and second-language proficiency on L2 reading comprehension. What follows is an overview of the findings followed by a discussion of the research questions. Strengths and limitations as well as implications for research and practice are also discussed in addition to directions for future research.

Participants were adolescents and adults who were fluent in Spanish, but were less than fully proficient in English as measured by the Woodcock-Muñoz Language Survey-Revised ([WMLS-R]; Woodcock et al., 2005). Those at the lower English CALP (cognitive academic language proficiency) level scored on average at a fourth-grade equivalent level (i.e., L2 CALP 2), and those at the higher English CALP level scored on average at a sixth-grade equivalent level (i.e., L2 CALP 3). The questions of interest were (1) whether students who read a text first in Spanish and then in English would show superior comprehension of the text compared to students who read the text twice in English, (2) whether performance would differ for students at higher and lower English proficiency levels, (3) whether effects would be evident on recall, question and answer (Q&A), and cloze measures of reading comprehension, and (4) whether performance would differ depending on whether students thought aloud as they read the text the second time in English or simply read the text aloud without comment.

Results revealed that among students with lower English proficiency, reading the text first in Spanish boosted their reading comprehension more than reading the text twice in English. Effects were evident on the recall and Q&A tests but not on the cloze

test. The think aloud procedure made a difference on the Q&A and cloze tests but not on the recall test. When students thought aloud, those who read the text first in Spanish accurately answered more questions than those who read the text twice in English, but no difference occurred when they read the text without comment. On the other hand, when students thought aloud, they recalled fewer cloze words than when they read the text without comment regardless of the language of the text. In contrast, none of the variables differentially affected the performance of students with higher English proficiency. Regardless of the language of the text or the think-aloud procedure, performance was similar on the three reading comprehension measures.

Comparison of performance across the measures of reading comprehension revealed that for all groups the percentage of correct responses was higher in the Q&A and cloze tasks than in the recall task. These results indicate that it was harder to recall text information completely than to recall answers to targeted questions and to fill in missing content words in the text. However, when students were asked which type of test they considered the hardest, their responses reflected a perception that differed from their performance. The greatest number of students regarded the cloze as the hardest task, somewhat fewer students rated retelling as the hardest task, and many fewer students rated the Q&A as the hardest task.

Correlations among the three reading comprehension measures were substantial but not sufficiently high to indicate they were measuring the same construct. WMLS-R language and literacy component tests in Spanish and English explained significant variance on these measures. Among students who read the text in both languages, WMLS-R component tests in Spanish explained significant variance on all three

measures, whereas WMLS-R components in English explained significant variance only on the cloze task. In contrast, among students who read the text only in English, WMLS-R components in Spanish did not explain significant variance, whereas WMLS-R component tests in English explained unique significant variance on all the measures.

Student's responses during the think-aloud procedure were categorized according to the strategies they used when they read the text for the second time in English. The most common strategies were paraphrasing, drawing inferences, translating into Spanish, and reacting to the text. Students who had already read the text in Spanish used more reader-based strategies (e.g., making inferences and reacting to the text), and fewer text-based strategies (e.g., paraphrasing) than students who had read the text only in English. Students used Spanish and English equally often to respond in the think-aloud even when they had read the text only in English.

Effect of L1 Input and L2 Proficiency on Reading Comprehension

The main research question was whether language-minority adults would comprehend printed text better if they read it in L1 followed by L2 than if they read the text twice only in L2. Three possible outcomes were entertained:

1. It was expected that students with both lower- and higher-English-proficiency would benefit from L1 input because they all were less than fluent in L2; and that students receiving the L1 input would perform better than students who read the text only in L2.

2. Another possible outcome was that students with both lower- and higher-English-proficiency would benefit from L1 input, but that students with lower English proficiency would benefit the most from L1 input since they needed the most language

support, whereas L1 input would neither hinder nor boost performance in L2 for students with higher English proficiency since they were expected to be above the critical threshold of linguistic competence and would be able to transfer L1 skills to L2.

3. Alternatively, it was also possible that L1 input would not benefit students and that the advantage of repeated readings would be greater only for the L2-only condition since comprehension was measured in L2.

The results of the study provided support for the second hypothesis. ANOVAs revealed that only students with lower English proficiency benefitted substantially from reading the text in Spanish. Students with lower English proficiency who read the text first in Spanish and then in English performed significantly better on the recall task. Performance of this group was also superior on Q&A task but only within the think-aloud procedure. When analyzing the performance of students with higher English proficiency across reading conditions, no significant differences in performance were found on each of the reading comprehension tasks.

These results were interpreted as providing support for the hypothesis that when the learner has adequate proficiency in the L1, reading a text in the native-language prior to reading the text in the second language serves as a cognitive scaffold by providing the reader with background knowledge that is then activated when the text is read in L2. When reading the text in L1, students learn about overall content meaning as well as individual key vocabulary. These abstract knowledge structures that are stored in memory are called schemata (Carrell, 1984), and are known to affect both comprehension and memory (Anderson, 1984).

It was expected that when students encountered the text in L2, they would make connections between what was read in L1 and the information presented in L2 to aid comprehension. Successfully cross-referencing information across L1 and L2 texts was expected to lead to the resolution of word-level and phrase-level difficulties. Bernhart (1987) showed that students dedicate a large amount of cognitive resources to bottom-up processes while reading in L2. As the cognitive burden is lessened in lower-level reading processes (e.g., lexical and syntactic analyses) by providing L1 input, students are able to dedicate more cognitive resources to higher-level reading processes (e.g., integration of meaning and comprehension).

The results also showed that lower English proficiency did not necessarily hinder the student's ability to use L1 input as a scaffold to aid comprehension of text in L2, and ultimately used it to boost learning and performance. Furthermore, for students with higher English proficiency, reading the text in both Spanish and English neither benefited nor hindered their performance compared to reading the text twice in English. Other research with French-English and English-Russian college students have also found that students with lower L2 proficiency benefit significantly from reading bilingual materials; and even though high-ability students appear not to benefit above and beyond reading the text twice in English, reading bilingual materials is not detrimental to their performance (Jarvis & Jensen, 1982; Parent & Belasco, 1970).

On the other hand, students with lower English proficiency who read the text twice in English were unable to transfer native-language knowledge and skills to aid in the comprehension of L2 text. Therefore, it seems that these students were at or below the lower threshold of L2 competence, which according to Cummins (1979, 2000)

learners must surpass in order to be able to successfully transfer already-acquired knowledge and reading skills from the L1 to the L2. Moreover, providing students at this level of L2 competence appropriate and immediate L1 support served as a cognitive scaffold, which was not available to students in the English-only reading conditions.

It is important to note that the superior performance of students who were less proficient in English on two of the reading comprehension measures did not arise because they had been exposed to both L1 and L2 versions of the text and thus had available two languages in which to express their answers compared to students who had been exposed only to L2 text and thus had been provided with only English answers. Students were required to respond only in English on these tests and only correct English responses were scored. Because one of the groups had read the English text twice, one might expect that they would have better memory for English answers and thus outperform the group who was exposed to English only once. However, the opposite outcome rather than this occurred, indicating that the recall and Q&A tasks were tapping deep rather than surface level comprehension of the text.

Whereas text language influenced recall and Q&A scores of lower proficiency students, text language did not matter to their performance on the cloze test. On the former tasks, students could express their answers using their own English words, whereas on the cloze task they were restricted to recalling the exact English words from the text. In this case, one might expect that reading text twice in English would provide an advantage for memory of words. Although there was a slight difference favoring the EE condition over the SE condition (see Figure 4), the difference was not statistically

significant. The absence of a difference suggests a deeper level of comprehension was assessed by the cloze test rather than just word memory.

In addition, it is important to note the potential effect of text difficulty level and the decoding abilities of these students. CALP 2 students were reading at the 9th-grade level in L2, whereas CALP 3 students were reading above the 12th-grade level. Average L1 decoding ability across all groups was at college-level (see Table 1: Word Reading). It is important to note that the superior performance of students who were more proficient in English on comprehension measures did not arise because they had greater decoding skill in English and thus were able to process the text better. The text had a fifth-grade readability level, and thus it was below the reading level of the students. Students with lower English proficiency in the SE reading group performed at about the same level as the students in the two CALP 3 groups on all comprehension measures, and significantly better than CALP 2 students in the EE reading group on two comprehension measures (see Table 7). Furthermore, it was surprising to find that L2 decoding ability was not related to comprehension (see Table 11). Therefore, differences in performance cannot be attributed to differences in L2 decoding ability or level of text difficulty in relation to CALP levels. In addition, performance across CALP 3 groups was equivalent. Therefore, superior performance of CALP 2 students in the SE group on comprehension measures, and hence the benefit of L1 input, cannot be attributed to differences in L1 decoding ability.

Effect of Think-Aloud on Reading Comprehension

Another matter of interest was the potential effect that thinking aloud while reading the text would have on reading comprehension. On the one hand, it was possible

that having to interrupt the reading process to report one's thoughts would hinder comprehension because of the cognitive burden that the think-aloud procedure would add to an already challenging cognitive task. On the other hand, having to stop and comment on what was being read would increase the student's awareness and comprehension by requiring the student to actively interact with and think about the text. A third possibility was that the think-aloud procedure would have no effect on comprehension, and thinking aloud would merely provide a window for observing on-line processes that normally occur as text is read and comprehended. The results supported all three possibilities.

