EXPLORING PREDICTORS OF READING COMPREHENSION FOR ENGLISH LANGUAGE LEARNERS

by

CAMERON GRACE ODDONE

(Under the Direction of Scott Ardoin)

ABSTRACT

Past reading research with English Language Learners (ELLs) in elementary school suggests that a number of component factors contribute to reading comprehension performance. Although comprehensive models of reading comprehension have been proposed, such as the Simple View of Reading (SVR), which explains reading comprehension as the combination of decoding and linguistic comprehension, further research is required to determine if these models adequately describe reading performance for ELLs. The purpose of the present study was two-fold. First, seeing as a recently developed version of a sentence repetition task would be used in the present study as a potential predictor of reading comprehension, the researcher aimed to closely examine the roles that working memory and oral language play in English and Spanish versions of this sentence repetition task. Second, the sentence repetition task was incorporated, along with a comprehensive measure of English Language Proficiency (ELP), decoding, and working memory, to determine whether these predictors could combine to explain greater variance in reading comprehension than the SVR did alone.

Eighty-four first- and second-grade ELL students were administered measures for all potential predictive variables, as well as reading comprehension. Results suggested that both
working memory and oral language contributed to performance on the English sentence repetition task, but that only working memory was a significant contributor to performance on the Spanish sentence repetition task. When investigating the SVR, findings indicated that although the SVR was an adequate reading comprehension model, adding ELP and working memory significantly improved the percentage of variance in reading comprehension that was explained. These findings have implications for development of universal screening and progress monitoring tools utilized within schools for ELLs struggling with reading comprehension. Additionally, these findings emphasize the importance of ELP and working memory in reading comprehension outcomes for ELLs.

INDEX WORDS: English Language Learners, Reading, Comprehension, Simple view of reading, Language proficiency, Children
EXPLORING PREDICTORS OF READING COMPREHENSION FOR
ENGLISH LANGUAGE LEARNERS

by

CAMERON GRACE ODDONE
B.A., Emory University, 2010
M.A., University of Georgia, 2015

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial
Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA
2017
EXPLORING PREDICTORS OF READING COMPREHENSION FOR ENGLISH LANGUAGE LEARNERS

by

CAMERON GRACE ODDONE

Major Professor: Scott Ardoin
Committee: Amy Reschly
Stacey Neuharth-Pritchett
Paula Mellom

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
August 2017
ACKNOWLEDGEMENTS

The research reported here was supported by the University of Georgia Department of Educational Psychology’s Owen Scott Scholars Award. The opinions expressed are those of the authors and do not represent views of the University of Georgia.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td></td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td></td>
<td>vii</td>
</tr>
<tr>
<td>1</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>ELLs and Reading Problems</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>A Model of Reading Comprehension for ELLs</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Within-Group Diversity of ELLs</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Purpose</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>REVIEW OF THE LITERATURE</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Acquiring Language and Literacy Skills in English</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Assessing ELLs in U.S. Schools</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Universal Screeners</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Within-Group Diversity of ELLs</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Measures of ELP</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>The Simple View of Reading</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Oral Language</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>Working Memory</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>Decoding</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>Summary and Purpose</td>
<td>44</td>
</tr>
</tbody>
</table>
Research Questions and Hypotheses ................................................................. 46

3  METHOD ........................................................................................................ 63
   Participants and Setting .............................................................................. 63
   Measures .................................................................................................... 65
   Procedure ................................................................................................... 72
   Procedural Fidelity and Inter-Observer Agreement ................................... 73
   Data Analyses ............................................................................................ 74

4  RESULTS ....................................................................................................... 79
   Sentence Repetition Task Findings .............................................................. 79
   Explaining Variance in Reading Comprehension ....................................... 80

5  DISCUSSION ................................................................................................. 83
   Limitations and Future Directions .............................................................. 87
   Summary and Implications ....................................................................... 88
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1: Descriptive Statistics</td>
<td>82</td>
</tr>
<tr>
<td>Table 2: Correlation Matrix</td>
<td>82</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

English language learners (ELLs) are a rapidly growing population in U.S. schools. Data reported by the National Center for Education Statistics indicated that the number of ELLs in U.S. public schools increased from the 2002-2003 school year (an estimated 4.1 million students, or 9%) to the 2011-2012 school year (an estimated 4.7 million students, or 10%; Aud et al., 2013). Nationally, the rate of increase for ELLs is well above the rate of increase for the overall school population. A 105% increase in the ELL population was noted between the 1990-1991 school year and the 2000-2001 school year in comparison to a 12% total enrollment increase (Kindler, 2002). Similar to nationwide data, the percentage of ELLs in the state of Georgia is growing rapidly. Between the 1994-1995 school year and the 2004-2005 school year, Georgia’s ELL student population grew 292%, placing it amongst the top 10 states with the largest ELL growth percentages (Payán & Nettles, 2008). According to the National Clearinghouse for English Language Acquisition (2015), in the 2013-2014 school year, Georgia’s public schools served approximately 98,600 ELLs in kindergarten through twelfth grade.

The ELL term generally refers to students whose native language is not English and who score low on a measure of English language proficiency (ELP). However, a highly diverse group of students are categorized under the ELL label and may vary on a number of characteristics such as immigration status, age, and native language spoken. Some ELLs are born in the U.S. and others immigrate to the U.S. at different points in their schooling careers. The largest numbers of students with limited English proficiency are enrolled in the elementary
grades (Kindler, 2002). Therefore, in a policy brief, the Foundation for Child Development argues for a high concentration of research efforts relating to ELL instruction and assessment to focus on the early school years (Russakoff, 2011). ELLs also represent a wide range of spoken languages; however, a report by the Migration Policy Institute (Ruiz, Hooker, & Batalova, 2015), utilizing data collected by the U.S. Census Bureau in 2013, indicated that Spanish was the most common language spoken by ELLs, with 71% of ELLs speaking Spanish as their home or first language. With this in mind, the present dissertation study targets Spanish-speaking ELLs at the early elementary school level.

Due to federal and state legislation enacted in the past 15 years, ELLs have begun to receive more attention from educators and researchers with the goal of determining appropriate teaching and assessment practices. The No Child Left Behind (NCLB) Act of 2001 emphasized that all students, including ELLs, must meet state-specified standards in reading and math every year in grades 3-8 and once in high school. In addition to holding schools, districts, and states accountable for ELLs’ academic skill levels, NCLB required that all states measure ELLs’ progress on ELP assessments and develop ELP standards if none already existed (Bailey & Huang, 2011). In December 2015, the Every Student Succeeds Act replaced NCLB. The Every Student Succeeds Act (2015) maintains commitment of federal funds to support ELLs’ academic and language education and enacts necessary changes to NCLB to more appropriately address the needs of ELLs.

Under NCLB, Adequate Yearly Progress was required for ELLs in the designated academic areas as well as in ELP. However, the separation of the academic goals under Title I and the ELP goals under Title III discounted the connection between developments in ELP and progression of academic skills measured in English (August, Gándara, Hakuta, Linquanti, &
O’Day, 2015). According to the National Education Agency’s (2008) Working Group on ELL Policy, a longitudinal analysis of state accountability data indicated that ELLs were more likely to reach academic performance standards once they became proficient in English, suggesting the presence of a link between the goals under Title I and Title III. Thus, in 2015, the Working Group on ELL Policy released a brief suggesting that ELP assessment should be considered in conjunction with academic content performance to ensure that ELLs exiting language assistance programs were prepared for academic course content presented in English (August et al., 2015). The Every Student Succeeds Act (2015) sanctioned these suggested changes by moving accountability measures for ELLs from Title III to Title I, indicating an increased emphasis on reporting academic performance in combination with ELP gains.

The Working Group on ELL Policy further suggested that accountability procedures should include setting reasonable timeframes for ELLs to reach full ELP and requirements for states to monitor both current and former ELLs throughout their school career (August et al., 2015; National Education Agency, 2008). Although NCLB required that states demonstrate progress for ELLs using ELP assessments, the Working Group on ELL Policy also recommended that states report the number of ELLs not meeting full ELP after five years in English language assistance programs (i.e., long-term ELLs). Along with reporting numbers of long-term ELLs, it was proposed that states set targets to decrease percentages of long-term ELLs, as well as plans within districts for providing high-quality supports to these students (August et al., 2015; National Education Agency, 2008). To more accurately assess ELL programs and provide support for students at all levels of English proficiency, a continuation of monitoring for former ELLs beyond the existing two-year period post-exit from language assistance programs was suggested. This would allow for better monitoring of student progress,
including graduation rates, and prevent the large fluctuation in ELL group performance observed as more proficient students were exited from the ELL group and less proficient students entered (August et al., 2015; National Education Agency, 2008). Changes under the Every Student Succeeds Act (2015) now require states to report numbers of long-term ELLs and continue monitoring former ELLs’ academic performance for up to four years.

**ELLs and Reading Problems**

Inclusion of ELLs in state assessment and accountability procedures drew greater attention to the academic and language needs of ELLs as well as the achievement gap between ELLs and native English speakers (EL1s; August et al., 2015; National Education Agency, 2008). The 2015 results from the National Assessment of Educational Progress indicated that despite the reading achievement progress evident for ELLs since 1998, ELLs continue to underperform in comparison to their EL1 peers (U.S. Department of Education, 2015). The reading measures revealed that ELLs were behind EL1s by 37 scale score points (range 0-500) in fourth grade and 45 points in eighth grade. Currently, 68% of fourth-grade ELLs are reading below a basic reading level, a statistic that has not changed significantly since 2005 (U.S. Department of Education, 2015).

Findings from national and state standardized assessments indicating that ELLs consistently fall behind EL1s in reading achievement are corroborated by research findings in this area. Research suggests that even with quality instruction, ELLs often do not match EL1s in higher-level reading tasks such as reading comprehension (for a review, see August & Shanahan, 2006). Although ELLs often fall behind EL1s in reading comprehension, they frequently reach equivalency in word-level abilities (e.g., phonological awareness, spelling, decoding), provided they have sufficient exposure to quality teaching techniques (August & Shanahan, 2006).
In order to combat poor comprehension outcomes for ELLs, it is important to identify ELLs who may be at-risk for reading difficulties at an early age and provide necessary instruction and interventions. Through universal screening processes beginning at the start of formal schooling, ELLs who are falling behind grade-level expectations can be identified. However, questions regarding the utility of various universal screening measures for ELLs often arise (Gersten et al., 2007). To more fully understand which type of universal screening literacy measures should be used with ELLs, a more complete understanding of successful reading predictors for ELLs is required.

The present dissertation study investigates factors that affect ELLs’ reading comprehension. Chapter 2 provides readers with an in-depth review of extant research on the measurement of reading skills of ELLs and how factors such as ELP, oral language, working memory, and decoding may differentially impact reading comprehension for ELLs in comparison to their EL1 peers. The extensive research base on factors impacting reading comprehension for EL1s will be summarized and an extension of this research including ELLs will be highlighted.

A Model of Reading Comprehension for ELLs

The simple view of reading (SVR) is a widely accepted and well researched model of reading comprehension which suggests that reading comprehension is best predicted by a combination of decoding and listening comprehension (e.g., Catts & Weismer, 2006; Tilstra, McMaster, Van den Broek, Kendeou, & Rapp, 2009). The SVR was originally proposed as a model for EL1s (Gough & Tunmer, 1986) and then was applied to ELLs (Gottardo & Mueller, 2009; Hoover & Gough, 1990; Proctor, Carlo, August, & Snow, 2005; Verhoeven & van Leeuwe, 2012). The present study extends research in the area of predictive models for reading
comprehension in ELLs by expanding the SVR to include working memory and ELP, attempting to capture a broader picture of young ELLs’ emerging reading comprehension and the component skills that predict success in this area.

For young ELLs, decoding (e.g., Nakamoto, Lindsey, & Manis, 2007), oral language (e.g., Gottardo & Mueller, 2009), working memory (e.g., Swanson, Orosco, & Lussier, 2015), and ELP (e.g., Kieffer, 2008) are important indicators for reading comprehension. ELP is a key factor ignored in a majority of the existing literature. The tendency to underestimate ELL intragroup variability by failing to disaggregate factors such as ELP is a limitation of the existing research examining models of ELL reading comprehension (Artiles, Rueda, Salazar, & Higareda, 2005; Gutiérrez & Vanderwood, 2013). The present study addresses this limitation by including ELP in the model of reading comprehension.

Within-Group Diversity of ELLs

Research indicates that ELP level is important to consider when assessing ELLs’ reading skills (Artiles, Rueda, Salazar, & Higareda, 2005; August et al., 2015; Gutiérrez & Vanderwood, 2013); however, a majority of studies have treated the ELL group as homogenous and have not separated the group by ELP level. Failing to recognize the diversity within the ELL group causes problems relating to external validity because it is unclear to which subcategory of ELLs the results may apply. Examining within-ELL distinctions has immense practical value for educators, because with more detailed information they can better understand appropriate expectations for subgroups of ELLs and more accurately identify when a student is at-risk (Gutiérrez & Vanderwood, 2013).