As mentioned before, analysis of performance on the reading comprehension measures showed that think-aloud effects were limited to students with lower English proficiency. Among these students, results showed that thinking aloud while reading the text had multiple effects, both positive and negative, conditioned by other factors. On the one hand, it was found that students who performed the think-aloud procedure and read the text in both languages performed significantly better on the Q&A task compared to students who performed the think-aloud and read the text only in English. In contrast, students with lower English proficiency who did not perform the think-aloud procedure and read the text in both languages performed similarly on the Q&A task to students who did not perform the think-aloud and read the text only in English. Likewise, performance across TA *vs.* No TA and language of text conditions was comparable. On the Q&A task, students were asked to provide brief answers to questions pertaining to information directly found in the text. Perhaps, performing the think-aloud after reading the text in L1 allowed them to rehearse the overall ideas in the text with more confidence as they

confirmed what was already read in L1, thus allowing them to answer the questions more accurately.

On the other hand, it was found that among the students with lower English proficiency, those in the think-loud condition performed significantly worse on the cloze task compared to students not in the think-aloud condition. This was true regardless of the language of the text they read. On this task, students were asked to provide the exact word as it appeared in the text, thus it had the strongest memory requirement. Perhaps, having to interrupt the reading process to report one's thoughts added an extra cognitive burden that negatively impacted the student's ability to recall specific words and information to complete the cloze task. Perhaps thinking aloud led to memory interference between what was read and what they thought about as they read the text, thus finding it hard to retrieve the original words from memory. In addition, the time that elapsed between reading of the text and completion of the cloze task was greater for students who performed the TA than for students who read the text without comment. This allowed for exact word memory to erode more in the former condition.

Thinking aloud while reading the text had no effect on free recall among students with lower English proficiency or on the overall performance of students with higher English proficiency. In this case, requiring students to think-aloud while reading the text only extended reading time, but it did not have an effect on reading comprehension.

Only two Second Language Acquisition (SLA) studies to date have empirically tested the issue of reactivity of thinking aloud while reading in L2 on reading comprehension (Bowles & Leow, 2005; Leow & Morgan-Short, 2004). Both studies found that thinking aloud had no effect on reading comprehension when participants were

instructed, as they were in this study, to “simply verbalize” their thoughts as they read the text. One plausible explanation for the mixed results could stem from differences in L2 proficiency among participants or the instruments used to assess reading comprehension. Bowles and Leow measured reading comprehension with a multiple-choice test, and their participants were highly L2 proficient college students. Leow and Morgan-Short’s tested college students with very low L2 proficiency, and they measured reading comprehension with three tasks also in the form of multiple-choice tests. In addition, tests were constructed in the students’ L1, and they answer in L1. In contrast, participants in the present study had moderate levels of L2 proficiency. Reading comprehension was assessed with other measures, all of which were written in L2 and for which answers were required in L2.

Performance Across Reading Comprehension Tasks

Another research question was whether students would perform differently across the various reading comprehension tasks. A possible outcome was that because the recall task required the greatest language production skill in English, performance would be weakest in this task. In line with this prediction, the results showed that students performed the lowest on the recall task across all experimental groups, whereas performance on the Q&A task was comparable to performance on the cloze task.

According to Macdonald-Ross and Scott (1995), scores above 60% represent “satisfactory” comprehension, scores between 40% and 60% represent “partial” comprehension, and scores below 40% represent “inadequate” comprehension. If reading comprehension had been measured using only one task, for example, only the recall task, we would have concluded that, overall, students did not understand the text very well

because they were only able to recall about 42% to 55% of the ideas in the text.

However, having the Q&A and cloze data revealed that the students had a relatively good understanding of the text in that they were able to comprehend about 61% to 77% of the text. In addition, effects of language of text and L2 proficiency on reading comprehension differed depending on the reading comprehension task. These results underscore the importance of employing multiple measures when assessing reading comprehension as performance on an individual task only provides a limited picture of reading comprehension processes.

The reading comprehension tasks administered varied in the retrieval cues and language processes required to complete them. They required students to construct rather than to recognize correct responses. This may explain the significant effect of test type found among both lower- and higher-English-proficiency groups. Although these may be more authentic measures of comprehension, they are usually more difficult (Brantmeier, 2005; Wolf, 1993); and for language-minority students, these could be even more challenging. Furthermore, there are variations in the difficulty levels within constructed response tasks. For instance, it has been found that tasks with a heavy production component (e.g., verbal free-recall) may underestimate comprehension (Goldman, Reyes, & Varnhagen, 1984). Shohamy (1984) also found that testing method has a differential effect on reading comprehension when comparing performance on multiple-choice versus open-ended questions. In her study, students with low levels of L2 proficiency had difficulty completing the tasks as evidenced by their lower scores, while students with higher L2 proficiency were barely affected. Nevertheless, advanced level students may still be affected by test type. For example, Wolf (1993) found that advanced L2 learners

as well as less advanced L2 learners performed significantly better when measuring reading comprehension with a multiple-choice test than with open-ended questions or a cloze test.

It is reasonable to expect that at the beginning levels of learning a second language, when reading L2 texts students will have more difficulties comprehending and integrating the information read, and producing the information during assessment. This suggests that perhaps students with lower L2 proficiency may be more susceptible to variations in instructional or reading modalities. In contrast, more advanced learners may be less affected because they can manipulate language better. In the present study, although sample means were higher for students with higher L2 proficiency than for students with lower L2 proficiency, these students' ability to demonstrate what they understood was still affected by test type in the same way that students with lower L2 proficiency were affected.

Relationship Between L1 and L2 Skills and Reading Comprehension

Research shows that both L2-specific and non-L2 specific factors seem to be involved in L2 reading (Alderson, 1984). According to Cummins (1979), language-specific skills are not easily transferred from L1 to L2, but there may be moderate to high correlations between literacy-related language skills because CALP underlies the acquisition of these. As mentioned before, interdependence of CALP across languages implies that L1 and L2 CALP are manifestations of the same underlying dimension -- a common underlying proficiency (CUP). Several research studies have shown significant correlations between Spanish and English literacy skills (Durgunoglu, Nagy, & Hancin-Bhatt, 1993; González, 1989; Hakuta & Diaz, 1985; Nagy, McClure, & Mir, 1997; Reese,

Garnier, Gallimore, & Goldenberg, 2000; Umbel & Oller, 1995), lending support to Cummins' interdependence hypothesis. In the present study, significant correlations among various components of Spanish and English proficiency were found, including verbal analogical reasoning, word reading, and writing skill, providing further support to Cummins' hypothesis.

In an attempt to uncover relationships between performance on the reading comprehension measures and measures of Spanish and English proficiency, correlation coefficients were computed. Results of the correlational analyses by language of text revealed several significant, although contrasting, relationships between the reading comprehension and proficiency measures. For the students who received L1 input, several L1 proficiency skills but few L2 proficiency skills were related to performance on the various reading comprehension tasks. In contrast, several L2 proficiency skills but few L1 proficiency skills were related to performance on comprehension tasks among students who read the text only in English. Regression analyses showed that for students who read the text in both languages, skills related to language proficiency in L1 reliably explained a significant amount of variance in scores on all reading comprehension measures, whereas L2 skills only explained a significant amount of variance on cloze performance. In contrast, among students who read the text only in English, L1 skills did not explain significant variance, whereas L2 skills explained a significant amount of variance on all tasks.

Research has shown that different reading comprehension tests tap multiple cognitive processes in different ways (Cutting & Scarborough, 2006; Francis, Fletcher, Catts, & Tomblin, 2005; Keenan, Betjmann, & Olson, 2008; Kendeou, Savage, & van

den Broek, 2009; Nation & Snowling, 1997; Vellutino, Tunmer, Jaccard, & Chen, 2007). In the present study, each reading comprehension task seemed to place somewhat different demands on L1 and L2 linguistic skills depending upon the conditions under which students read the text. Overall, WMLS-R L1 and L2 tests combined accounted for up to 36% of the total variance on a given measure of reading comprehension, leaving a significant portion of variance still unexplained.

Strategy Use as Revealed by Think-Aloud Protocols

A qualitative analysis of the strategies employed by the students as revealed through the think-aloud was conducted in an attempt to understand how native-language use facilitated the comprehension of L2 input. Analyses of the verbal protocols revealed that students accessed both languages equally when reading the L2 text regardless of the language in which they read the text first. In addition, it was found that students tended to use the same kind of strategies across both levels of English proficiency. Interestingly, students who read the text in both languages used significantly more reader-based strategies (e.g., translating and reacting to the text) and fewer text-based strategies (e.g., re-reading text), and paraphrased the text significantly less often than students who read the text only in English.

In relation to the current study, reader-based strategies are analogous with top-down strategies and processes that allow students to monitor their comprehension. On the other hand, text-based strategies are analogous with bottom-up strategies and processes that allow students to understand specific linguistic units (Block, 1986; Davis & Bistodeau, 1993; Jimenez, Garcia, & Pearson, 1996). Differences in the types of strategies used by students across reading conditions (see Table 13) show that students

who read the text in the both languages processed text more deeply during the TA procedure than students who read the text only in English. Greater use of constructive elaborative inferences by students lower proficiency students in the SE group suggests that students who read the text in both languages processed the text at a deeper level, possibly explaining why their performance was better compared to students with lower English proficiency who read the text twice in English.