Few studies have acknowledged the ELP heterogeneity of the ELL group by examining the effects of ELP on reading performance in English, but the research that does exist will be
discussed in Chapter 2. In order to address the lack of research in this area, the current study will systematically include ELLs with a range of ELP levels to assess how ELP impacts performance on an outcome measure of reading comprehension in English.

**Purpose**

Due to the staggering underperformance of ELLs on reading comprehension assessments, it is important for researchers and educators to better understand factors that may contribute to this deficit. The present study aims to investigate potential contributing factors to ELL reading comprehension outcomes, including decoding skills, oral language ability (in Spanish and English), working memory skills, and ELP. This dissertation study will examine how these independent variables uniquely and cumulatively predict reading comprehension for ELLs in first and second grade, a time period in which measurement of these skills is appropriate and a range of ability is expected.

Chapter 2 presents a comprehensive review of the literature and rationale for the dissertation study. The chapter will first review extant literature on reading skill deficits for elementary-aged ELLs. Next, the chapter will highlight difficulties presented when assessing ELLs in U.S. schools, including the use of accommodations and testing in English versus a native language. These topics are critical to discuss because the researcher is interested in parsing out various constructs (e.g., ELP) that may unintentionally be measured when testing in English for students at lower ELP levels. Importantly, the chapter will also cover the limited research available discussing the impact of ELP on reading ability. Then, an in-depth analysis of research in the areas of contributing factors to ELL reading comprehension will be discussed. The SVR will be presented as an example of a comprehensive model, followed by a thorough
investigation of the specific component measures that will be utilized within the present study: ELP, oral language, working memory, and decoding.

The methodology and analyses for the current study are presented in Chapter 3. First- and second-grade ELLs, who speak Spanish as their first language, were recruited for the study. A crucial aspect of participant recruitment involved systematically acquiring students with a range of ELP levels, based on scores from a school district-administered measure of ELP. Participants completed measures of reading comprehension, working memory, decoding, and an English and a Spanish version of a sentence repetition task assessing oral language abilities. Additionally, the researcher gathered key demographic information (e.g., age, gender, preschool attendance) about the participants from their school files. Ultimately, analyses were conducted to investigate whether adding working memory and ELP improved the predictive value of the SVR; but first, analyses investigating the underlying skills involved in the sentence repetition task were investigated. Findings (a) distinguish the differential roles of oral language and working memory in English and Spanish administrations of the sentence repetition task, (b) reveal whether the English or Spanish administration of the sentence repetition task is a better representation of oral language ability, and (c) identify whether working memory and ELP add predictive value to the SVR and explain variance in reading comprehension.
References


U.S. Department of Education, Institute of Education Sciences, National Center for Education

CHAPTER 2
REVIEW OF THE LITERATURE

National standardized testing and a substantial empirical research base indicate that a large proportion of ELLs struggle to learn to read in English. However, it is unclear to what extent these reading difficulties are a result of limited ELP and to what extent they are due to learning disabilities (for a review, see Klingner, Artiles, & Barletta, 2006). This distinction is important for special education eligibility teams within schools to consider as they seek to determine which students should qualify for services. Although typical reading development patterns and risk indicators for EL1s have been widely studied, extant research with ELLs is limited. Without this necessary information, it is difficult for school staff to determine which ELLs are at-risk for reading problems and which students should qualify for special education services due to inadequate response to interventions.

In a systematic review of the empirical literature, Klingner et al. (2006) investigated factors that could potentially differentiate ELLs who struggle to read due to inadequate ELP and those who struggle due to an actual reading disability. Conclusions from the review indicated that more descriptive work is needed to better understand the specific learning needs for subgroups of ELLs. The authors emphasized that the majority of studies have not disaggregated their ELL populations by ELP level, which leads to confusion for educators over whether the research findings can apply to the ELLs with whom they are working. Therefore, Klingner et al. called for additional research in the area of ELLs’ literacy skills so that educators can better
understand the characteristics, development, and risk factors for ELLs who have difficulty learning to read in English.

**Acquiring Language and Literacy Skills in English**

Acquiring literacy skills in a second language is a complicated process. Although a number of EL1s struggle with early literacy skills, a larger proportion of ELLs tend to have difficulty in this area. One proposed reason for this difference is that ELLs must simultaneously develop their English language skills alongside their reading skills, causing additional challenges not faced by EL1s (Gersten et al., 2007).

Although ELLs are learning to read in a second language, some aspects of evidenced-based early reading instruction for ELLs remain the same as those recommended for their EL1 peers. For example, research emphasizes the importance of explicitly teaching certain literacy components such as decoding, spelling, writing, fluency, and comprehension to both ELLs and EL1s (August, McCardle, & Shanahan, 2014; August & Shanahan, 2006, 2010). Some suggestions for enhancement or alteration of reading instruction specifically for ELLs include incorporating students’ native language into instructional routines, using visual cues to clarify meaning, and providing extra emphasis on vocabulary development, comprehension checks, and clear presentation of ideas (August et al., 2014; August & Shanahan, 2010). With the appropriate evidenced-based instruction, research reviews by August and Shanahan (2006, 2010) suggest that ELLs and EL1s perform equally well on word-level reading measures that assess phonological awareness, spelling, and decoding. However, ELLs typically fall behind on higher-level reading comprehension and overall reading achievement measures.

The degree to which ELLs are perceived to be on par or behind EL1s may depend, in part, on the type of measure administered and the subsequent conclusions that can be drawn from
the scores. For example, extensive research completed with EL1s indicates that there is a strong connection between measures of oral reading fluency (ORF) and overall reading achievement (Fuchs, Fuchs, Hosp, & Jenkins, 2001; Reschly, Busch, Betts, Deno, & Long, 2009). Consequently, ORF measures (i.e., Curriculum based measurement - Reading), which are relatively quick and simple to administer, are considered to be good universal screening tools for identifying EL1s with reading difficulties and progress monitoring their response to intervention (Deno, 2003). However, a more limited body of research suggests that the relationship between ORF and reading achievement may not be as direct for ELLs (Crosson & Lesaux, 2010; Klein & Jimerson, 2005; Quirk & Beem, 2012). ELLs may develop text-reading skills without fully understanding what they are reading, indicating that ORF measures may not accurately represent ELLs’ overall reading abilities. ELLs’ slower pace of development for reading comprehension in comparison to ORF may be connected to the complexity of language involved in reading and interpreting text. The academic language present in classroom reading is generally different and more difficult to comprehend than conversational language.

The distinction between conversational and academic language was first proposed by Cummins (1979; 1981b). Cummins coined the terms basic interpersonal communication skills (BICS) and cognitive academic language (CALP) to refer to two different categories of language. BICS refers to the conversational English skills that ELLs generally acquire through discourse with peers and adults who speak English. Alternatively, the term CALP encompasses the academic language that is specific to classroom instruction and academic reading and often is less context embedded than conversational language, providing fewer cues for meaning (Cummins, 1981b). CALP can be separated into three subcategories based on the context in which the vocabulary appears: general academic vocabulary, context-specific academic
vocabulary, and specialized academic vocabulary (Bailey & Huang, 2011). Cummins (1981a) stated that while BICS often develops within the first two years of exposure to English, CALP generally takes five to seven years to fully develop. This leads to challenges as ELLs attempt to catch up to their EL1 peers in understanding academic language used in the school setting (Cummins, 2008). Research findings from Cummins (1984) indicated that proficiency with BICS can be mistakenly likened to an ELL’s readiness for academic learning and assessment in English, when in fact more time is required for CALP to develop.

The BICS/CALP distinction helped draw attention to the specific type of English language skills that were required for success in school. Abedi (2007) explained that without effective use of academic language, ELLs are unable to extract meaning from their academic content classes and progress satisfactorily. This knowledge regarding the importance of academic English gave ELP test developers more specific focus in assessment targets. Modern measures of ELP heavily emphasize the measurement of academic English across various academic subject areas (Abedi, 2007). In addition to providing useful information to ELP test developers, knowledge regarding the development of academic language is also valuable for academic achievement test developers. With the understanding that ELLs take more time to acquire academic language, test developers can take certain precautions to support fair test development such as reducing unnecessarily complicated academic vocabulary. Moreover, once the tests are developed, certain test-taking accommodations can be employed to further assist ELLs (Abedi, Hofstetter, & Lord, 2004).

Assessing ELLs in U.S. Schools

Testing accommodations. The Committee on Educational Excellence and Testing Equity (under the direction of the National Research Council and the Board on Testing and
Assessment) published a summary report on testing ELLs in U.S. schools (National Research Council, 2000). This report was intended to review the current research base and provide recommendations to policy makers. The report specified that in order to offer ELLs the teacher attention and school resources they require, it was essential to include these students in accountability testing. Although committee members agreed that teachers and schools should be held accountable for the education that they provide to ELLs, it was acknowledged that it is no easy feat to accurately assess ELLs’ academic ability. Many assessments administered to ELLs in English may in fact measure ELP as opposed to the targeted academic skills. Testing accommodations can aid in allowing ELLs to demonstrate their knowledge; however, accommodations can often be inappropriately applied, resulting in inaccurate test scores (National Research Council, 2000).

Testing accommodations for ELLs include alterations to the testing process or the test itself. These accommodations provided across subject areas are intended to allow ELLs to overcome the language barrier and demonstrate their knowledge without giving an advantage over students who do not receive the accommodations (Abedi et al., 2004). Abedi et al. (2004) reviewed the literature on test accommodation strategies for ELLs. The authors recommended that when contemplating potential accommodations, researchers and educators consider if the accommodation (a) is effective in reducing the performance gap between ELLs and EL1s, (b) alters the construct being measured, (c) has a differential impact based on background variables (e.g., ELP level), and (d) is feasible to implement (e.g., low cost). Conclusions from the review indicated that rewording test items to reduce unnecessary linguistic complexity and providing customized dictionaries (e.g., only including words that appear in the test) were promising assessment accommodation strategies for ELLs. However, Abedi et al. (2004) cautioned against
uniformly applying one type of accommodation to all ELLs, as some accommodations were more effective than others based on characteristics of the students.

An updated meta-analytic review from Kieffer, Lesaux, Rivera, and Francis (2009) built upon the findings from Abedi et al. (2004). Kieffer et al. (2009) evaluated the effectiveness and validity of seven common accommodations provided to ELLs on large-scale assessments: (a) reduced linguistic complexity, (b) English dictionaries or glossaries, (c) bilingual dictionaries or glossaries, (d) tests in the native language, (e) dual language test booklets, (f) dual language questions for English passages, and (g) extra time. Ultimately, Kieffer et al. determined that providing access to English dictionaries or glossaries was the only accommodation that had an overall positive effect on ELLs’ testing outcomes across the studies analyzed. However, even this accommodation reduced the ELL/EL1 achievement gap by only moderate values of 10% to 25%. As such, the present testing accommodations available to ELLs were not found to be very effective. Operating under the assumption that a valid test accommodation for ELLs should have no significant effect on achievement scores of EL1s, Kieffer et al. indicated that the accommodations studied were valid.

**Testing in English vs. Spanish.** The translation of English assessments into the student’s native language is a testing accommodation that is extensively debated. In addition to the difficulty of equating tests in terms of content, difficulty level, reliability, and validity, creating translated tests is time consuming and expensive; thus, translated tests are not widely used (Abedi et al., 2004; Hofstetter, 2003). Beyond the problems associated with test translation, the debate involves the appropriateness of using an assessment originally developed in the ELL’s native language. The crux of the debate generally centers on the ELL’s foundational literacy and language skills in their native language and their language of instruction. If the student learned
the subject matter through English instruction, testing in their native language may confuse the child and harm rather than help scores (Abedi et al., 2004; Hofstetter, 2003; Kieffer et al., 2009; Townsend & Collins, 2008). However, when ELLs have strong foundational reading and language skills in their first language, in addition to academic content area instruction in their first language, then testing in their first language is believed to be appropriate (Abedi et al., 2004; Hofstetter, 2003; Kieffer et al., 2009).

The question remains, however, as to the suitability of testing in English if ELP is low, despite receiving content area instruction in English. Abedi (2002) suggested that tests administered in English and normed on EL1s may inadvertently be measuring ELP as opposed to the desired content area knowledge for students with low English proficiency. Additionally, Abedi indicated that the impact of ELP on test outcomes varies by content area, with the greatest performance differences evident in areas with a high linguistic demand (e.g., reading) as opposed to those with a low linguistic demand (e.g., math). However, these academic areas can also be linked, as ELLs’ English reading proficiency has significant associations ($p < .01$) with their performance on other content area tests such as math (Hofstetter, 2003).