In general, students' main strategies for interpreting information were to paraphrase the text into more familiar terms and to draw inferences. According to Pressley and Afflerbach (1995), paraphrasing the text into more familiar terms as well as constructing elaborative inferences and certain forms of affective reactions to text (e.g., instantiating relevant prior knowledge) are strategies that help readers interpret text and form mental representations of the meaning of the text, known as *macrostructure*. Translating text into native-language terms is also accomplished by L1 paraphrases and paraphrased translations, which also allow students to confirm the ideas conveyed in the text (Kern, 1994).

The use of native language while reading text in a second language as a way to aid comprehension has been previously noted (Block, 1986; Jimenez, Garcia, & Pearson, 1986; Upton, 1997). Similar to results in prior studies investigating strategy use among second language learners, L1 paraphrase and translation were strategies found to be used frequently. According to Kintsch and Rawson (2005), a situation model is needed in order to be able to make correct inferences about what is being read. Translating and paraphrasing text into Spanish suggests that students are consciously attempting to interpret the explicit meaning of the text. In this way, phrase-level and discourse-level

meanings can be represented and integrated in L1, hence facilitating the overall reading comprehension process. Based on the students' comments during the exit interview, it appeared that paraphrasing the text in L1 served various purposes. Among them, it helped students confirm what they thought the sentence meant, and it helped them monitor and facilitate the comprehension of individual words and ideas. Although the ANOVA was not statistically significant, the correlational analyses indicated that L1 paraphrases and translations were positively related to performance on the recall and cloze tasks.

It seems that providing L1 input allowed students with lower English proficiency to dedicate more cognitive resources for higher-level processes while reading the L2 text. Perhaps, students who read the L1-L2 versions of the text struggled less with the L2 text and were able to form a coherent situation model because the L1 input facilitated the resolution of word-related and idea-related difficulties, even though all students used Spanish to an equal extent. Furthermore, to the extent that L1 helped students develop a situation model suggests that higher-level processes are not language specific, as stipulated by the linguistic interdependence hypothesis (Cummins, 1979).

The finding that all groups used L1 equally during the TA procedure shows that students spontaneously relied in L1 as they attempted to understand L2 text. However, when L2 proficiency was below a certain level of linguistic competence, spontaneous attempts to transfer L1 knowledge and strategic reading skills to L2 reading were short-circuited (see Table 7). In the absence of L1 input, performance in L2 suffered as evidenced by the poorer performance of the CALP 2 students who read the text twice in L2 (EE). On the other hand, for CALP 2 students in the SE reading group, L1 input

facilitated L1 skills to transfer when the connection was made explicit, hence boosting performance. In fact, CALP 2 students in the SE reading group performed at about the same level as the students in the two CALP 3 groups, so scaffolded transfer from L1 to L2 raised the scores of the CALP 2 SE group to the level of reading competence supported by spontaneous transfer from L1 to L2 in the two CALP 3 groups. CALP 3 students were above the threshold and did not benefit from reading the text in L1 before L2 because their L1 skills were already free to transfer spontaneously when they read the L2 text, as evidenced by non-differences in performance between CALP 3 groups. When allowed during the TA procedure, all groups used L1 as often as L2 regardless of the text read first. This shows that they were all transferring from L1 to L2, albeit some more successfully than others, and were not just processing the information in L2.

As a whole, the results showed that students who are in the middle stages of learning a second language access both language systems often as they read in L2, whether or not L1 support is provided. This suggests that students' native-language abilities are beneficial to understanding the L2 text at all levels of learning a second language and at all levels of L2 proficiency. However, providing students at the lower levels of L2 proficiency with L1 support allows them to gain greater comprehension.

Summary Comments

To sum up, the results of this study showed that the relative contribution of L1 input changes as a function of L2 proficiency. For students with lower English proficiency, reading the text in both languages helped significantly in comprehending the text, at least as measured by the recall and Q&A tasks. Once students attained sufficient proficiency in L2, L1 input neither benefited nor hindered reading comprehension. These

results suggest that memory representations of the text are not language-specific and that meaning can transfer from L1 to L2 independently of surface form (i.e., the language in which the text is presented). These results are consistent with Cummins (1979) linguistic interdependence hypothesis that stipulates that higher-level processes are not language-specific. The present findings also suggest that students with very limited L2 proficiency can successfully transfer knowledge acquired in the native-language to influence cognitive functioning in L2 when scaffolded by L1 input, whereas otherwise transfer is short-circuited. This latter finding indicates that CALP 2 students in the EE group were at or below a lower threshold of L2 competence, as proposed by Cummins (1979).

Furthermore, it was found that text processing, in terms of strategy use, did not differ much as a function of L2 proficiency. Specifically, all students regardless of their level of English proficiency or reading condition used L1 often while interacting with the text. Although students used L1 often to facilitate their comprehension, they were not always successful as evident by the poor performance of lower proficiency students who read the text only in English. Nevertheless, the use of native language served as a metacognitive skill useful at both levels of L2 proficiency.

Finally, it was found that students performed differently across the different assessment tasks, and that students' comprehension of the text was underrepresented on the recall task. Research has shown that different reading comprehension tests tap multiple cognitive processes in different ways. Students with lower English proficiency seem to be most affected by test type since they have the greatest language difficulties. However, even though students with higher English proficiency can manipulate language better, their ability to demonstrate comprehension may still be affected by test type.

Consequently, it is important to employ multiple measures when assessing comprehension while keeping in mind the potential interaction between the cognitive demands of the task and the level of L2 proficiency of the students.

Implications for Instruction

This study has great value for research and practice given the current demand for ESL services. As research has shown, comprehension capabilities in the second language precede production capabilities in the second language. This study points to the need to use multiple measures to assess reading comprehension. Ideally, reading comprehension assessments should be constructed so that they require varying degrees of language production since the ultimate goal is to assess comprehension and not production.

In addition, translation should not be viewed as something to be discouraged, but as an important metacognitive skill that helps students construct meaning and monitor comprehension. Instructors should expect some dependence on L1 at all proficiency levels and should foster a positive image toward the use of L1. Although it may seem common sense, students may not always be clear about the potential benefits of native-language use while reading in a second language (Jimenez, 1997). Translating or paraphrasing text in L1 is a common and powerful strategy that can be incorporated in the repertoire of strategies available to L2 learners.

Finally, this study points to the role of instruction in facilitating transfer and how to adapt teaching methodologies to accommodate the needs of this population. Research has shown no adverse effects on English from bilingual programs and/or instructional time dedicated to L1 literacy skill development (Cummins, 2000). This study shows no adverse effects from using bilingual reading materials to aid L2 reading comprehension.

Helping students gain or activate relevant background knowledge by reading material in a language they can easily understand before reading it in the second language helps to increase their comprehension and memory. The results of this study support the use of bilingual reading materials at least with Spanish-speaking students with adequate levels of L1 proficiency, especially at the lower level of L2 competence.

Strengths and Limitations of the Present Study

One of the goals of this study was to try to address the limitations encountered in previous research. In prior research, L2 proficiency has often been gauged according to the number of semesters the student had been studying the L2, their scores on a few multiple-choice questions, or the student's ability to read a list of words. The present study measured more aspects of L2 and L1 oral language and lexical development including expressive vocabulary, reasoning using lexical knowledge, reading ability, and applied language proficiency using a psychometrically-sound standardized test. In addition, this study examined on-line reading processes as students read a text aloud through the use of a think-aloud procedure. Finally, reading comprehension was assessed with multiple tasks. The use of a randomized control-group design added to the strength of this study.

This study also had several weaknesses that limit the generalizability of the results. Our focus was on language-minority adults who were enrolled in remedial reading courses in the U.S. and who were native-speakers of Spanish, possibly limiting the generalization of findings to other language populations. Future research should examine whether L1 input contributes to L2 reading comprehension among students of various age levels, language backgrounds, and proficiency levels. For instance, would

similar results be obtained with children or younger adolescents who may have lower levels of L1 competence? Future research should draw a larger sample of students with varying levels of L1 proficiency. This way, it might be possible to identify how much proficiency in L1 is sufficient to allow for transfer of skills from L1 to L2. The use of one reading passage also limits the generalizability of the results. Future research should examine whether results hold with texts having different structures, topics, and length. For instance, would similar results be obtained if students read narrative texts? Given that students are more familiar with the structure of narrative texts; it is possible that students may be able to create a mental story grammar with more facility, allowing for better reading comprehension. In addition, students were not told how comprehension was going to be assessed prior to introducing the tasks, so they may not have read the passage carefully or purposefully. It is possible that students may modify their reading behaviors according to how they will be tested for comprehension. The potential effect of priming students on how comprehension will be assessed deserves further attention. With regard to testing methods, the rational-deletion cloze task was scored using the exact word method. Other scoring methods may be used. It is possible that findings may change if less stringent scoring methods are used. Despite the limitations of this study, the findings provide unique and valuable information about L2 reading comprehension processes and native-language influences.

Directions for Future Research

Several possible directions for future research stem from this study. Replications with students of diverse ages, language backgrounds and proficiency levels, using reading materials with different organizational structures and syntactic complexity,

multiple think-aloud modalities, and other testing methods might be pursued. One question that future research could address is the effect of oral versus silent reading on comprehension. How would performance differ if students were allowed to read the text silently as opposed to reading the text aloud? It is possible that among students with lower L2 proficiency who have limited command of the language, reading aloud may hinder comprehension because the extra task of pronouncing words may create an added cognitive burden. Additionally, students may focus more on pronouncing the words accurately, which can also hinder comprehension. Another question for possible research is whether reading one-sentence at a time distorts the normal reading process. Perhaps reading processes would differ if students read full pages of text and were able to look back at previously read text when there was a breakdown in comprehension, which is a known comprehension-monitoring strategy among successful readers (Ericson & Simon, 1993).

APPENDIXES

Appendix A

Written Informed Consent Form

Dear Participant,

My name is Astrid S. Rodríguez. I am a student in the Ph.D. Program in Educational Psychology at the Graduate School and University Center of the City University of New York (CUNY), and Principal Investigator of this project, entitled “The Influence of Cross-Linguistic Input and L2 Proficiency on L2 Reading Comprehension among Spanish-Speaking Adults Learning English as a Second Language.”

I am inviting you to participate in this project. This is a research study to investigate whether readers are helped to understand a story when they read versions of the story in their first and second languages. This study is expected to advance our knowledge about the effects of language on reading comprehension among adults who are learning English as a second language. Your task will be to read a story and complete three tasks to see what you understood about the story. My interest is to see whether the language of presentation affects comprehension.

We will meet twice, each for about 45 minutes during off-class hours. During the first meeting, I will ask you to complete some tasks that will tell me about your experiences with the English language. During the second meeting, you will read a short story twice and then complete some tasks to see how you understood the story. All of this information will be kept strictly confidential, and will be stored in a locked file cabinet that only I and my advisor can open. You are free not to answer any of my questions during the initial interview and to withdraw from participating in the study at any time without consequence.

The study poses no risks to you. The things you will be asked to do are just like those occurring in any educational setting such as your ESL class. The benefit of your participation is to help us understand more about how to help adult readers improve their comprehension when they read stories in English. There will be 80 adults taking part in this study.

I may publish the results of the study, but names of people, or any identifying characteristics, will not be used in any of the publications. If you would like a copy of the study, please provide me with your address and I will send you a copy in the future.

If you have any questions about this research, you can contact me at (646) 312-1384, (917) 353-3240 or arodriguez1@gc.cuny.edu, or my advisor Dr. Linnea C. Ehri at (212) 817-8294 or lehri@gc.cuny.edu. If you have questions about your rights as a participant in this study, you can contact Kay Powell, IRB Administrator, The Graduate School and University Center, City University of New York, (212) 817-7525, kpowell@gc.cuny.edu.

If you agree to volunteer in this project, please sign this form. You can keep the second copy of this form for your records. Your participation is greatly appreciated! You will be compensated \$20 for your time and agreeing to participate in this study. Thank you.

Participant's Signature	Date	Principal Investigator's Signature	Date
-------------------------	------	------------------------------------	------

Address

(En Español al reverso)

Formulario de Consentimiento Informado Escrito

Estimado Participante,

Mi nombre es Astrid S. Rodríguez. Soy un estudiante en el Ph.D. Programa en Psicología Educativa en el *Graduate School and University Center* de la universidad de la ciudad de Nueva York (CUNY), e Investigador Principal de este proyecto, titulado “*The Influence of Cross-Linguistic Input and L2 Proficiency on L2 Reading Comprehension Among Spanish-Speaking Adults Learning English as a Second Language.*”

Le estoy invitando a que participe en este proyecto. Esto es un estudio para investigar si podemos ayudar a lectores a entender una historia cuando leen versiones de la historia en su primer y segundo idioma. Se espera que este estudio avance nuestro conocimiento sobre los efectos de la lengua en la comprensión de la lectura entre adultos que están aprendiendo inglés como segunda lengua. Su tarea será leer una historia y completar tres tareas para ver lo que usted entendió sobre la historia. Mi interés es ver si el idioma de presentación afecta la comprensión.

Nos reuniremos dos veces, cada vez por cerca de 45 minutos durante horas libres de clases. Durante la primera reunión, le pediré que complete algunas tareas que me dirán sobre sus experiencias con el idioma inglés. Durante la segunda reunión, usted leerá una historia corta dos veces y después completará algunas tareas para ver cómo usted entendió la historia. Toda esta información será mantenida estrictamente confidencial, y será almacenada en un gabinete bajo llave el cual solamente yo y mi consejero puedan abrir. Usted está libre de no contestar cualquiera de mis preguntas durante la entrevista inicial y de retirar su participación en el estudio en cualquier momento sin consecuencia.

El estudio no plantea ningún riesgo a usted. Las cosas que le pediré que haga son como éstas que ocurren en otros centros educativos tal como su clase de ESL. El beneficio que su participación provee es ayudarnos a entender más sobre cómo ayudar a lectores adultos a mejorar su comprensión cuando leen historias en inglés. Habrá 80 adultos participando en este estudio.

Puede ser que en un futuro publique los resultados del estudio, pero los nombres de la gente, o alguna característica que identifiquen los participantes no será utilizada en publicaciones.

Si usted quisiera una copia del estudio, provea por favor su dirección y le enviaré una copia en el futuro.

Si usted tiene alguna pregunta sobre este estudio, usted puede contactarme al (646) 312-1384, (917) 353-3240 o arodriguez1@gc.cuny.edu o a mi consejera, Dr. Linnea C. Ehri al (212) 817-8294 o lehri@gc.cuny.edu. Si usted tiene preguntas sobre sus derechos como participante en este estudio, puede contactar a Kay Powell, la Administradora de al *IRB, Graduate School and University Center, CUNY*, al (212) 817-7525, kpowell@gc.cuny.edu.

Si usted acepta voluntariamente participar en este proyecto, por favor firme esta forma. Usted puede guardar la segunda copia de esta forma para sus expedientes. ¡Su participación se aprecia grandemente! Usted será compensado(a) \$20 por su tiempo y por aceptar a participar en este estudio. Gracias.

Firma del Participante	Fecha	Firma del Investigador Principal	Fecha
------------------------	-------	----------------------------------	-------

Dirección

Appendix B

Demographics and Language Background Interview Script

Hello. My name is Astrid S. Rodríguez. I am a doctoral student at the CUNY Graduate School. Thank you for helping with this project. I am going to ask you a few questions so that I can learn a little about you and your language learning experiences with Spanish and English. Please answer the questions honestly. You can answer in Spanish or in English. Please tell me if you do not understand something so I can explain what I mean. Do you have any questions? Would you like me to ask the questions in Spanish or in English?

1. Where were you born?
2. What is the highest grade you attended?
3. Did you have to repeat a grade due to language issues when you came to this country?
4. How old were you when you moved to the United States?
5. How many years have you lived in the United States?
6. Did you learn English in your country? If yes, was it through formal schooling? How was English taught?
7. How often did you speak and write in English in your country?
8. Did you attend school here?
9. Have you participated in special programs to help develop your skills in English? If yes,
 - a. What programs have you attended?
 - b. How long did you attend the program for?
10. How many ESL classes have you taken in this program?
11. Now that you live in U.S., what percent of your daily communications is in English?

That is the end of my questions. Do you have any questions you would like to ask me? Is there anything else you would like to tell me about yourself?

(Spanish version of the interview)

Hola. Mi nombre es Astrid S. Rodríguez y soy una estudiante en la escuela graduada de CUNY. Gracias por ayudarme con este proyecto. Voy a hacerte unas preguntas para así poder aprender un poco sobre tus experiencias aprendiendo el idioma español e inglés. Por favor contesta las preguntas honestamente. Me puedes contestar en español o en inglés. Por favor dime si no entiendes algo para así poder explicar a lo que me refiero. ¿Tienes alguna pregunta? ¿Quisieras que te haga las preguntas en español o en inglés?

1. ¿Dónde naciste?
2. ¿Cuál es el grado más alto que atendiste?
3. ¿Tuviste que repetir algún grado a consecuencia de problemas con el idioma cuando viniste a este país?
4. ¿Qué edad tenías cuando te trasladaste a los Estados Unidos?
5. ¿Cuántos años has vivido en los Estados Unidos?
6. ¿Aprendiste inglés en tu país? Si sí, ¿fue formalmente en la escuela? ¿Cómo el inglés fue enseñado?
7. ¿Cuán a menudo hablabas y/ó escribías en inglés en tu país?
8. ¿Atendiste a la escuela aquí?
9. ¿Has participado en programas especiales para desarrollar tus habilidades/destrezas en inglés? Si sí,
 - a. ¿Qué programas has atendido?
 - b. ¿Por cuánto tiempo?
10. ¿Cuántas ESL clases has tomado en este programa?
11. Ahora que vives en los Estados Unidos, ¿Cuan frecuente interactúas con otros en inglés diariamente?

He llegado al final de mis preguntas. ¿Tienes tu alguna pregunta que te gustaría hacerme?

¿Alguna otra cosa que te gustaría decirme acerca de ti?