It may be that more instructional time is simply required for ELLs to gain the necessary foundational language skills before they can demonstrate their knowledge on English standardized tests. This notion is written into the federal law, as ELLs who have not attended U.S. schools for a full academic year are exempt from having their academic scores on state assessments count for Adequate Yearly Progress purposes (Abedi et al., 2004; National Education Agency, 2008). With time allowed for language acquisition, educators can reasonably expect that foundational skills (e.g., phonological awareness, syntactic awareness) demonstrated in the student’s first language would transfer to their second language (Durgunoğlu, 2002;
Lindsey, Manis, & Bailey, 2003). It should be emphasized, however, that if ELLs have not acquired foundational skills in their native language and are also demonstrating below average performance in their second language, adjustments in instruction may be required (Gersten et al., 2007). With time and appropriate instruction, ELLs have the ability to make vast gains in both ELP and academic performance; however, in order to identify which ELLs are progressing accordingly and which ELLs require more specialized attention, more information is required regarding the best methods for screening ELLs for academic difficulties.

**Universal Screeners**

Response to intervention (RTI) is a multi-tiered service delivery model in which students are provided with increasingly intense interventions based on their movement through the tiers of services. One of the assumptions of RTI is that all students are provided with effective, research-based instruction in their general education classroom. However, the quality and appropriateness of Tier 1 practices have been called into question when considering how they are implemented with ELLs (Brown & Doolittle, 2008; Thorius & Sullivan, 2013). Although more information is needed on effective instructional practices for ELLs, the focus within the present study will move beyond general education instruction to another component in the RTI process: universal screening.

A key element of RTI is the early identification of students who are and are not at-risk for reading problems through universal screening of all students beginning at an early age. Universal screening allows educators to make more accurate decisions regarding instructional support. Similar to findings for EL1s, early identification and intervention is beneficial for ELLs (Haager & Windmueller, 2001; Lesaux & Siegel, 2003). Recommendations from the Institute of Educational Sciences Practice Guide tailored towards literacy education for ELLs in the
elementary grades (Gersten et al., 2007) indicated that formative evaluations focusing on phonological processing, letter knowledge, and word and text reading accurately predict later reading performance across skill areas. Additionally, the practice guide suggests that the same English literacy screening measures can be used with EL1s and ELLs (Gersten et al., 2007).

Research findings discussed earlier regarding the value of matching the language of the test to the language of instruction (e.g., Abedi et al., 2004; Hofstetter, 2003; Kieffer et al., 2009) also apply when considering universal screeners and progress monitoring tools. Townsend and Collins (2008) found that Spanish-speaking ELLs who received literacy instruction in English showed greater growth on English versions as opposed to Spanish versions of screening measures of letter naming fluency, letter-word identification, and nonsense word fluency across the first-grade year. Furthermore, the English versions of the reading screening measures were more sensitive in identifying ELLs who were at-risk for reading difficulties in English (Townsend & Collins, 2008).

One type of literacy screening measure that is commonly used once students gain a basic level of text-reading ability is curriculum-based measurement in reading (CBM-R). CBM-R requires students to read text aloud for one minute while an examiner counts the number of words read correctly. CBM-R was found to be an appropriate measure of oral reading fluency for ELLs (Al Otaiba et al., 2009; de Ramirez & Shapiro, 2006). In a study by de Ramirez and Shapiro (2006), results indicated that CBM-R, administered in both English and Spanish, was sensitive to growth from first grade through fifth grade for Spanish-speaking ELLs in a bilingual program.

Beyond utilizing CBM-R as a direct measure of oral reading fluency, some studies also suggest a strong relationship between CBM-R and reading achievement measures for ELLs as
well as EL1s (Baker & Good, 1995; McMaster, Wayman, & Cao, 2006; Muyskens, Betts, Lau, & Marston, 2009; Wiley & Deno, 2005). However, problems can arise when CBM-R scores are extrapolated to reading achievement scores for ELLs (Crosson & Lesaux, 2010; Klein & Jimerson, 2005; Quirk & Beem, 2012). Research suggests that bias is present when using CBM-R scores to predict reading comprehension scores for ELLs, as CBM-R scores are more representative of the ELLs’ word reading ability than overall reading proficiency (Crosson & Lesaux, 2010; Klein & Jimerson, 2005). Furthermore, results from Quirk and Beem (2012) demonstrated that a majority (55%) of their ELL participants earned ORF scores that were significantly higher than would be predicted based on their low reading comprehension scores. The researchers therefore cautioned against using reading fluency scores as a proxy for reading comprehension measures for ELLs (Quirk & Beem, 2012).

Computer adaptive measures of broad reading abilities linked to grade-level standards are another type of screening measure that can be used with students in kindergarten through twelfth grade (Christ et al., 2014; Northwest Evaluation Association, 2009; Renaissance Learning, 2010). Christ et al. (2014) developed the Adaptive Reading (aReading) assessment as part of the FastBridge Learning suite of assessment tools. The aReading items developed for students in kindergarten through fifth grade target concepts of print, phonological awareness, phonics, vocabulary, and comprehension. Inclusion of these target domains is related to research that indicates that these foundational components of reading are most predictive of future reading skills (Christ et al., 2014; Stanovich, 1986; Vellutino & Scanlon, 1991). Similar emphasis is placed on the importance of screening for foundational reading skills with ELLs in an attempt to predict future reading success (Gersten et al., 2007).

It is important to note that simply screening students for reading difficulties is not
sufficient for effecting change. These formative assessments are meant to guide instructional
decisions through provision of interventions to students who require these services to meet
reading achievement expectations (Gersten et al., 2007). Similar to best-practice
recommendations for assessment measures, the interventions provided to students should match
their language of instruction and be comprehensive across the components of reading
achievement (Linan-Thompson, Vaughn, Prater, & Cirino, 2006). With this type of intervention,
Linan-Thompson et al. (2006) found that ELLs made substantial reading gains, over and above
those observed in ELLs receiving services through the existing intervention programs for
struggling readers available in the schools. However, because ELLs constitute an extremely
diverse group of students with varying educational needs, more information is required regarding
within-group differences for ELLs so that assessment and intervention can be more targeted to
specific sub-groups of ELLs.

Within-Group Diversity of ELLs

Although it is recognized that ELLs in U.S. schools are an extremely diverse population,
as they speak a wide variety of languages, have different degrees of exposure to English
language and U.S. culture, and come from a range of socioeconomic backgrounds, a majority of
studies investigating ELLs’ academic abilities include all ELLs in one group. In doing so,
Artiles, Rueda, Salazar, and Higareda (2005) argued that these studies are overestimating the
homogeneity of ELLs, and that instead, researchers should disaggregate the ELL group based on
factors such as ELP level. By recognizing intragroup dissimilarities, we can better identify
subpopulations of ELLs and understand their unique learning needs.

A better understanding of the within-group diversity of ELLs can be gained by examining
a number of demographic variables such as cultural exposure, language type, and socioeconomic
status. For example, Betts, Bolt, Decker, Muyskens, and Marston (2009) illustrated that the amount of time that an ELL’s family has lived in the U.S. can affect reading achievement. Although Betts et al. suggested that cultural exposure was a significant predictor of reading outcomes, language type (Somali or Spanish) was not a driving factor for shaping young ELLs’ reading abilities (Betts et al., 2009). Similar to cultural exposure, socioeconomic status is a demographic variable linked to a number of ELL reading outcome measures (D'Angiulli, Siegel, & Maggi, 2004). In fact, socioeconomic status may be a potential confound with measures of primary language, ELP, and reading achievement, as there tends to be a higher concentration of ELLs residing in disadvantaged communities (D'Angiulli et al., 2004; Kieffer, 2008; Roberts, Mohammed, & Vaughn, 2010).

Beyond investigating differences in demographic variables, differences in language skills, such as ELP, that may affect reading performance should also be acknowledged within the ELL population. Al Otaiba et al. (2009) conducted a large-scale study that recognized the heterogeneity of ELP within the ELL groups and considered the role that ELP level plays in English reading skill. Al Otaiba et al. used second- and third-grade Latino students from high-poverty schools, receiving English-only reading instruction, grouped into three ELP categories: students receiving English as a second language (ESL) services \( (n = 2,182) \), ESL-exited students \( (n = 965) \), and students who never needed ESL services \( (n = 1,857) \). The large sample size allowed researchers to make comparisons within ELP group based on special education status (i.e., students not receiving special education services, students with learning disabilities, and students with speech and language impairments). CBM-R passages were administered to measure students’ ORF. The results indicated that ELLs with lower ELP (i.e., ESL students) performed at a lower level on ORF measures than their Latino peers with higher ELP (i.e., ESL-
exited students, students who never needed ESL services) across all special education designations. Therefore, the authors concluded that ELP was an important indicator of ORF skills. However, a weakness of this study was that the students’ history of ESL services was used as a designation of ELP, which is a somewhat rudimentary measure, as there is a wide variability of ELP levels observed within the group of students receiving ESL services. Comparing performance on a comprehensive measure of ELP to performance on a reading outcome measure could have improved the study.

Similar to Al Otaiba et al. (2009), Gutiérrez and Vanderwood (2013) recognized the importance of exploring how differences in ELP may affect English reading performance for ELLs. However, Gutiérrez and Vanderwood addressed limitations of previous research (Al Otaiba et al., 2009) by using a comprehensive measure of ELP to separate ELLs into groups based on ELP level. Furthermore, Gutiérrez and Vanderwood investigated the relationship between ELP and English reading ability as well as rate of growth in early literacy skills. The study included 260 second-grade Latino ELLs stratified by ELP level (i.e., Beginning, Early Intermediate, Intermediate, Early Advanced, Advanced) based on results from the district-administered California English Language Development Test (California Department of Education, 2006). To measure literacy skills, Dynamic Indicators of Basic Early Literacy Skills (Good & Kaminski, 2002) measures of ORF, phonological awareness (Phoneme Segmentation Fluency), and decoding (Nonsense Word Fluency) were administered in the fall, winter, and spring of the participants’ second-grade year.

Through growth curve analyses, Gutiérrez and Vanderwood (2013) determined that in the fall of second grade, ELLs from each of the five ELP levels significantly differed ($p < .005$) from each other in initial ORF ability, with the exception of the two highest levels of ELP: Early
Advanced and Advanced. In other words, a higher ELP level was associated with reading words more quickly and accurately from text aloud in English at the start of second grade. However, in growth rate comparisons amongst students at the five ELP levels, Gutiérrez and Vanderwood found that the only significant difference in ORF growth across the second-grade year existed for students at an Advanced ELP level in comparison to students at a Beginning ELP level. Students with Advanced ELP levels displayed significantly steeper \((p < .005)\) ORF growth slopes than students with Beginning ELP levels. ORF growth rates were 0.82, 0.95, 0.97, 1.10, and 1.30 words per week for students with Beginning, Early Intermediate, Intermediate, Early Advanced, and Advanced ELP levels, respectively.

Although there were significant differences in growth rates in ORF between students at Beginning and Advanced ELP levels, no significant differences were detected in growth rates on phonological awareness or decoding across the five ELP levels (Gutiérrez & Vanderwood, 2013). With regard to initial performance in the fall of second grade, Gutiérrez and Vanderwood (2013) found that students in the Beginning ELP level significantly differed \((p < .01)\) from students in the Early Advanced and Advanced ELP levels on measures of decoding. Gutiérrez and Vanderwood also investigated initial differences in phonological awareness between ELLs at various ELP levels. The authors found that although students with Advanced ELP were able to correctly identify approximately eight more segmented sounds on average than students with Beginning ELP, this difference was not significant at the \(p < .01\) level.

Overall, results from Gutiérrez and Vanderwood (2013) indicated that it was important to desegregate the ELL group and consider students’ ELP level when interpreting ELLs’ ORF and decoding screening results. The fact that the ELLs in this study were grouped based on scores from the state measure of ELP (i.e., California English Language Development Test), which was
a comprehensive measure with evidence supporting reliability and validity, was a strength of the study. Although results from this study were valuable because reading skill comparisons between ELLs at different ELP levels were possible, the study could have been strengthened by including a larger sample of ELLs at the Advanced ELP level. The Advanced ELP group included only 10 students as compared to other ELP categories, which ranged in size from 30 to 90 students. Additionally, the study was limited in that a measure of reading comprehension was not administered. Instead, the extent to which ELP affected three component skills of reading was assessed, leaving the reader wondering what the affect of ELP may be on participants’ understanding of text.

A study by Kieffer (2008) moved beyond investigation of the impact of ELP on ORF to explore the impact of ELP on a reading achievement measure. The comprehensive reading achievement outcome measure used in this study was a methodological strength, as it was developed by a panel of experts to include items from published standardized tests that directly targeted basic reading skills, vocabulary, and five types of reading comprehension skills (Kieffer, 2008). In this study, students who lived in homes where a language other than English was spoken, but entered kindergarten with full ELP \( (n = 746) \), were compared to ELL \( (n = 1,134) \) and EL1 \( (n = 15,362) \) students. The researchers found that ELLs had slower growth rates in English reading achievement from kindergarten to fifth grade whereas students with language minority backgrounds that had full ELP upon school entry performed similarly to EL1 students on reading achievement measures across the elementary grades. This difference in growth rate led to large differences in reading achievement by the fifth grade, with ELLs falling behind their peers in comparison groups. However, the effects of initial ELP were reduced from large to moderate when demographic risk factors (e.g., SES) were controlled. In sum, findings from Kieffer, in
addition to Gutiérrez and Vanderwood (2013) and Al Otaiba et al. (2009), suggest that ELP should be considered when assessing English reading outcomes, as differences in level and growth can arise between ELL groups at varying levels of ELP.

**Measures of ELP**

In order to accurately disaggregate ELL groups by ELP, valid and reliable measures of ELP are required. With the implementation of NCLB in 2001, states were required to measure ELP for all ELLs on an annual basis, and thus national efforts to develop new ELP assessments began (Abedi, 2007). New ELP assessments are aligned to ELP standards, emphasize academic English (i.e., CALP), are comparable across grades, and are tiered within grade levels. In a review of the currently available ELP assessments, Abedi (2007) explained that scores from ELP assessments allow schools to determine placement in English language assistance programs, monitor students’ language acquisition progress, and make informed decisions regarding instructional objectives and exit from support programs.

One of the summative, criterion-referenced ELP assessments developed in the new wave of measures following NCLB was Assessing Comprehension and Communication in English State to State for English Language Learners (ACCESS; WIDA, 2016). ACCESS was first developed and field tested in 2004 and was employed in three states beginning in 2005 (Abedi, 2007). ACCESS is now a widely used measure of ELP. During the 2013-2014 school year, ACCESS was administered to 1,372,806 students in grades K-12 from 33 states, including Georgia (Center for Applied Linguistics, 2015a).

Although state measures of ELP are widely used within school systems across the U.S., limited research utilizing these measures exists beyond trials to assess reliability and validity. In fact, although other state measures of ELP have been used in research studies (e.g., Beal, Adams,
& Cohen, 2010; Gutiérrez & Vanderwood, 2013), to date no studies have utilized ACCESS scores. Reports (Abedi, 2007; Center for Applied Linguistics, 2015a; Gottlieb & Kenyon, 2006; Kenyon, 2006) demonstrate the strong reliability and validity of ACCESS scores; however, the academic literature is extremely limited in utilizing this credible data source. More research employing ACCESS data is needed to better understand the impact that ELP level may have on various academic outcomes, including reading. The present study will use ACCESS scores as a predictor variable that may help to explain variance in reading outcome scores for ELLs.

**The Simple View of Reading**

The simple view of reading (SVR) describes reading comprehension as the product of decoding and linguistic comprehension (Gough & Tunmer, 1986). Therefore, when deficits in either decoding or linguistic comprehension are present, difficulties with reading comprehension arise. Within the SVR model, decoding is defined as efficient recognition of novel words and linguistic comprehension is defined as understanding of orally presented information (Hoover & Gough, 1990). Dividing the complex process of reading into two parts (i.e., decoding, linguistic comprehension) simplifies the paradigm of what is required for successful reading, but, it does not discount the overall complexities involved in the reading process (Hoover & Gough, 1990). The SVR garnered wide support as a model for reading comprehension with EL1s (e.g., Catts, Herrera, Nielsen, & Bridges, 2015; Catts & Weismer, 2006; Joshi & Aaron, 2000; Tilstra, McMaster, Van den Broek, Kendeou, & Rapp, 2009), and it was tested with ELLs and found to be an appropriate predictive model for reading comprehension with this population as well (Gottardo & Mueller, 2009; Hoover & Gough, 1990; Proctor, Carlo, August, & Snow, 2005). The studies investigating the applicability of the SVR with ELLs used primarily Spanish-speaking participants across the elementary grades.
Initial support for the use of the SVR model with ELLs came from Hoover and Gough (1990), who conducted a longitudinal study tracking students from first through fourth grades. Results indicated that across grades, a significant portion of variance in reading comprehension was explained by the product of decoding and linguistic comprehension. Additionally, the relative contribution of decoding as a determinant of reading comprehension appeared to be stronger in grades one and two, while linguistic comprehension was stronger in grades three and four. These results supported the view that reading comprehension could be simply represented by the combination of ability in decoding and linguistic comprehension for ELLs in the elementary grades (Hoover & Gough, 1990).

Proctor et al. (2005) also utilized an ELL sample to test the SVR model but expanded upon Hoover and Gough’s (1990) study by adding measures of component skills under the categories of decoding and linguistic comprehension. Measures of decoding included pseudoword reading and real-word fluency, and measures of linguistic comprehension included vocabulary knowledge and listening comprehension. A sample of 135 fourth-grade ELLs were involved in this study. Results suggested that listening comprehension was the strongest determinant of reading comprehension. The decoding measures were also significantly correlated with reading comprehension, but less so than the linguistic comprehension measures, potentially due to the restricted range of decoding scores observed in the relatively older and more advanced fourth-grade sample (Proctor et al., 2005).

Gottardo and Mueller (2009) added to the SVR ELL literature by administering Spanish and English versions of decoding and linguistic comprehension measures to their participants. These researchers utilized a sample of 131 first-grade ELLs who received all instruction in English. Results indicated that only the English measures of decoding and linguistic
comprehension were significantly related to English reading comprehension. The lack of a direct connection between the Spanish measures and English reading comprehension was attributed to the fact that the participants received their reading instruction entirely in English (Gottardo & Mueller, 2009).

Although the SVR ELL literature utilized participants drawn from a range of instructional programs across the country that employed Spanish and English literacy instruction to varying degrees (Gottardo & Mueller, 2009; Hoover & Gough, 1990; Proctor et al., 2005), some consistent conclusions exist across studies. Uniformly, findings indicated that English measures of decoding and linguistic comprehension were each strong predictors of reading comprehension; however, models that contained both decoding and linguistic comprehension were found to have the best fit (Gottardo & Mueller, 2009; Hoover & Gough, 1990; Proctor et al., 2005).

A limitation of the studies exploring the application of the SVR with ELLs (e.g., Gottardo & Mueller, 2009; Hoover & Gough, 1990; Proctor et al., 2005) is that they present the ELLs as one homogenous group without acknowledging important within-group variations, such as ELP. Furthermore, they have failed to investigate how variations in ELP may affect the predictive value of decoding and linguistic comprehension on reading comprehension. Given that significant differences were observed in reading outcome measures between groups of ELLs with different ELP levels (Al Otaiba et al., 2009; Gutiérrez & Vanderwood, 2013; Kieffer, 2008), it would seem that ELP may be an important determinant of reading ability that should be included in any comprehensive predictive model.
Oral Language

**Oral language and reading.** The SVR indicates that linguistic comprehension is a critical predictor of reading comprehension, and some SVR models (Catts & Weismer, 2006; Gottardo & Mueller, 2009) further deconstruct linguistic comprehension to include oral language abilities. Similarly, oral language is an important component in comprehensive assessments of ELP, which generally cover domains of reading, writing, listening, and speaking; with the listening and speaking scores combining to form an oral language composite (Abedi, 2007; Center for Applied Linguistics, 2015b). Oral language is not uniformly defined across research studies, but most researchers agree it is comprised of receptive and expressive language skills and can be separated into multiple components (Abedi, 2007; Roth, Speece, & Cooper, 2002). Components of oral language include phonology (sound in language), semantics (word meaning knowledge and vocabulary), syntax (word order and grammatical rules), morphology (identifying units of words), and pragmatics (social use of language; Arañas & Christ, 2014a; Catts, Fey, Zhang, & Tomblin, 1999; Storch & Whitehurst, 2002). Similar to the variety seen in definitions of oral language, a wide variety of measurement systems for oral language are utilized across studies (e.g., sentence repetition tasks [Arañas & Christ, 2014a], tests assessing vocabulary knowledge and syntactic processing [Gottardo & Mueller, 2009], narrative language samples [Miller et al., 2006]).

Although studies vary in oral language definitions and measurement systems, some general conclusions exist regarding the relationship between oral language and early reading skills. Oral language was found to be an important foundational skill for early reading success in EL1 students (Catts et al., 1999; Kendeou, Van den Broek, White, & Lynch, 2009; Roth et al., 2002; Storch & Whitehurst, 2002), and a similar relationship was demonstrated with ELLs
(Crosson & Lesaux, 2010; Gottardo & Mueller, 2009; Kieffer, 2012; Miller et al., 2006; Nakamoto, Lindsey, & Manis, 2007; Proctor et al., 2005). However, within ELL populations, the language of reading instruction (i.e., native language, English) can potentially affect the connection between oral language and reading outcomes (Gottardo & Mueller, 2009; Miller et al., 2006).

It appears that when literacy instruction is provided in the native language, oral language skills in both the child’s native language and in English are uniquely predictive of their English reading abilities, as was demonstrated by Miller et al. (2006). In this cross-sectional study, 1,531 Spanish-speaking ELLs in kindergarten through third grade received primary instruction in Spanish and gradually transitioned to English after third grade. Participants completed measures of oral language (narrative language sample), reading comprehension, and word reading efficiency in both Spanish and English. Results suggested that oral language abilities in both Spanish and English predicted reading comprehension and word reading efficiency in both languages (Miller et al., 2006). Alternatively, in a study completed by Gottardo and Mueller (2009), literacy instruction was provided in English to 131 Spanish-speaking ELLs in first and second grade. In this longitudinal study, only English measures of oral language were significantly related to English reading comprehension. Furthermore, in the Gottardo and Mueller study, intercorrelations between English and Spanish oral language measures were nonsignificant ($r = -0.04$ to $0.22$).

To further investigate how oral language abilities may be linked to reading outcomes for ELLs, Kieffer (2012) conducted a longitudinal study with a nationally representative sample of 295 Spanish-speaking ELLs. This study specifically examined the effect of early oral language skills in Spanish and English on later English reading achievement and reading growth rates.
Kieffer’s study sample differed from Miller et al. (2006) and Gottardo and Mueller (2009) in that a nationally representative sample was collected through access to a large cohort study database. Therefore, although the sample is characteristic of Spanish-speaking ELLs in the U.S., little is known about the individual participants, including their language of reading instruction.

The oral language study measures in the Kieffer (2012) study were collected in kindergarten and included a Spanish and English version of three subtests taken from the Pre-Language Assessment Scales 2000 Form C (PreLAS; Duncan & DeAvila, 1998): a listening comprehension task, an expressive vocabulary task, and a story retell task. The reading outcome measure collected in third, fifth, and eighth grade was an adaptive test constructed by a panel of experts that included items from the National Assessment of Educational Progress, measures from two longitudinal studies conducted by the National Center for Educational Statistics, and other published standardized tests (Najarian, Pollack, & Sorongon, 2009). This reading achievement assessment measured reading comprehension and vocabulary, as well as basic reading skills in the earlier grades (Kieffer, 2012).

Results from Kieffer (2012) indicated that both Spanish and English oral language skills predicted reading achievement in third, fifth, and eighth grades. However, Spanish oral language was not significantly predictive after controlling for English oral language. The author noted that these findings may differ from results of prior research (e.g., Miller et al., 2006) due to fundamental differences between the study samples, such as language of instruction. However, due to the selection of participants from a large national database, little information on language of instruction was available. Although Kieffer found that early oral language skills were predictive of later reading abilities, they did not significantly predict growth rates for reading between third and fifth grade. The author, therefore, suggested that although oral language skills
are important for learning to read in the early grades, oral language skills do not support quicker reading skill growth rates in later elementary and middle school years (Kieffer, 2012).

As is evident from Kieffer (2012), a limitation of research on oral language and reading is that the majority of studies (e.g., Miller et al., 2006; Proctor et al., 2005) employ complex batteries of oral language measures that would not be appropriate for school-based use. Schools require quick, standardized measures to better understand students’ oral language abilities, but few measures exist (Arañas & Christ, 2014a; Gardner, Froud, McClelland, & van der Lely, 2006). One available type of measure for oral language that is quick and easy to administer is sentence repetition tasks, which focus on syntactical understanding.

**Sentence repetition tasks.** Sentence repetition tasks were developed for the purpose of assessing oral language with EL1s (Arañas & Christ, 2014a; Clay, 1971; Gardner et al., 2006) and ELLs (Dockrell, Stuart, & King, 2010; Manis, Lindsey, & Bailey, 2004; Nakamoto et al., 2007; Stuart, 1999; Verhoeven, 1994). Verhoeven (1994) conducted a three-year longitudinal study that tracked Turkish-speaking, first-grade students (n = 98) learning Dutch as their second language while attending school in the Netherlands. When administered sentence repetition tasks in both Turkish and Dutch, the participants consistently performed better in their first language (i.e., Turkish) across the three years. Verhoeven scored the sentence repetition task for correctness of function words, word-final markers, and clause structure and summed those scores for an overall syntactical knowledge value. Although the participants demonstrated gains in syntactical knowledge in both languages across the three years, this gain was relatively larger in their second language (i.e., Dutch). Additionally, results indicated that in first grade, performances on the sentence repetition task in the primary and secondary languages were significantly correlated; however, this relationship was not significant in second and third grade.
Therefore, the author concluded that developments in syntactic knowledge in the first and second language are essentially independent processes (Verhoeven, 1994).