Appendix C

Reading Passage

(Adapted from Marshall, K., 2000)

Danger Under Water

A few years ago, some divers were swimming near a beach in Florida when they noticed something strange in the water. They realized that it was a submarine that had been sunk many years before. The divers memorized the number written on the side of the submarine and swam back to the beach. When they got home, they called a place in Washington that has a list of all the ships sunk during the Second World War. The people in Washington said that this was a German submarine that had been sunk off the coast of North Carolina during the war. The divers were amazed to hear this. The submarine had drifted more than a thousand miles from North Carolina to the place in Florida where they had found it. Although the submarine had sunk, some air remained trapped inside, and it didn't sink all the way to the bottom. Instead, it drifted around the Atlantic Ocean for years. Finally, the air inside the submarine escaped, and it sank to the bottom off the coast of Florida. The people in Washington also learned that the submarine was carrying torpedoes filled with mercury. When the government in Florida heard about this, it was worried because the wreck was near a public beach. Mercury is such a strong poison that only a small amount can make people crazy or kill them. If it touches a person's skin, it can seep through cuts and poison the body. The government tried to figure out a way to get rid of the poison torpedoes. The problem was that whatever they did might cause the torpedoes to break open and let the mercury escape. The government decided to leave the submarine in the ocean and hope that no mercury would escape.

(Spanish version of the reading passage)

Peligro Bajo el Agua

Hace algunos años, unos buzos estaban nadando cerca de una playa en la Florida cuando ellos notaron una figura extraña en el agua. Ellos se dieron cuenta que era un submarino que había sido hundido muchos años atrás. Los buzos memorizaron el número escrito en el lado del submarino y nadaron de vuelta a la playa. Cuando llegaron a la casa, ellos llamaron a un lugar en Washington que tiene una lista de todos los barcos hundidos durante la Segunda Guerra Mundial. Las personas en Washington dijeron que éste era un submarino Alemán que se había hundido cerca de la costa de Carolina del Norte durante la guerra. Los buzos estaban sorprendidos al oír

esto. El submarino había estado a la deriva por más de mil millas, desde Carolina del Norte al lugar en la Florida en donde lo habían encontrado. A pesar de que el submarino se había hundido, un poco de aire permaneció atrapado adentro, y no se hundió completamente hasta el fondo. En su lugar, estuvo a la deriva a través del Océano Atlántico por años. Finalmente, el aire adentro del submarino escapó, y se hundió al fondo cerca de la costa de la Florida. Las personas en Washington también aprendieron que el submarino llevaba adentro torpedos llenos de mercurio. Cuando el gobierno de la Florida oyó sobre esto, se preocupó por que el naufragio estaba cerca de una playa pública. El mercurio es un veneno tan fuerte que sólo una pequeña cantidad puede hacer que la gente se vuelva loca o que mueran. Si hace contacto con la piel de una persona, puede filtrarse a través de alguna cortadura y envenenar el cuerpo. El gobierno trató de descifrar alguna manera de deshacerse de los torpedos venenosos. El problema era que cualquier cosa que ellos hicieran podía causar que los torpedos se rompieran y permitiera que el mercurio escapara. El gobierno decidió dejar el submarino en el fondo del océano, con la esperanza de que el mercurio no escape.

Appendix D

General Instructions: Meeting # 2

Hello. Let's begin. I am going to read the instructions. If there is something you do not understand I'll be happy to tell them to you in Spanish. I am going to tape record the session to help me remember what happened. Would you like me to read the instructions in Spanish or in English?

I. Text Reading Activity: First Reading

I am interested in studying how readers understand short stories. You will be asked to read a short story out loud to me. You will get to read the story twice. Think about what happens so the story will make sense to you and you can remember it.

Group X (L1-L2) – The first time you read the story, it will be written in Spanish. The second time you read the story, it will be written in English.

Group Y (L2-L2) - The story will be written in English.

The story will be presented on this computer. Each line of the story will appear on a separate screen. The title will appear on the first screen. To see each line of the story, press the spacebar. Please read the story aloud to me. If you come to some hard words, I won't be able to help you, so please read them the best you can. Think about what you are reading so you can understand the story, because I'll be asking you some questions about the story later on. Let's begin.

[Prompt to participant if help is requested: "Do the best you can." "As lo mejor que puedas."]

(Spanish version)

Estoy interesada en saber cómo lectores entienden historias breves. Te pediré que leas en voz alta una breve historia. Vas a leer la historia dos veces.

Group X (L1-L2) – La primera vez leerás la historia en español. La segunda vez leerás la historia en inglés.

Group Y (L2-L2) – La historia estará escrita en inglés.

La historia será presentada en esta computadora. Cada línea de la historia aparecerá en una hoja diferente. El título aparecerá en la primera hoja. Para ver cada línea, presiona la tecla de espacio. Por favor lee la historia en voz alta. Si encuentras una palabra que se te es difícil no te podré ayudar so por favor léelas lo mejor que puedas. Piensa en lo que estás leyendo para que puedas entender la historia porque te voy a hacer unas preguntas acerca de la historia luego. Comencemos.

II. Text Reading Activity: Second Reading Think-Aloud Condition

Now you are going to read the same story aloud to me a second time around.

Group X (L1-L2) - This time, the story will be written in English.

I will ask you to read it in a special way this time. I'll explain how. Stop me if something doesn't make sense and I'll explain. This time I would like you to "think aloud" after you read each sentence. That is, I want you to tell me everything that comes to mind after you read each sentence. For example, maybe you think about what will happen next. Maybe you think of a different way to explain what is happening. Maybe you don't understand so you talk about what is confusing or what words you don't know. I want you to tell me every thought that comes to mind as you read the story even if it doesn't seem related to the story. This will help me understand how you are making sense of the story and how it relates to your ideas and experiences. You will see a red star at the end of each sentence to remind you to stop and say everything that you are thinking about at the moment. This will tell me whether and how you are understanding the story. You may do this in your language of preference. Try to understand the story as best you can, because I'll be asking you some questions when you finish. Do you have any questions? Are you ready? Remember to read the story aloud to me and think-aloud after you read each sentence. If you come to some hard words, I won't be able to help you, so please read them the best you can. Let's begin.

[Prompt given three seconds after reading sentence to participant to stimulate her/him to "think aloud" if silent: "What are you thinking?"]

[Prompt given if participant does not stop to "think aloud" and reads-on: "Remember to tell me everything that comes to mind after you read each sentence."]

(Spanish version)

Ahora vas a leer la misma historia en voz alta una segunda vez.

Group X (L1-L2) – Esta vez la historia va a estar escrita en inglés.

Te voy a pedir que esta vez leas la historia en una forma especial. Te voy a explicar cómo. Interrúmpeme si algo no te hace sentido y te explicare. Esta vez, voy a pedirte que “pienses en voz alta” después que leas cada oración. Es decir, quiero que me digas todo lo que te venga a la mente después de leer cada oración. Por ejemplo, quizás estas pensando acerca de lo que va a pasar después. Quizás piensas acerca una manera diferente de explicar lo que estás pasando. Quizás quieres hablar de algo que es confuso o palabras que no sabes. Quiero que me digas cada pensamiento que tengas cuando leas la historia aunque no parezca ser relacionado con la historia. Así me ayudarás a entender cómo es que tu estas tratando de entender la historia y como se relaciona con tus ideas y experiencias. Vas a ver una estrella roja al final de cada oración como recordatorio que debes parar y decir todo lo que estés pensando en el momento. Esto me dirá como y sí estas entendiendo la historia. Puedes usar el idioma que desees. Quiero que trates de comprender la historia lo mejor que puedas porque te voy a pedir que contestes unas preguntas cuando acabes. ¿Tienes alguna pregunta? ¿Estás listo(a)? Recuerda de leer la historia en voz alta y de decirme lo que estés pensando después de que leas cada oración. Si encuentras alguna palabra que es difícil no te podré ayudar, so por favor léelas lo mejor que puedas. Comencemos.

II. Text Reading Activity: Second Reading No Think-Aloud Condition

Now you are going to read the same story aloud to me a second time around.

Group X (L1-L2) - This time, the story will be written in English.

Try to understand the story as best you can, because I'll be asking you some questions when you finish. Do you have any questions? Are you ready? Remember to read the story aloud to me. If you come to some hard words, I won't be able to help you, so please read them the best you can. Let's begin.

(Spanish version)

Ahora vas a leer la misma historia en voz alta una segunda vez.

Group X (L1-L2) –Esta vez la historia va a estar escrita en inglés.

Quiero que trates de comprender la historia lo mejor que puedas porque te voy a pedir que contestes unas preguntas cuando acabes. ¿Tienes alguna pregunta? ¿Estás listo(a)? Recuerda de leer la historia en voz alta. Si encuentras alguna palabra que es difícil no te podré ayudar, so por favor léelas lo mejor que puedas. Comencemos.

III. Recall Task

I would like you to tell me everything you remember from the story you just read. Pretend that you are telling the story to someone who has never read it or heard it. I would like you to retell the story in English. You can begin when ready.

Prompt if L1 is used: “Remember, use English to tell me the story.”

Prompt at the end: “Can you remember anything else?” [If not,] “OK. Let’s continue.”