Whereas Verhoeven (1994) investigated the performance of one ELL group on a sentence repetition task administered in their first and second language, Komeili and Marshall (2013) included EL1s in addition to ELLs, but only administered the sentence repetition task in English. The purpose of this study was to investigate the repetition accuracy and error pattern differences between groups. A small sample of 18 EL1s and 18 Farsi-English bilingual children that ranged in age from six to twelve years old participated in the study. Komeili and Marshall found that EL1s repeated English sentences significantly more accurately than the bilingual children; however, once receptive vocabulary skills were controlled for, this difference was no longer significant. These results suggest that sentence repetition is generally easier when the sentences are fully understood due to higher proficiency with the language. However, a limitation of this study is that no measure of ELP was administered and therefore, researchers were relying on the crude measure of monolingual/bilingual designation to determine language proficiency. In addition to administering a more comprehensive measure of ELP, administering the sentence repetition task in the participants’ first and second languages, as Verhoeven (1994) did, would have added to the understanding of how language affected task performance.

One version of a sentence repetition task that includes two different language forms is the Oral Language Sentence Repetition (OLSR) task (Arañas & Christ, 2014b), which is a recently developed sentence repetition measure based on recommendations from Clay (1971). Both the English and Spanish versions were designed for kindergarten and first-grade children and are comprised of orally presented sentences that increase in syntactical difficulty as the task progresses. The child is asked to repeat the sentences word for word and receives a score
ranging from zero to two for each sentence based on the number of errors committed. The OLSR task is intended to measure children’s receptive oral language, and more specifically their syntactical knowledge or understanding of grammar and the structure of language (Arañas & Christ, 2014b). As is evident in the description of the behaviors required during the OLSR task, there appears to be a memory component involved. In fact, some earlier studies have classified other variations of sentence repetition tasks within test batteries of short-term or phonological working memory (Chiappe, Siegel, & Wade-Woolley, 2002; Nakamoto et al., 2007; Plaza, Cohen, & Chevrie-Muller, 2001; Shankweiler et al., 1995). Furthermore, one review emphasized that sentence repetition task performance requires both well-developed memory and language processing systems (Bley-Vroman & Chaudron, 1994).

There are opposing viewpoints within the literature as to whether sentence repetition tasks tap working memory systems or are more representative of a broad oral language ability (Klem et al., 2015). The first position is that sentence repetition tasks measure working memory capacity, which is uniquely linked to the development of language skills in children (Alloway & Gathercole, 2005; Alloway, Gathercole, Willis, & Adams, 2004). Alternatively, another faction argues that sentence repetition tasks are best described as measuring a unitary language construct (Klem et al., 2015; Moll, Hulme, Nag, & Snowling, 2015). Although there are differences between the two camps, both agree that sentence repetition tasks have a strong relationship with reading ability (Alloway & Gathercole, 2005; Moll et al., 2015).

Recently, Klem et al. (2015) conducted a study examining the longitudinal relationship between sentence repetition performance and other measures of language ability. Beginning at age four, 216 EL1s were measured at three time points over two years, utilizing tests of sentence repetition, vocabulary, and grammatical knowledge. No measures of working memory were
administered in this study, but Klem et al. argued that if sentence repetition was measuring working memory, which influences the rate of growth of language skills, then sentence repetition would be a longitudinal predictor of growth in language skills. If instead, sentence repetition were a measure of broad language ability, it would load strongly on a unitary language factor. Results suggested that there was support for sentence repetition as an underlying unitary construct of language ability, rather than as a separate memory construct (Klem et al., 2015).

To date, research investigating the role of oral language and working memory in sentence repetition task performance has largely occurred with EL1 children (e.g., Alloway & Gathercole, 2005; Klem et al., 2015; Moll et al., 2015). Therefore, there is little understanding of how these factors may come into play with ELL populations. Even more information could be garnered by administering a sentence repetition task in the ELLs’ first and second languages in conjunction with a working memory task to investigate if there was a differential relationship between working memory and sentence repetition task performance in the different languages.

**Oral language summary.** Oral language is an essential component of language proficiency that involves an understanding of the structure and grammatical rules of language. Schools require quick and easy measures of oral language, seeing as it is an important indicator of early reading success across EL1 (e.g., Roth et al., 2002) and ELL populations (e.g., Kieffer, 2012). Thus, sentence repetition tasks were proposed as an appropriate screening tool for oral language skills because they were simple to administer (Arañas & Christ, 2014b). However, the question remains whether sentence repetition tasks more strongly measure oral language or working memory (Klem et al., 2015) and how this relationship may differ for administrations in ELLs’ native language and English. Furthermore, the connection between the sentence
repetition task and reading comprehension remains unexplored within the elementary-aged ELL population.

**Working Memory**

In order to distinguish the effects of working memory from oral language, it is important to consider working memory as a separate and unique predictor of reading performance. Working memory is considered to be a cognitive process involved in information storage and manipulation (Baddeley, 1992). This simultaneous storage and processing of information was found to be an important factor in second language vocabulary learning (Baddeley, Papagno, & Vallar, 1988; Gorman, 2012). Thus, working memory is one important skill that can lay the foundation for second language development.

In addition to contributing to language acquisition, working memory also plays a critical role in reading outcomes. In fact, working memory is reportedly a unique predictor of reading outcomes in EL1s (Cain, 2006; Cain, Oakhill, & Bryant, 2004; Carretti, Cornoldi, De Beni, & Romanò, 2005) and ELLs (Swanson, Orosco, & Lussier, 2015; Swanson, Orosco, Lussier, Gerber, & Guzman-Orth, 2011). Although findings indicate that ELLs fall behind EL1s in reading comprehension, they typically perform equally to EL1 peers on measures of underlying cognitive processes such as working memory (Geva & Farnia, 2012).

Geva and Farnia (2012) found that there was no significant difference in performance of EL1s and ELLs who completed a working memory assessment (Backward Digit Span subtest of the Wechsler Intelligence Scale for Children-Third Edition [WISC-III; Wechsler, 1991]) in second and fifth grade. EL1s \(n = 149\) and ELLs \(n = 390\) were included in this study and all measures were administered in English. The ELLs received instruction in English in their Canadian schools, but spoke a language other than English at home (e.g., Punjabi, Cantonese,
Portuguese). In addition to performing equally well on the English working memory task, both ELLs and EL1s in this study demonstrated working memory performance that was significantly correlated \((p < .01)\) across second grade and fifth grade administrations, indicating that for both groups working memory was a relatively stable construct. For both groups, fifth-grade working memory scores were also significantly correlated with fifth grade reading comprehension performance (EL1s \(r = .24\); ELLs \(r = .32\)) at the \(p < .01\) level. Unfortunately, statistical analyses were not conducted to investigate if the correlations differed significantly by group.

When investigating working memory abilities with ELLs, it is important to understand if there are differences between these skills in students’ first and second languages. In a study by Abu-Rabia and Siegel (2002), fourth to eighth graders \((N = 56)\) who spoke Arabic as their native language and received English instruction in their Canadian schools completed working memory tasks in both languages. The working memory task involved supplying a missing final word to orally presented sentences, then repeating all missing words from sets of various sizes. Results indicated that working memory skills in the first and second languages were significantly correlated at the \(p < .001\) level (Abu-Rabia & Siegel, 2002). Unfortunately, this study did not include measures of reading skills; therefore, the extent to which these working memory measures might explain variance in reading performance for ELLs is unknown.

In order to investigate how working memory and reading abilities may be related for Spanish-speaking ELLs, Swanson et al. (2015) completed a three-year longitudinal study. In this study, working memory and reading achievement measures were administered in both Spanish and English. A total of 410 ELLs in grades one \((n = 133)\), two \((n = 132)\), and three \((n = 145)\) were included in the study. The working memory measures included a short-term memory component (i.e., phonological loop) and an executive memory component (i.e., controlled
attention, processing). The short-term memory measure was gathered from scores on Forward and Backward Digit Span tasks from the WISC-III (Wechsler, 1991) as well as experimenter-created word span and pseudoword span tasks that involved recalling sets of words/pseudowords that gradually increased in set size. The executive memory component of working memory was gathered from scores on experimenter-created measures of conceptual span (organizing sequences of words into categories following a distractor), listening sentence span (recalling final words of lists of sentences following a distractor), rhyming span (recalling sequences of rhyming words following a distractor), and updating (recalling the last three digits of lists of three, five, seven, or nine digits) tasks. These working memory measures were compared to criterion measures of word identification and passage comprehension from the Woodcock–Muñoz Language Survey–Revised (Woodcock, Muñoz-Sandoval, & Alverado, 2005), receptive vocabulary as measured by the Peabody Picture Vocabulary Test (Dunn & Dunn, 1981), and oral language scores from the Illinois Test of Psycholinguistic Abilities III (Hammill, Mather, & Roberts, 2001) and the Expressive One-Word Picture Vocabulary Test (Brownell, 2001).

Results from Swanson et al. (2015) indicated that performance on both the short-term and executive components of working memory (measured in English) at time one was significantly ($p < .001$) related to all English reading and English language acquisition measures three years later. Additionally, growth in English reading and language acquisition was related to a diminished relationship between the outcome measures and short-term and executive working memory, suggesting that working memory is most essential in the initial reading skill acquisition phases. Finally, ELL participants were within the average range on norm-referenced English measures of reading and vocabulary across the three testing periods. However, the sample was less proficient in Spanish achievement measures across the study time period, as the results
showed below average scores on Spanish reading and vocabulary measures. The authors reason that these results were due to the fact that the sample received school instruction primarily in English (Swanson et al., 2015). A limitation of this study was that ELP was not measured or considered as a potential variable that could affect the relationship between working memory and English reading outcomes.

**Working memory summary.** Working memory, which involves temporary storage and manipulation of information, is an essential foundational ability that has a close relationship to second language learning (Baddeley, Papagno, & Vallar, 1988; Gorman, 2012) and reading performance for ELLs (Swanson et al., 2015; Swanson et al., 2011). Specifically, working memory is thought to be most essential for developing early reading skills, but it also has a lasting effect on reading performance in later elementary grades (Swanson et al., 2015). When considering working memory performance of ELLs, it is important to note that ELLs typically perform the same as EL1s on measures of working memory (Geva & Farnia, 2012), and there is a strong relationship between working memory skills in their first and second languages (Abu-Rabia & Siegel, 2002). The research on the connection between working memory and reading comprehension for ELLs is limited in that ELP has not been considered as a factor that could affect this relationship.

**Decoding**

Decoding is another important component process of reading that is closely associated with reading comprehension, perhaps even more so than working memory within EL1 populations (Goff, Pratt, & Ong, 2005). The National Reading Panel identified decoding as an essential component of reading comprehension and therefore emphasized the importance of phonemic awareness instruction for promoting successful reading (National Institute of Child
Health and Human Development, 2000). The SVR also highlights decoding, along with linguistic comprehension, as a critical factor in predicting reading comprehension skills (Gough & Tunmer, 1986; Hoover & Gough, 1990). Typically, within SVR models, decoding independently explains a unique portion of variance in reading comprehension, but the product of decoding and linguistic comprehension explains a larger portion of the variance (e.g., Hoover & Gough, 1990; Joshi & Aaron, 2000). For example, Joshi and Aaron (2000) reported that decoding explained 35% of the variance in reading comprehension scores, and the product of decoding and linguistic comprehension explained 48% of the variance in their sample of third-grade EL1 participants.

Decoding is also noted as an important predictor of English reading success for ELLs (Gottardo & Mueller, 2009; Hoover & Gough, 1990; Nakamoto et al., 2007; Proctor et al., 2005). Although research indicates that decoding is an important skill to consider, some suggest that for ELLs in later elementary grades (e.g., fourth grade), decoding skills may be less predictive of reading comprehension than other component skills such as oral language (Proctor et al., 2005). But in the earlier elementary grades, decoding skills are highly correlated with skilled reading (Hoover & Gough, 1990; Nakamoto et al., 2007).