(Spanish version)

Me gustaría que me digas todo lo que recuerdes acerca de la historia que acabas de leer. Pretende que tienes que decirle la historia a alguien que nunca la ha leído o escuchado. Me gustaría que relataras la historia en inglés. Comienza cuando estés listo(a).

Prompt if L1 is used: “Recuerda de relatarme la historia en inglés.”

Prompt at the end: “¿Puedes recordar algo mas? [If not,] “Esta bien, continuemos.”

IV. Q&A Task

Now you are to answer 8 questions. These questions are about information presented in the story you just read. Please read each question carefully and say your answer to me as completely as possible. Be sure to say your answer in English only. Each question will be on a separate slide. Click on “next” to go to the next question.

(Spanish version)

Ahora vas a contestar 8 preguntas. Estas preguntas son acerca de información en la historia que acabas de leer. Por favor lee cada pregunta cuidadosamente y dime tu contestación lo más completamente posible en voz alta. Recuerda de decir tus contestaciones en inglés solamente. Cada pregunta aparecerá en una hoja diferente en la computadora. Oprime en “next” para avanzar a la próxima pregunta.

V. Cloze Task

Now you are going to see the same story again, but this time some of the words will be missing. Your job is to fill in the missing words to complete the sentences. Try to remember the words and ideas as they appear in the story or make your best guess if you can't remember the exact word. Write the answers in the blank spaces. Try not to leave anyone blank. Remember to answer in English.

(Spanish version)

Ahora vas a ver la misma historia otra vez, pero esta vez algunas palabras están omitidas. Tu tarea es proveer las palabras omitidas para completar la oración según estaba escrita en la historia. De no acordarte de la palabra exacta, has tu mejor esfuerzo para completar la oración basada en las ideas en la historia. Escribe las contestaciones en el espacio en blanco. Trata de no dejar ningún espacio en blanco. Recuerda de contestar en inglés.

VI. Exit Interview (see Appendix I)

Appendix E

Recall Task Scoring Sheet

Number of idea units: 24 items

- ___ some divers were swimming
- ___ a beach in Florida
- ___ they noticed something strange in the water
- ___ it was a submarine
- ___ that had been sunk many years before
- ___ the divers memorized the number written on the side of the submarine (and swam back to the beach)
- ___ they called a place in Washington (when they got home)
- ___ that has a list of all the ships sunk during the Second World War
- ___ this was a German submarine
- ___ that had been sunk off the coast of North Carolina during the war
- ___ the divers were amazed to hear this
- ___ the submarine had drifted (more than a thousand miles) or (from North Carolina to Florida)
- ___ some air remained trapped inside
- ___ it didn't sink all the way to the bottom but drifted around the Atlantic Ocean for years
- ___ finally, the air inside the submarine escaped and it sank to the bottom
- ___ the submarine was carrying torpedoes
- ___ filled with mercury
- ___ the government in Florida was worried because the wreck was near a public beach
- ___ mercury is such a strong poison
- ___ that only a small amount can make people crazy or kill them
- ___ if it touches a person's skin, it can seep through cuts and poison the body
- ___ the government tried to figure out a way to get rid of the poison torpedoes
- ___ the problem was that torpedoes might break open and let the mercury escape
- ___ the government decided to leave the submarine in the ocean

Appendix F

Comprehension Questions on Q&A Task and Scoring Key

Maximum Score: 9 points

1. What did the divers discover? (1 point) Where? (1 point)
 - A submarine near a beach (in the ocean off the coast of) in Florida. (2 points)
 - A submarine (ship) in a beach (in the ocean) in Florida. (1.5 point)
 - A submarine in the ocean (in Florida). (1.5 point)
 - A ship in the ocean. (1 point)

2. Who or where did the divers call in Washington? (1 point)
 - The divers called a place/office in Washington that has a list of all the ships sunk during WWII. (1 point)
 - People that knew about ships (had the names of ships) sunk/lost during WWII. (1 point)
 - People that knew about ships/submarines. (.5 point)
 - A place where they have a list of ships/submarines sunk. (.5 point)
 - A place where they know/have information about the ships/submarines that went to war. (.5 point)
 - The government. (.25 points)

3. How did the people in Washington know which submarine the divers had found? (1 point)
 - Through the number/code written on (the side of) the submarine, which the divers had memorized. (1 point)

4. Why didn't the submarine sink straight to the bottom when it was hit during the war? (1 point)
 - Because some air had remained trapped inside. (1 point)

5. How many miles did the submarine drift? (1 point)
 - More than a thousand miles. (1 point)
 - From North Carolina to Florida. (1 point)

- 1,000 miles (or thousands of miles). (.5 point)
6. Why was the submarine dangerous? (1 point)
- Because it was carrying torpedoes filled with mercury. (1 point)
 - Because it had torpedoes with poison. (.75 point)
 - Because it had cans of mercury inside. (.75 point)
 - Gas can explode/spread and hurt people. (.5 point)
 - Because it had poison. (.5 point)
 - Because it had mercury. (.5 point)
 - Because it had torpedoes. (.5 point)
7. Who might get hurt? (1 point)
- People near/swimming in/at/using/that goes to the beach. (1 point)
 - People in Florida. (.75 point)
 - People. (.25 point)
8. What is the problem with trying to pull the submarine up to the surface? (1 point)
- That whatever they did might cause the torpedoes to break open/crack and let the mercury escape (or come out/affect people). (1 point)
 - The mercury can/would come out/escape. (.5 point)
 - The poison can/would spread in the water. (.5 point)
 - It was easy to/could break/crack. (.5 point)
 - The poison can create danger for people. (.5 point)

Appendix G

Cloze Task and Scoring Key

Name: _____

Danger Under Water

A few years ago, some _____ were swimming near a beach in _____ when they noticed something strange in the _____. They realized that it was a _____ that had been sunk many years _____. The divers memorized the number _____ on the side of the _____ and swam back to the _____. When they got home, they _____ a place in Washington that has a _____ of all the ships sunk during the _____ World War. The people in _____ said that this was a German _____ that had been sunk off the _____ of North Carolina during the _____. The divers were amazed to _____ this. The submarine had drifted _____ than a thousand miles from North _____ to the place in Florida where _____ had found it. Although the _____ had sunk, some air remained trapped _____, and it didn't sink all the way to the _____. Instead, it drifted around the Atlantic _____ for years. Finally, the air _____ the submarine escaped, and it _____ to the bottom off the _____ of Florida. The people in _____ also learned that the submarine was carrying _____ filled with mercury. When the _____ in Florida heard about this, it was _____ because the wreck was near a _____ beach. Mercury is such a strong _____ that only a small amount can make _____ crazy or kill them. If it _____ a person's skin, it can seep through _____ and poison the body. The _____ tried to figure out a _____ to get rid of the poison _____. The problem was that whatever they _____ might

cause the torpedoes to break _____ and let the mercury escape. The _____
decided to leave the submarine in the _____ and hope that no mercury would
_____.

A few years ago, some **divers** were swimming near a beach in **Florida** when they noticed something strange in the **water**. They realized that it was a **submarine** that had been sunk many years **before**. The divers memorized the number **written** on the side of the **submarine** and swam back to the **beach**. When they got home, they **called** a place in Washington that has a **list** of all the ships sunk during the **Second** World War. The people in **Washington** said that this was a German **submarine** that had been sunk off the **coast** of North Carolina during the **war**. The divers were amazed to **hear** this. The submarine had drifted **more** than a thousand miles from North **Carolina** to the place in Florida where **they** had found it. Although the **submarine** had sunk, some air remained trapped **inside**, and it didn't sink all the way to the **bottom**. Instead, it drifted around the Atlantic **Ocean** for years. Finally, the air **inside** the submarine escaped, and it **sank** to the bottom off the **coast** of Florida. The people in **Washington** also learned that the submarine was carrying **torpedoes** filled with mercury. When the **government** in Florida heard about this, it was **worried** because the wreck was near a **public** beach. Mercury is such a strong **poison** that only a small amount can make **people** crazy or kill them. If it **touches** a person's skin, it can seep through **cuts** and poison the body. The **government** tried to figure out a **way** to get rid of the poison **torpedoes**. The problem was that whatever they **did** might cause the torpedoes to break **open** and let the mercury escape. The **government** decided to leave the submarine in the **ocean** and hope that no mercury would **escape**.

* The words on bold were omitted for the cloze task.

Appendix H

Exit Interview

Now that you have finished reading the story and completing the tasks, I would like to ask you a few more questions about the tasks so that I can find out how to make them better next time. Please answer the questions honestly. You can speak in Spanish or English. Please ask me if you don't understand something so that I can explain what I mean. Do you have any questions? You can also feel free to ask me any questions at the end.

1. Was this story difficult or easy to understand? Why? What made it easy or difficult?
2. Were you familiar with this story?
3. Were there any words in the story that you didn't understand?
4. If you had to rank the tasks from hard to easy, which task was easiest? Hardest?
5. What made the easy task so easy for you?
6. What made the hardest task so hard for you?
7. When you retold the story to me, could you have done better if you told me in Spanish?
8. On the Questions & Answers task - do you feel you could have done better if you had answered the questions in Spanish?
9. On the fill in the blank task - do you feel you could have done better if you had been allowed to complete the sentences in Spanish?

(Spanish version of the exit interview)

Ahora que has terminado de leer la historia y completado todas las tareas, quisiera hacerte algunas preguntas acerca de las tareas de modo que pueda descubrir cómo hacerlas mejor en una próxima ocasión. Por favor contesta las preguntas honestamente. Me puedes hablar en español o inglés. Por favor pregúntame si no entiendes algo para así poder explicar a lo que me refiero. ¿Tienes alguna pregunta? También me puedes hacer preguntas cuando terminemos.

1. ¿Fue la historia difícil o fácil de entender? ¿Por qué? ¿Qué la hizo fácil o difícil?
2. ¿Habías escuchado esta historia anteriormente?
3. ¿Hubieron algunas palabras en la historia que no entendiste?
4. ¿Si tuvieras que sortear las tareas de fácil a difícil, que tarea fue mas fácil? ¿Más difícil?
5. ¿Qué hizo la tarea mas fácil tan fácil para ti?

6. ¿Qué hizo la tarea más difícil tan difícil para ti?
7. Cuando me relataste la historia, ¿Piensas que hubieras podido hacer mejor si me hubieras dicho en español?
8. En la tarea de “Preguntas & Respuestas” - ¿Piensas que hubieras podido hacer mejor si hubieras contestado las preguntas en español?
9. En la tarea de “Llena Blancos”- ¿Piensas que hubieras podido hacer mejor si hubieras completado las oraciones en español?

References

- Alderson, J. C. (1984). Reading in a foreign language: A reading problem or a language problem? In J. C. Alderson & A. H. Urquhart (Eds.), *Reading in a foreign language* (pp. 122-135). New York: Longman.
- Alderson, J. C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.
- Alvarado, C. G., Ruef, M. L., & Schrank, F. A. (2005). *Comprehensive manual. Woodcock-Muñoz Language Survey-Revised*. Itasca, IL: Riverside Publishing.
- American Fact Finder. (2008). U.S. Census Bureau, 2006-2008 American Community Survey 3-Year Estimates. Retrieved on April 27, 2010, from <http://factfinder.census.gov>
- Anderson, R. C. (1984). Role of the reader's schema in comprehension, learning, and memory. In R. C. Anderson, J. Osborn, & R. J. Tierney (Eds.), *Learning to read in American schools: Basal readers and content texts* (pp. 469-482). Hillsdale, NJ: Lawrence Erlbaum Associates.
- August, D., & Shanahan, L. (Eds.). (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on language minority children and youth*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Baker, C., & Jones, S. P. (1998). *Encyclopedia of bilingualism and bilingual education*. Clevedon, UK: Multilingual Matters.
- Bernard, W. (1950). Psychological principles of language learning and the bilingual reading method. *Modern Language Journal*, 35(5), 87-96.
- Bernhardt, E. B. (1987). Cognitive processes in L2: An examination of reading behaviors. In J. Lantolf & A. Labarca (Eds.), *Delaware symposia on language studies: Research in second language acquisition in the classroom setting* (pp. 35-50). Norwood, NJ: Ablex.
- Bernhardt, E. B. (1991). *Reading development in a second language: Theoretical, empirical and classroom perspectives*. Norwood, NJ: Ablex.

- Block, E. (1986). The comprehension strategies of second language readers. *TESOL Quarterly*, 20(3), 463-494.
- Brantmeier, C. (2005). Effects of reader's knowledge, text type, and test type on L1 and L2 reading comprehension in Spanish. *The Modern Language Journal*, 89(i), 37-53.
- Brisbois, J. E. (1995). Connections between first- and second-language reading. *Journal of Reading Behavior*, 27(4), 565-584.
- Bowles, M. A., & Leow, R. P. (2005). Reactivity and type of verbal report in SLA research methodology: Expanding the scope of investigation. *Studies in Second Language Acquisition*, 27, 415-440.
- Carlisle, J. F., Beeman, M., Davis, L. H., & Spharim, G. (1999). Relationship of metalinguistic capabilities and reading achievement for children who are becoming bilingual. *Applied Psycholinguistics*, 20(4), 459-478.
- Carrell, P. L. (1984). Schema theory and ESL reading: Classroom implications and applications. *The Modern Language Journal*, 68(4), 332-343.
- Clarke, M. A. (1979). Reading in Spanish and English: Evidence from Adult ESL students. *Language Learning*, 29(1), 121-150.
- Clarke, M. A. (1980). The short circuit hypothesis of ESL reading--or when language competence interferes with reading performance. *Modern Language Journal*, 64(2), 203-209.
- Cohen, J. (1988). *Statistical Power Analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates
- Collier, V. P. (1987). Age and rate of acquisition of second language for academic purposes. *TESOL Quarterly*, 21(4), 617-641.
- Collier, V. P. (1989). How long? A synthesis of research on academic achievement in a second language. *TESOL Quarterly*, 23(3), 509-531.
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49(2), 222-251.
- Cummins, J. (1980). The cross-lingual dimensions of language proficiency: Implications for bilingual education and the optimal age issue. *TESOL Quarterly*, 14(2), 175.
- Cummins, J. (1981). Empirical and theoretical underpinnings of bilingual education. *Journal of Education*, 163(1), 16-29.

- Cummins, J. (2000). *Language, power and pedagogy: Bilingual children in the crossfire*. Tonawanda, NY: Multilingual Matters Ltd.
- Cummins, J. (2003). *Bilingual education in the United States: Power, pedagogy, and possibility*. Retrieved on October 25, 2008, from <http://www.iteachilearn.com/cummins/bilingualedus.html>
- Cummins, J. (2010). *BICS and CALP*. Retrieved on September 2, 2010, from <http://www.iteachilearn.com/cummins/bicscalp.html>
- Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and their cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading, 10*(3), 277-299.
- Davis, J. N., & Bistodeau, L. (1993). How do L1 and L2 reading differ? Evidence from think aloud protocols. *Modern Language Journal, 77*(iv), 459-472.
- Diaz, R. M., & Klinger, C. (1991). Towards an explanatory model of interaction between bilingualism and cognitive development. In E. Bialystok (Ed.), *Language processing in bilingual children* (pp. 167-192). Cambridge University Press.
- Dulay, H., & Burt, M. (1980). Aspects of bilingual education for LES/NES students. In R. Cervantes (Ed.), *Bilingual program, policy, and assessment issues*. Sacramento: California State Department of Education Office of Bilingual Bicultural Education.
- Durgunoglu, A. Y., Nagy, W. E., & Hancin-Bhatt, B. J. (1993). Cross-language transfer of phonological awareness. *Journal of educational psychology, 85*(3), 453-465.
- Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K., et al. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes among older struggling readers. *Review of Educational Research, 79*(1), 262-300.
- Education Commission of the States. (2004). *Hispanic achievement* (Vol. 6, No. 3). Denver, CO: Author. Available online at <http://www.ecs.org/clearinghouse/53/97/5397.pdf>
- Ehri, L. C. (1998). Grapheme-phoneme knowledge is essential for learning to read words in English. In J. Metsala and L. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 3-40). Mahwah, NJ: Lawrence Erlbaum Associates.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (Rev. ed.). Cambridge, MA: MIT Press.

- Francis, D. J., Fletcher, J. M., Catts, H. W., & Tomblin, J. B. (2005). Dimensions affecting the assessment of reading comprehension. In S. G. Paris & S. A. Stahl (Eds.), *Children's reading comprehension and assessment* (pp. 369–394). Mahwah, NJ: Lawrence Erlbaum Associates.
- Friesen, D. C., & Jared, D. (2007). Cross-language message- and word-level transfer effects in bilingual text processing. *Memory & Cognition*, 35(7), 1542-1556.
- Fuchs, L. S., Fuchs, D., & Maxwell, L. (1988). The validity of informal reading comprehension measures. *Remedial and Special Education [RASE]*, 9(2), 20-28.
- Goldman, S. R., Reyes, M., & Varnhagen, C. K. (1984). Understanding fables in first and second languages. *NABE: The Journal for the National Association for Bilingual Education*, 8(2), 835-866.
- González, L. A. (1989). Native language education: The key to English literacy skills. In D. J. Bixler-Márquez, G. K. Green, & J. L. Ornstein-Galicia (Eds.), *Mexican-American Spanish in its societal and cultural contexts* (pp. 209– 224). Brownsville: Pan American University.
- Greene, B. B. (2001). Testing reading comprehension of theoretical discourse with cloze. *Journal of Research in Reading*, 24(1), 82-98.
- Greene, J. P. (1997). A meta-analysis of the Rosell and Baker review of bilingual education research. *Bilingual Research Journal*, 21(2), 1-22.
- Harris, A. J., & Sipoy, E. R. (1980). *How to Increase Reading Ability*. New York: Longman.
- Hadley, A. O. (1993). *Teaching language in context*. Boston: Heinle & Heinle.
- Hakuta, K., Butler, Y. G., & Witt, D. (2000). *How long does it take English learners to attain proficiency?* (Report No. FL 026-180). Santa Barbara: University of California Linguistic Minority Research Institute.
- Hakuta, K., & Diaz, R.M. (1985). The relationship between degree of bilingualism and cognitive ability: A critical discussion and some new longitudinal data. In K. E. Nelson (Ed.), *Children's language* (Vol. 5). Hillsdale NJ: Lawrence Erlbaum Associates.
- Hancin-Bhatt, B., & Nagy, W. (1994). Lexical transfer and second language morphological development. *Applied Psycholinguistics*, 15(3), 289-310.
- Haynes, J. R. (1989). *The complete problem solver*. Hillsdale, NJ: Lawrence Erlbaum Associates.