To investigate the relationship between English word decoding and English reading comprehension for ELLs, Nakamoto et al. (2007) tracked 303 Spanish-speaking ELLs from first through sixth grades. These students were learning through a transitional bilingual curriculum in which they began kindergarten learning in both Spanish and English, then were transitioned to primarily English instruction in first grade once they had gained competency with the English language. It is important to note, however, that the extent to which these results generalize to ELLs without access to this type of program may be limited.
To measure decoding, Nakamoto et al. (2007) utilized the Word Attack and Letter-Word Identification subtests of the Woodcock-Johnson Tests of Achievement (Woodcock & Johnson, 1989; Woodcock, McGrew, & Mather, 2001), which requires students to name increasingly difficult real and nonsense words. Reading comprehension was measured with the Passage Comprehension subtest of the Woodcock-Johnson Tests of Achievement (Woodcock & Johnson, 1989; Woodcock, et al., 2001). Results from Nakamoto et al. suggested that, for Spanish-speaking ELLs, decoding and reading comprehension were more highly correlated in first grade than in sixth grade; however, a significant ($p < .001$) correlation existed at both time points.

**Decoding summary.** Decoding is an important component process of reading for ELLs, which is closely linked to successful reading comprehension performance, especially in the early elementary grades (Nakamoto et al., 2007). Within the SVR, decoding is recognized as a necessary skill that, in conjunction with linguistic comprehension, predicts reading comprehension (Gough & Tunmer, 1986; Hoover & Gough, 1990). Although the relationship between decoding and reading comprehension has been explored with ELLs through the SVR model (e.g., Proctor et al., 2005), this literature has not included ELP as a variable in analyses. Therefore, it is unclear to what extent differences in reading comprehension may be explained by decoding and linguistic comprehension (i.e., traditional SVR model) versus ELP.

**Summary and Purpose**

Research with ELLs in the elementary grades suggests that a number of component factors (e.g., decoding, oral language, working memory) contribute to overall reading outcomes. Many of these early indicators of reading comprehension have been included in comprehensive models, such as the SVR, and findings suggest that reading comprehension skill can be explained as the product of decoding and linguistic comprehension (Gottardo & Mueller, 2009; Hoover &
Gough, 1990; Proctor et al., 2005). However, it is unclear to what extent ELP may affect how component skills impact higher-order reading success, as the majority of studies (e.g., Gottardo & Mueller, 2009; Kieffer, 2012; Nakamoto et al., 2007; Proctor et al., 2005; Swanson et al., 2015) have not segregated ELL participants by ELP level. Considering that initial differences have been detected in reading achievement performance between Spanish-speaking students who do and do not require ESL services (Kieffer, 2008), it is important to extend this research to investigate within-ELL group differences based on variations in ELP. With more detailed information regarding how ELP may affect reading comprehension outcomes, educators can more accurately make instructional decisions for different subgroups of ELLs. Therefore the first purpose of the present study is to systematically recruit ELLs with varying levels of ELP and administer measures of oral language (in Spanish and English), working memory, decoding, and reading comprehension. This will allow the researcher to investigate how these component factors independently and cumulatively affect reading comprehension and the role ELP may play in this relationship. Additionally, by exploring whether working memory and ELP add value to the predictive model of reading comprehension for ELLs, the ELL SVR literature will be extended.

Another purpose of the present study is to more closely investigate the sentence repetition task, a commonly used measure of syntactical oral language ability. Although it is clear from extant research (e.g., Arañas & Christ, 2014b, Moll et al., 2015, Verhoeven, 1994) that sentence repetition tasks measure oral language, there is disagreement on whether sentence repetition tasks measure a unitary language construct or if they are primarily a measure of working memory, which is then uniquely linked to oral language skills (Klem et al., 2015). In order to extend this literature base, the present study will investigate the roles that working memory and
oral language play in explaining variance in English and Spanish administrations of a sentence repetition task. Furthermore, this study aims to examine whether the English or Spanish version of the sentence repetition task is a better representation of oral language skills for ELLs receiving English instruction.

**Research Questions and Hypotheses**

The present study examines the following research questions:

1. *Which administration of the sentence repetition task (Spanish or English) is more closely related to other measures administered in this study? And more specifically, which version of the sentence repetition task is a better indicator of oral language skills, as measured by the Oral Language Composite of the ACCESS test?*

   Given past research findings suggesting that the match between language of instruction and language of assessment is important (Abedi et al., 2004; Hofstetter, 2003; Kieffer et al., 2009), as is the match between assessment languages (Gottardo & Mueller; 2009), it is predicted that the English OLSR will be more highly correlated with other English measures administered in this study. Furthermore, it is hypothesized that the English administration of the sentence repetition task will be a better indicator of English oral language skills. This research question is important to address because whichever predictor variable (i.e., English OLSR, Spanish OLSR) correlates more strongly with the criterion variable (i.e., ACCESS Oral Language Composite), will be used in the final stage of analyses.

2. *Are oral language skills, as measured by the Oral Language Composite of the ACCESS test, and working memory significant unique contributors to English and Spanish administrations of a sentence repetition task (i.e., the OLSR)? Is oral language or working memory a stronger contributor in each language? Is working memory more highly correlated with the English or
Spanish administration of the sentence repetition task?

The Oral Language Composite of the ACCESS test is comprised of a combination of the listening and speaking subscales (Abedi, 2007; Center for Applied Linguistics, 2015b), which are skills that are required for successful performance on sentence repetition tasks (Abedi, 2007; Roth, Speece, & Cooper, 2002). It is expected that the Oral Language Composite of the ACCESS test will be significantly correlated with performance on English and Spanish administrations of the OLSR; however, this relationship will be stronger for the English OLSR, given the match between languages. Furthermore, it is hypothesized that the sentence repetition task is not solely measuring oral language skills, but that working memory also plays a role in performance on this measure (Alloway & Gathercole, 2005; Alloway, Gathercole, Willis, & Adams, 2004), albeit a smaller role. It is expected that memory systems will be taxed most prominently when proficiency with the language is poor. Thus, it is anticipated that ELLs will require more working memory input to perform well on the English versus the Spanish version of the OLSR.

3. Do working memory and ELP explain variance in reading comprehension above and beyond variance explained by oral language and decoding, the variables typically used in SVR models?

It is expected that measures of oral language, decoding, working memory, and ELP will each individually predict English reading comprehension. This hypothesis is based on research with ELLs in the elementary grades that investigated the impact of oral language (e.g., Gottardo & Mueller, 2009), decoding (e.g., Nakamoto et al., 2007), working memory (e.g., Swanson et al., 2015), and ELP (e.g., Kieffer, 2008) on higher-order reading outcomes. Additionally, the SVR model is hypothesized to explain a large amount of variance in reading comprehension. However, it is predicted that working memory and ELP will explain variance in reading
comprehension above and beyond that which is explained by the SVR. It is expected that a large portion of variance in reading comprehension will be explained by the four proposed predictive variables (i.e., oral language, decoding, working memory, ELP).
References


oral language assessments (Master’s thesis). University of Minnesota, Minneapolis, Minnesota.


Center for Applied Linguistics (2015b). The WIDA consortium English language proficiency assessment for grades 1-12: Test and item design plan for the annual summative and on-demand screener.


WIDA Consortium (2012). 2102 Amplification of the English Language Development


CHAPTER 3

METHOD

Participants and Setting

Data were collected with first- and second-grade ELLs recruited from four elementary schools, and 25 different classroom teachers, in a Georgia school district. District-wide, 10% of students were ELLs, ethnic make-up of students was 61% White, 16% Hispanic, 13% Black, 5% Asian, and 5% Multiracial, and 58% of students qualified for free or reduced-price lunches. A school staff member who was authorized to access student records identified ELLs who were eligible to participate in the study. Eligibility was determined through the ELL designation in the students' school records. Thus, all participants were receiving English to Speakers of Other Languages (ESOL) services during the school year in which data were collected, and were required to take the Assessing Comprehension and Communication in English State to State for English Language Learners (ACCESS; WIDA, 2016) test, which is a measure of ELP. Students who had not attended school within the district since the beginning of the school year were excluded from participation.

Participants (N = 84) were primarily male (63.1%), largely attended preschool (82.1%), ranged in age from 6.70 years to 8.96 years (M = 7.68 years, SD = 0.60), and were enrolled in either first grade (n = 51) or second grade (n = 33). Of note, although a majority of participants had prekindergarten experiences, Georgia does not provide ESOL services at the prekindergarten level. Six participants (7.1%) qualified for Special Education services. Participants’ ACCESS scores ranged from level 2.1 to 6.0 (M = 4.4, SD = 0.96), indicating that all students had at least
the basic understanding of English required to understand task instructions provided in the present study and some students had reached a level of English proficiency to qualify them to exit ESOL services for the next school year. Sample size was determined by using an a priori power analysis for multiple regression. This sample size allowed for detection of medium effects ($f^2 = .15, \alpha = .05, \beta = .80, 4k$) in multiple regression analyses, according to Cohen’s (1988; 1992) effect size tables.

The native language for all ELLs participating in this study was Spanish, and all identified as Hispanic. A brief survey regarding the frequency of Spanish and English spoken at home was included on the Spanish consent form for the parents to complete. Of the students participating in the study, 46.4% spoke only Spanish at home, 39.3% spoke Spanish and a little English, 10.7% spoke Spanish and English equally, and 3.6% spoke English and a little Spanish. Within the classroom context, all participants received English instruction.

Consent forms were sent home in English and Spanish to the parents of eligible ELLs. All students that returned consent forms, regardless of their parents’ decision, were able to choose a reward from the researcher’s prize box. If consent forms are not returned, a second consent form was sent home, in case the students failed to deliver it to parents. In addition to providing consent for participation, participants’ parents also provided consent for a review of school records. The records review allowed the experimenter to collect participants’ scores from the winter 2016 administration of the ACCESS to utilize in participant selection and data analyses. Participants were systematically selected in an attempt to include students that sufficiently represented each ELP level. The records review also allowed the experimenter to gather demographic information (e.g., gender, age, race, time in school district). Only students
with written consent from a parent or legal guardian and a signed student assent form participated in the study.

**Measures**

**ACCESS.** The ACCESS is a test of ELP that is administered once per year to ELLs in WIDA Consortium states (33 states in 2013-2014). Administration of the ACCESS test meets federal mandates for annual ELP testing (Center for Applied Linguistics, 2015a). As Georgia is one of the WIDA Consortium states, the ACCESS test was administered to all study participants during the seven-week test window in winter 2016, approximately two to four months prior to the data collection for remaining study measures.

The basis for ACCESS test development is heavily rooted in research on academic language (Abedi, 2007; Fox & Fairbairn, 2011; WIDA Consortium, 2012), as well as ELP standards that specifically address language proficiency in academic settings, including Social and Instructional Language, Language of Language Arts, Language of Mathematics, Language of Science, and Language of Social Studies (Kenyon, 2006). In order to capture academic English abilities across multiple modalities, scores on the ACCESS are provided in four language domains: reading, writing, listening, and speaking, as well as composite scores for Oral Language, Literacy, Comprehension, and Overall. The composite scores are weighted as follows: the Oral Language Composite is 50% listening and 50% speaking, the Literacy Composite is 50% reading and 50% writing, the Comprehension Composite is 70% reading and 30% listening, and the Overall Composite is 15% listening, 15% speaking, 35% reading, and 35% writing (Abedi, 2007; Center for Applied Linguistics, 2015b).

Composite scores in all areas are available as both scale scores and proficiency level scores. Raw scores are converted to scale scores through a statistical equating and scaling
process, which allows for scoring consistency across test forms and grade-level clusters (Center for Applied Linguistics, 2015a). In order to create a test that was developmentally appropriate for various ages, ACCESS includes test forms for five grade level clusters (K, 1-2, 3-5, 6-8, and 9-12), which the test developers refer to as the vertical test dimension (Center for Applied Linguistics, 2015a). Vertical scales, which measure results across grades on the same scale, are ideal for tracking growth over time (Kenyon, MacGregor, Li, & Cook, 2011). ACCESS also includes a horizontal dimension, which allows for varying difficulty of test items based on the varying levels of ELP within each grade level cluster; Tier A (beginning ELLs), Tier B (intermediate), and Tier C (advanced; Center for Applied Linguistics, 2015a; Kenyon, 2006). Due to variations in tiers and grade level, students are exposed to different test items; however, as a result of equating and scaling, scale scores mean the same regardless of which test items students take (Center for Applied Linguistics, 2015a). Thus, scale scores are used for all analyses within the present study.

Proficiency level scores interpret students’ scale scores based on application of English language development standards. Unlike scale scores, proficiency level scores are dependent upon the student’s grade (Center for Applied Linguistics, 2015a). Therefore, within the present study they were solely used as a descriptor of the study sample rather than for statistical analyses. Proficiency level scores are commonly used within school settings to make interpretations regarding students’ ELP. ACCESS proficiency level scores fall into one of five levels: Entering (level 1), Beginning (level 2), Developing (level 3), Expanding (level 4), Bridging (level 5), as well as one final exit stage known as Reaching (level 6; Kenyon, 2006). Specific procedures for exiting students from language assistance services vary across states (Fox & Fairbairn, 2011); however, proficiency level scores on the Overall Composite of the
ACCESS are often a primary factor in these decisions (Abedi, 2007).

From its inception, ACCESS has undergone rigorous testing to ensure that the scores are reliable and valid as measures of ELP for all grades and ability levels (Center for Applied Linguistics, 2015a). To empirically assess content validity, data from the initial field test of 6,500 students were analyzed and it was determined that the items were appropriately ordered to represent the five proficiency levels defined by the standards (Abedi, 2007; Kenyon, 2006). Concurrent validity was assessed utilizing data from a bridge study involving 4,985 students in grades K-12. Results demonstrated moderate to high correlations between four older ELP tests and ACCESS, indicating appropriate concurrent validity (Abedi, 2007; Gottlieb & Kenyon, 2006; Kenyon, 2006).

The annual technical report for the 2013-2014 administration of ACCESS, prepared by the Center for Applied Linguistics (2015a), indicated that reliability for the Overall Composite score was very high across the grade level clusters. Specifically, the reliability for grades 1-2 was .943, which is important to note given the present study’s first- and second-grade sample. Similarly, the 2013-2014 administration of ACCESS indicated a very high accuracy level for decisions about student placement using the Overall Composite score. High accuracy was demonstrated across all grade and proficiency level cut scores; however, because first and second grade was the population of interest for the present study, accuracy scores for those grades are reported here. In first grade, the level 1/level 2 cut was .985; 2/3 cut, .938; 3/4 cut, .930; 4/5 cut, .970; and 5/6 cut, .991. In second grade, the level 1/level 2 cut was .985; 2/3 cut, .956; 3/4 cut, .911; 4/5 cut, .936; and 5/6 cut, .985 (Center for applied Linguistics, 2015a).

Beginning in the 2015-2016 school year, ACCESS version 2.0 was administered in the schools. Test creators from the Center for Applied Linguistics (2015b) explained that although
the measured constructs and performance criteria remained the same, a number of improvements were made from the first to second version of the test. First, ACCESS 2.0 is now computer-administered, with the exception of the Writing subtest for students in first through third grades, which remains in paper and pencil format. Second, all subtests are now group administered, including the Speaking subtest, which previously required individual administration. Third, grade clusters are modified slightly to provide separate forms for grades 1, 2-3, 4-5, 6-8, and 9-12; however, comparisons across grades remain appropriate. Fourth, test design and functionality are streamlined through features such as a next arrow that prevents unintentional skipping of questions, a progress bar indicating test-taker progress through the subtest, and flexible design format allowing for variation in text size, volume, and speed of video depending on student needs. Testing, spread across three sessions, takes approximately 2 hours and 15 min, which is comparable to the first version of the test (Center for Applied Linguistics, 2015b).

Working memory assessment. The Digit Span subtest from the Wechsler Intelligence Scale for Children, Fifth Edition (Wechsler, 2014) was administered to all participants as a measure of working memory. This subtest is appropriate for use with children ages six through 16. The Digit Span subtest includes three subparts: Digit Span Forward, Digit Span Backward, and Digit Span Sequencing. Participants’ working memory score will consist of the summed raw scores from each of the three subparts. Given the relatively low cultural and linguistic demand of digit span tasks (Flanagan, Ortiz, & Alfonso, 2013) and strong correlations between working memory scores in ELLs’ first and second languages (Genesee & Geva, 2006), working memory tasks in the present study were administered in English. Reliability coefficients ranged from 0.90 to 0.92 for the ages that were tested (Wechsler, 2014).

Digit Span Forward. During Digit Span Forward, an examiner read digits aloud while
the participant listened and then repeated back the digits in the same order. The digits were read aloud at a rate of one digit per second. The items of this subtest became increasingly more difficult, beginning with two digits that must be repeated and increasing to nine digits on the last item. This subtest included eight items with two trials per item. A participant received a score of “1” for each trial with digits repeated in the correct order. A score of “0” was assigned when a participant repeated the numbers incorrectly or did not respond. The examiner discontinued testing when scores of “0” were earned on both trials within an item.

**Digit Span Backward.** Procedures for Digit Span Backward were the same as in Digit Span Forward, except that participants were required to repeat the digits in reverse order. To familiarize the participants with the procedures, two practice trials were administered in which participants practiced repeating orally presented digits in reverse order. Corrective feedback was provided as necessary on the practice trials. Discontinue criteria were the same as described in Digit Span Forward.

**Digit Span Sequencing.** Digit Span Sequencing was the final portion of the working memory assessment. Procedures regarding item administration, practice trials, and discontinue criteria were the same as in Digit Span Backward. However, in order to earn a score of “1” on this subpart, participants were required to repeat the digits back in numerical order. On some trials the same digit was listed more than once and the participant was required to list that digit the correct number of times. One of the practice items included a trial with a digit that was listed more than one time and thus, the participants gained practice with this type of item administration before the experimental items began.

**Oral language assessment.** Participants completed the individually administered Oral Language Sentence Repetition task (OLSR; Arañas & Christ, 2014). The OLSR is part of the
FastBridge Learning suite of assessment tools, which is utilized in schools across 35 states, including a statewide adoption in Iowa (www.fastbridge.org). Specifically, the OLSR is one subtest of the earlyReading assessment in the FastBridge Learning system and is appropriate for screening with students up to third grade. All participants were administered the English and Spanish versions of the OLSR.

In an effort to standardize task instructions and item administration across participants, a pre-recorded video was played for each participant that included the directions, practice items, and test items read aloud by a young-adult male. A separate video was recorded for the English and Spanish versions of the OLSR. The English video featured an individual whose native language was English, and the Spanish video featured an individual whose native language was Spanish.

**English OLSR.** During the English OLSR administration, the video including task instructions, and item administration was played for the participant. Two practice sentences and 20 test sentences in English were administered to each participant. The participant was required to repeat back each sentence exactly as it was said in the video. Sentences became syntactically more complex as the test progressed. The participant earned a full score of “2” if the sentence was repeated word for word without any errors, a score of “1” if the participant made only one error, and a score of “0” if the participant made more than one error or did not respond within five seconds. Types of errors included: omissions, commissions, and switching words. Non-errors included: contractions, speech or dialectical differences, repeated words, and self-corrections. The examiner provided corrective feedback on the two practice items as needed. The participant’s score was the sum of scores on all test sentences. The English OLSR has strong psychometric properties; high reliability and validity were reported for a test sample of
first grade students (Arañas & Christ, 2014). For example, internal consistency was .99, two-week test-retest reliability was .86, and inter-rater reliability was .91 (Arañas & Christ, 2014).

**Spanish OLSR.** Procedures for the Spanish OLSR were the same as for the English OLSR, except the video directions and item administration were conducted in Spanish. Two practice sentences and 15 test sentences in Spanish were administered to each participant. Only examiners who were proficient Spanish speakers and passed the two practice trials and two certification trials available on the FastBridge Learning website administered and scored the Spanish OLSR in the present study. Reliability and validity info for the Spanish OLSR is not yet available.

**Decoding assessment.** The Nonsense Words subtest (Christ et al., 2014) of the earlyReading assessment in the FastBridge Learning system was individually administered to all participants. This test is designed to measure if participants can recognize English letter-sound correspondences and read them fluently. The Nonsense Words subtest is appropriate for use with students in K-third grade. The participant was asked to read pretend words that were listed on a page while an examiner tracked the time and errors committed. The participant was encouraged to read the nonsense words as whole words but was also permitted to read the sounds of each letter; both formats were considered correct. Incorrect reading, omissions, and 3 s hesitations were considered to be errors. If the participant hesitated for 3 s, the examiner provided the word and asked the participant to continue to the next word. The test began with a practice item and corrective feedback was provided if necessary. The test form included 50 nonsense words and the participant’s score was the number of words read correctly in 1 min. An adjusted score was calculated if a student finished in less than one min.

The Nonsense Words subtest of the FastBridge Learning system has strong psychometric
properties. Specifically, within a first-grade sample, alternate form reliability coefficients ranged from .69-.96, two-week test-retest reliability was .76, and inter-rater reliability was .99 (Christ et al., 2014). The concurrent criterion-related validity between the Nonsense Words subtest and the Iowa Test of Basic Skills (Hoover, Dunbar, & Frisbie, 2001), a group-administered reading achievement test, is .77 in first grade and .64 in second grade (January, Ardoin, Christ, Eckert, & White, 2016). Given that the participants received their literacy instruction in English and that researchers (Abedi, Hofstetter, & Lord, 2004; Townsend & Collins, 2008) recommend matching the language of testing to the language of instruction, the decoding assessment was administered in English.

**Passage Comprehension.** The Passage Comprehension subtest from the Woodcock-Johnson IV Tests of Achievement (Schrank, Mather, & McGrew, 2014) was administered to all participants as a modified cloze task measure of English reading comprehension. This subtest is appropriate for use with prekindergarten students through adults. The items begin with basic rebuses and gradually become more difficult, moving to pictures that tell about words, and finally silent reading of sentences and short passages, which require an oral response for a missing word. Administration began with item one for each student, and participants received a score of “1” for each correct response and a score of “0” for each incorrect response. The examiners tested by complete pages and discontinued testing when three consecutive scores of “0” were recorded. Raw scores were used for all analyses. Reliability coefficients ranged from 0.93 to 0.98 for the ages that were tested (Schrank, Mather, & McGrew, 2014).

**Procedure**

**English OLSR and Spanish OLSR.** Participants completed the English and Spanish OLSR with the assistance of an undergraduate or graduate student trained in testing procedures.
All subtests were individually administered. The order of the English OLSR and Spanish OLSR subtests was counterbalanced to control for ordering effects. The assessments were completed in a quiet location in the participating school. Administration time was approximately 10-15 minutes.

**Working memory, decoding, and reading comprehension assessment.** Participants worked with a trained graduate student to complete the working memory, decoding, and reading comprehension assessments. Administration order of these three measures was counterbalanced to control for ordering effects. The assessments were administered individually in a quiet location at participants’ school and took approximately 25-35 min. Once participants completed all study assessments, they received a small prize.

**Procedural Fidelity and Inter-Observer Agreement**

All assessments were audio recorded to allow for procedural fidelity and inter-observer agreement scoring by a second, trained individual. Procedural fidelity and inter-observer agreement data were collected for 33% of English OLSR, Spanish OLSR, and Nonsense Words assessments, as well as 10% of Digit Span and Passage Comprehension assessments.

Procedural fidelity was calculated using a checklist to ensure that all steps were completed correctly. Data are reported as total steps completed correctly divided by total possible number of steps multiplied by 100 to yield a percentage. Across examiners and study measures, procedural specifications were adhered to closely, with procedural fidelity equaling 98% or higher for all assessments (range 98-100).

Inter-observer agreement data were calculated to determine the extent to which two independent raters agreed on accuracy of participant responses. Agreement was defined as two independent raters giving the same score for a participant’s answer. Data are reported as number
of agreements divided by number of agreements plus disagreements multiplied by 100 to yield a percentage. Across examiners and study measures, inter-observer agreement coefficients were excellent. Inter-observer agreement averaged 92% for English OLSR (range = 75% to 100%), 91% for Spanish OLSR (range = 73% to 100%), 98% for Nonsense Words (range = 88% to 100%), 100% for Digit Span, and 100% for Passage Comprehension.

Data Analyses

To address research question one, Pearson correlations were conducted to determine whether significant relationships existed between the sentence repetition task measures and other measures administered in this study. In order to determine whether the correlation coefficients were significantly different from each other, coefficients were first transformed to z-scores using Fisher’s r-to-z transformation. Then the statistical significance between two coefficients was determined by using equations described by Steiger (1980) specific to dependent correlations.

Important correlations that were highlighted were the relationships between the ACCESS Oral Language Composite and both the English OLSR and the Spanish OLSR. The predictor variable (i.e., English OLSR, Spanish OLSR) with the strongest relationship with the criterion variable (i.e., ACCESS Oral Language Composite) was used as the oral language variable in the final stage of analyses. This step of replacing the ACCESS Oral Language Composite with one of the sentence repetition task variables was necessary in order to prevent violating the assumption of independence of observations when the ACCESS Overall Composite was included in the same analysis.

The second research question regarding the effects of oral language and working memory on English and Spanish administrations of the OLSR was addressed using regression analyses. First, a multiple regression analysis was conducted with the Oral Language Composite of the
ACCESS test and working memory as the independent variables and the English OLSR as the dependent variable. A similar multiple regression analysis was then conducted with the Spanish OLSR as the dependent variable. Finally, Pearson correlations between working memory and the English and Spanish OLSR were compared to determine whether working memory was more highly correlated with the sentence repetition task in one language over the other. The same procedure described above for statistical analysis of difference between correlation coefficients from dependent samples was also utilized here to answer research question two.