- Hock Seng, G., & Hashim, F. (2006). Use of L1 in L2 reading comprehension among tertiary ESL learners. *Reading in a Foreign Language, 18*(1), 29-54.
- Jarvis, D. K., & Jensen, D. C. (1982). The effect of parallel translations on second language reading and syntax acquisition. *Modern Language Journal, 66*, 18-23.
- Jiménez, R. T. (1997). The strategic reading abilities and potential of five low-literacy Latina/o readers in Middle School. *Reading Research Quarterly, 32*(3), 224-243.
- Jiménez, R. T., García, G. E., & Pearson, P. D. (1996). The reading strategies of bilingual Latina/o students who are successful English readers: Opportunities and obstacles. *Reading Research Quarterly, 31*(1), 90-112.
- Keenan, J. M., Betjemann, R. S., & Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading, 12*(3), 281-300.
- Kendeou, P., van den Broek, P., White, M. J., & Lynch, J. S. (2009). Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of Educational Psychology, 101*(4), 765-778.
- Kern, R. G. (1994). The role of mental translation in second language reading. *Studies in Second Language Acquisition, 16*(4), 441-461.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 92*(2), 163-182.
- Kintsch, W. (2004). The construction-integration model of text comprehension and its implications for instruction. In R. B. Rudell & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (5th ed., pp. 1270-1328). Newark, DE: International Reading Association.
- Kintsch, W., & Rawson, K. A. (2005). Comprehension. In M. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 209-226). Malden, MA: Blackwell.
- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*(5), 363-394.
- Kramer, V. R., & Schell, L. M. (1982). English auditory discrimination skills of Spanish-speaking children. *Alberta Journal of Educational Research, 28*(1), 1-8.

- Kutner, M., Greenberg, E., Jin, Y., Boyle, B., Hsu, Y., & Dunleavy, E. (2007). *Literacy in everyday life. Results from the 2003 National Assessment of Adult Literacy*. (NCES Publication No. 2007-480). U.S. Department of Education. Available online at <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2007480>
- Lasisi, M. J., Falodun, S., & Onyehalu, A. S. (1988). The comprehension of first and second-language prose. *Journal of Research in Reading*, 11(1), 26-35.
- Lee, J., & Schallert, D. L. (1997). The relative contribution of L2 language proficiency and L1 reading ability to L2 reading performance: A test of the threshold hypothesis in an EFL context. *TESOL Quarterly*, 31(4), 713-739.
- Lipson, M. Y., & Wixson, K. K. (2009). *Assessment and instruction of reading and writing disability: An interactive approach* (4th ed.). New York: Longman
- Leow, R. P., & Morgan-Short, K. (2004). To think aloud or not to think aloud: The issue of reactivity in SLA research methodology. *Studies in Second Language Acquisition*, 26, 35-57.
- Lucas, T., & Katz, A. (1994). Reframing the debate: The roles of native languages in English-only programs for language minority students. *TESOL Quarterly*, 28(3), 537-561.
- Macdonald-Ross, M., & Scott, B. (1995). *Results of the survey of OU students' reading skills: Text and readers programme* (Tech. Rep. No. 3). Walton, England: Open University, Institute of Educational Technology.
- Marshall, K. (2000). Danger under water. *The Kim Marshall series: Reading book 1* (p. 88). Cambridge, MA: Educators Publishing Service.
- Nagy, W. E., McClure, E. F., & Mir, M. (1997). Linguistic transfer and the use of context by Spanish-English bilinguals. *Applied Psycholinguistics*, 18(4), 431-452.
- Nation, I. S. P. (2005). Teaching and learning vocabulary. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning*. Rutledge Publishing.
- Nation, K., & Snowling, M. (1997). Assessing reading difficulties: The validity and utility of current measures of reading skill. *British Journal of Educational Psychology*, 67, 359-370.

- National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for instruction. Report of the subgroups.* (NIH Publication No. 00-4754). Washington, DC: Government Printing Office. Available online at <http://www.nationalreadingpanel.org/Publications/subgroups.htm>
- New York City Department of Education. (2008). *Language allocation policy guidelines. The LAP handbook for ELL programs.* Office of English Language Learners. Retrieved on October 27, 2008, from <http://schools.nyc.gov/Academics/ELL/KeyDocuments/Language+Allocation+Policy.htm>
- New York City Department of Education. (2009). *The 2008-2009 demographics of New York City's English language learners.* City of New York Department of Education. Retrieved on April 13, 2010, from http://schools.nyc.gov/NR/ronlyres/3A4AEC4C-14BD-49C4-B2E68EDF5D873BE4/78726/ELLDEMO2009_FINAL.pdf
- New York State Education Department. (2009). The teaching of language arts to limited English proficient/ English language learners: Learning standards for English as a second language. Office of Bilingual Education and Foreign Language Studies. Retrieved on August 11, 2010, from <http://www.emsc.nysed.gov/biling/resource/ESL/standards.html>
- Olshavsky, J. E. (1977). Reading as problem solving: An investigation of strategies. *Reading Research Quarterly, 12*(4), 654-674.
- Ordóñez, C. L., Carlo, M. S., Snow, C. E., & McLaughlin, B. (2002). Depth and breadth of vocabulary in two languages: Which vocabulary skills transfer? *Journal of Educational Psychology, 94*(4), 719-728.
- Parent, P. P., & Belasco, S. (1970). Parallel-column bilingual reading materials as a pedagogical device: An experimental evaluation. *Modern Language Journal, 54*(7), 493-504.
- Paulston, C. B. (1988). *International handbook of bilingualism and bilingual education.* Westport, CT: Greenwood Press.
- Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading.* Hillsdale, NJ: Lawrence Erlbaum Associates.
- Quiroga, T., Lemos-Britton, Z., Mostafapour, E., Abbott, R. D., & Berninger, V. W. (2002). Phonological awareness and beginning reading in Spanish-speaking ESL first graders: Research into practice. *Journal of School Psychology, 40*(1), 85-111.

- Reese, L., Garnier, H., Gallimore, R., & Goldenberg, C. (2000). Longitudinal analysis of the antecedents of emergent Spanish literacy and middle-school English reading achievement of Spanish-speaking students. *American Educational Research Journal*, 37(3), 633-662.
- Rumelhart, D. E. (1977). Toward an interactive model of reading. In R. B. Rudell & M. R. Rudell (Eds.), *Theoretical models and processes of reading* (pp. 864-894). Newark, DE: International Reading Association.
- Shohamy, E. (1984). Does testing method make a difference? The case of reading comprehension. *Language Testing*, 1(2), 147-170.
- Slavin, R. E., & Cheung, A. (2004). *A synthesis of research on language of reading instruction for English language learners*. Baltimore, MD: John Hopkins University.
- Sparkman, C. F. (1949). A new language must be spliced onto one's native language. *Modern Language Journal*, 33(5), 355-362.
- Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, 16, 32-71.
- Stanovich, K. E. (1984). Toward an interactive-compensatory model of reading: A confluence of developmental, experiential, and educational psychology. *Remedial and Special Education [RASE]*, 5(3), 11-19.
- Umbel, V.M., & Oller, D.K. (1995) Developmental changes in receptive vocabulary in Hispanic bilingual school children. In B. Harley (Ed.), *Lexical issues in language learning* (pp. 59– 80). Ann Arbor: Research Club in Language Learning.
- Upton, T. A. (1997). First and second language use in reading comprehension strategies of Japanese ESL students. *TESL-EJ*, 3(1), 1-27.
- U.S. Department of Education. (2006). *State-administered adult education program: Program year 2004-2005 enrollment*. U.S. Department of Education: Office of Vocational and Adult Education Division of Adult Education and Literacy. Retrieved May 15, 2008, from <http://www.ed.gov/about/offices/list/ovae/pi/AdultEd/aedatatables.html>
- Vellutino, F. R., Tunmer, W. E., Jaccard, J. J., & Chen, R. (2007). Components of reading ability: Multivariate evidence for a convergent skill model of reading development. *Scientific Studies of Reading*, 11(1), 3-32.
- Wolf, D. F. (1993). A comparison of assessment tasks used to measure FL reading comprehension. *The Modern Language Journal*, 77(4), 473-489.

- Woodcock, R. W., Muñoz-Sandoval, A. F., Rief, M. L., & Alvarado, C. G. (2005).
Woodcock-Muñoz Language Survey-Revised (WMLS-R), English Form A. Itasca,
IL: Riverside Publishing.
- Woodcock, R. W., Muñoz-Sandoval, A. F., Rief, M. L., & Alvarado, C. G. (2005).
Woodcock-Muñoz Language Survey-Revised (WMLS-R), Spanish Form. Itasca,
IL: Riverside Publishing.
- Zabucky, K. M., & Moore, D. (1999). Influence of text genre on adults' monitoring of
understanding and recall. *Educational Gerontology, 25*, 691-710.