Research question three, regarding the potential added value of working memory and ELP to the SVR, was answered with a hierarchical multiple regression analysis. First a multiple regression analysis was conducted with reading comprehension as the dependent variable and oral language and decoding as the independent variables. Then, working memory and ELP were added into the model to investigate the potential added value of these variables to explaining variance in reading comprehension. This subsequent step in the model was examined for relative change ($\Delta R^2$) in the proportion of variance explained in the reading comprehension outcome variable. Additionally, within the full model, the researcher examined which variables were significant unique predictors for reading comprehension performance.
References


Center for Applied Linguistics (2015b). The WIDA consortium English language proficiency assessment for grades 1-12: Test and item design plan for the annual summative and on-demand screener.


CHAPTER 4

RESULTS

Descriptive statistics for all study measures are provided in Table 1. Intercorrelations between all measures are reported in Table 2. Prior to running analyses, it was determined that the characteristics of the data set were consistent with all assumptions for multivariate analyses.

Sentence Repetition Task Findings

As indicated in Table 2, significant correlations were observed between all measures administered, with the exception of between the Spanish OLSR measure and the ACCESS Oral Language Composite. To address research question one, intercorrelations between study measures were examined. Correlations were greater between the English OLSR and all other measures administered than between the Spanish OLSR and all other outcomes. For example, examination of differences in correlations between English and Spanish OLSR with the ACCESS Oral Language Composite ($z = 5.56; p < .001$), Working Memory Total scores ($z = 2.87; p = .004$), as well as Passage Comprehension ($z = 4.46; p < .001$), all indicated significantly stronger correlations with the English OLSR measure.

In an effort to answer research question two and to better understand the constructs being measured by the OLSR measures, multiple regression analyses were conducted to evaluate the extent to which oral language, as measured by the oral language component of the ACCESS, and working memory predicted student performance on each OLSR task. Only Working Memory Total was found to be a significant unique predictor ($\beta = .303; p = .011$) of students’ Spanish OLSR performance. However, results indicated that both Working Memory Total scores ($\beta =
.388) and ACCESS Oral Language Composite scores ($\beta = .443$) were significant ($p < .001$) and
unique contributors to students’ English OLSR performance. Given that both variables were
found to significantly contribute to the English OLSR, further analysis was conducted to
determine which variable was a stronger contributor. The difference between the correlation
coefficients ($r_s = .563$ and .596) was not statistically significant ($z = 0.44; p > .05$); therefore,
there was not a meaningful difference between the contribution of working memory and oral
language to performance on English OLSR.

**Explaining Variance in Reading Comprehension**

Consistent with past research, results reported in Table 2 reveal that all measures
administered were significantly correlated with Passage Comprehension. The well-researched
SVR model indicates the importance of decoding and oral language in predicting reading
comprehension. In the current study, research question three was addressed through a
hierarchical multiple regression, which was performed to investigate the potential added value of
working memory and ELP to the SVR model. Students’ performance on the English OSLR was
used as the oral language variable in this model, given that it was a better predictor of oral
language than was the Spanish OLSR.

Step one of the regression indicated that 58.6% of the variability in Passage
Comprehension scores could be accounted for by variation in English OLSR and Nonsense
Words ($R^2 = .586; F[2,81] = 57.37; p < .001$), suggesting that the SVR was a suitable model for
predicting reading comprehension. However, there was an 11.6% increase ($\Delta R^2 = .116; p < .001$)
in predictive capacity when Working Memory Total and the ACCESS Overall Composite were
added into the model. The full model, including all four independent variables, explained 70.2%
of variance in students’ Passage Comprehension scores ($R^2 = .702; F[4,79] = 46.46; p < .001$).
Interestingly, within the full model, only English OLSR ($\beta = .266; p = .005$) and ACCESS Overall Composite ($\beta = .459; p < .001$) were found to be significant unique contributors to Passage Comprehension performance; Working Memory Total ($\beta = .149; p = .059$) and Nonsense Words ($\beta = .115; p = .139$) were not significant unique contributors.
Table 1

Descriptive Statistics

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCESS Overall Composite</td>
<td>312.99</td>
<td>24.41</td>
<td>252 - 367</td>
</tr>
<tr>
<td>ACCESS Oral Language Composite</td>
<td>346.37</td>
<td>28.63</td>
<td>271 - 404</td>
</tr>
<tr>
<td>Working Memory Total</td>
<td>17.38</td>
<td>4.40</td>
<td>0 - 27</td>
</tr>
<tr>
<td>English OLSR&lt;sup&gt;a&lt;/sup&gt;</td>
<td>23.39</td>
<td>6.50</td>
<td>6 - 39</td>
</tr>
<tr>
<td>Spanish OLSR&lt;sup&gt;b&lt;/sup&gt;</td>
<td>11.98</td>
<td>8.15</td>
<td>0 - 28</td>
</tr>
<tr>
<td>Nonsense Words</td>
<td>21.90</td>
<td>13.68</td>
<td>1 - 63</td>
</tr>
<tr>
<td>Passage Comprehension</td>
<td>22.06</td>
<td>4.89</td>
<td>9 - 31</td>
</tr>
</tbody>
</table>

<sup>a</sup>20 sentences administered, maximum possible score was 40. <sup>b</sup>15 sentences administered, maximum possible score was 30.

Table 2

Correlation Matrix

<table>
<thead>
<tr>
<th>Measure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ACCESS&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. ACCESS OL&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.85***</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. WM&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.48***</td>
<td>.40***</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. English OLSR</td>
<td>.70***</td>
<td>.60***</td>
<td>.56***</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Spanish OLSR</td>
<td>.20*</td>
<td>-.04</td>
<td>.24*</td>
<td>.35**</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. NW&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.55***</td>
<td>.35**</td>
<td>.48***</td>
<td>.49***</td>
<td>.24*</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>7. PC&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.78***</td>
<td>.61***</td>
<td>.58***</td>
<td>.73***</td>
<td>.29**</td>
<td>.57***</td>
<td>-</td>
</tr>
</tbody>
</table>

Note. ACCESS = ACCESS Overall composite; ACCESS OL = ACCESS Oral Language composite; WM = Working Memory Total score; OLSR = Oral Language Sentence Repetition; NW = Nonsense Words; PC = Passage Comprehension.

<sup>a</sup>Variables with significantly (<code>p < .05</code>) stronger correlations with English OLSR than Spanish OLSR

<sup>p < .05</sup>.  <sup>**p < .01</sup>.  <sup>***p < .001</sup>
CHAPTER 5
DISCUSSION

Extant research investigating the well-known model for reading comprehension, the SVR, suggests that the combination of decoding and oral language understanding provides good predictive value for reading comprehension skills for ELLs in the elementary grades (Gottardo & Mueller, 2009; Hoover & Gough, 1990; Proctor, Carlo, August, & Snow, 2005). However, numerous empirical studies also highlight a bevy of component factors, such as decoding and oral language, which are included in the SVR, but also working memory (e.g., Swanson, Orosco, & Lussier, 2015) and ELP (e.g., Kieffer, 2008) as important factors related to reading comprehension for young ELLs. Thus, the primary purpose of the present study was to investigate how these component factors independently and cumulatively predict reading comprehension skills, and more specifically, to extend the SVR literature by exploring whether working memory and ELP add value to this predictive model for reading comprehension in an ELL sample. Utilizing a comprehensive measure of ELP was an important feature of the present study because previous reading research with ELLs has largely overestimated the homogeneity of ELLs by not recognizing intragroup variations in ELP (Artiles, Rueda, Salazar, & Higareda, 2005).

Consistent with past research, results from the present study indicated that decoding, oral language, working memory, and ELP were all significantly related to reading comprehension performance. Furthermore, findings suggested that the SVR was an appropriate predictive model for reading comprehension in this sample, as the combination of decoding and oral language
explained a significant amount of variance in reading comprehension performance. However, in line with hypothesized outcomes, adding working memory and ELP to the SVR model did in fact allow for significantly greater variation to be explained in students’ reading comprehension, above and beyond the variance explained by decoding and oral language alone. This finding suggests that working memory and ELP are important factors to consider, in conjunction with decoding and oral language, when attempting to gain a full picture of factors related to reading comprehension for ELLs. A more in depth look at the full model including all four predictor variables (i.e., decoding, oral language, working memory, ELP) revealed that oral language and ELP were the most important, unique predictors of reading comprehension performance.

Within the previously discussed analyses, a sentence repetition task was used as the measure of oral language, which was consistent with the purpose for which the task was developed (Arañas & Christ, 2014) and with prior studies linking sentence repetition task performance to other measures of language ability (Klem et al., 2015). However, as a consequence of gaps in previous research, a degree of uncertainty remained as to whether sentence repetition tasks primarily measured oral language or working memory, especially within ELL samples, where the research was sparse. Therefore, another purpose of the present study was to investigate the contribution of oral language and working memory to explaining variance in ELLs’ performance on the OLSR. Furthermore, the researcher aimed to extend research in this area by examining the potentially varying roles that oral language and working memory may play in students’ scores on Spanish and English versions of the OLSR. For these analyses, the oral language composite of a comprehensive ELP measure (ACCESS) was used as the measure of oral language.
As expected, results indicated that both the ACCESS oral language composite and working memory were significant, unique contributors to students’ performance on the English OLSR. However, further analyses revealed that neither of these variables were significantly stronger predictors of students’ performance on the English OLSR than the other, thus suggesting that English OLSR performance is closely linked to both oral language and working memory scores for ELLs, but is not clearly a measure of one of these skills over the other. This result is inconsistent with findings from Klem et al. (2015), where results suggested that in a native English-speaking sample, performance on a sentence repetition task was indicative of language ability rather than a separate memory construct. However, the present study differs from Klem et al. in that an ELL sample was used and a direct measure of working memory was administered.

When similar analyses were conducted with the Spanish OLSR as the criterion measure, only working memory explained significant variance in student performance. The ACCESS oral language composite did not explain significant variance. This finding is possibly due to the fact that the ACCESS oral language composite measures English oral language skills, not Spanish language skills. These findings support those of past research suggesting the importance of a link between assessment languages for consistent results (Gottardo & Mueller, 2009). Furthermore, although skills acquired in one language generally transfer to a second language, some study participants may not have had a strong proficiency with the Spanish language, given that all of their academic instruction was provided in English and a small percentage of the sample was reported to speak only some Spanish at home, resulting in a weak relationship between the Spanish OLSR and the English measure of oral language. In fact, the Spanish
OLSR generally had weaker relationships with all study measures administered in English than did the English OLSR.

Although the relationships were not as strong as for the English OLSR, many significant intercorrelations were observed between the Spanish OLSR and other study measures, including English OLSR and reading comprehension, which was inconsistent with findings from Gottardo and Mueller (2009). Gottardo and Mueller found that for first and second grade Spanish-speaking ELLs who received literacy instruction in English, Spanish measures of oral language did not significantly correlate with English measures of oral language or English measures of reading comprehension. This inconsistency with present findings could be due to differences in types of oral language measures used; Gottardo and Mueller incorporated vocabulary knowledge with syntactic processing to measure oral language and in the present study only syntactic processing (i.e., OLSR) was used.

Although working memory was significantly related to performance on both versions of the OLSR, it was more strongly related to the English version. Consequently, for ELLs, working memory played a larger role when attempting to repeat sentences in their second language than when attempting to repeat sentences in their first language. This may be because sentence repetition task performance generally improves when vocabulary knowledge improves and the sentences are fully understood (Komeili & Marshall, 2013). Thus, ELLs with lower ELP who may not fully understand the English sentences must rely more heavily on working memory for the English as opposed to the Spanish OLSR. In extrapolating these findings to ELL performance on English reading comprehension measures, one could expect that ELLs’ working memory may be heavily taxed when attempting to repeat or understand English text because the understanding of the words they are reading may be poor.
Working memory was an important variable that was related to performance on both versions of the OLSR; however, the OLSR is more than just a measure of working memory. Strong relationships were detected between the OLSR tasks and reading comprehension, decoding, and ELP. Thus, significant predictive value was observed to stem from performance on the OLSR.

Limitations and Future Directions

Results from the present study should be interpreted in light of a number of limitations. First, although the researcher sought out participants with a range of ELP levels, none of the participants scored at the lowest extreme on the ELP measure, with the lowest ACCESS score of any participant being 2.1. This was largely due to the fact that all participants had been receiving English language services through the school system for at least one year. Furthermore, it was important that participants had at least basic knowledge of English so that they could understand task instructions provided in the present study. A second limitation was that the ELP measure utilized in this study (ACCESS) was not administered by the researchers but rather by school personnel. Therefore, the researcher is not privy to the fidelity of administration. However, there are strict guidelines and training procedures for administration that are defined by the ACCESS test creators. Furthermore, ACCESS is a computer-administered measure, which also allows for increased fidelity of administration. A related limitation is that the ACCESS was administered to all participants during a seven-week test window in winter 2016, approximately two to four months prior to collection of the other measures employed within the study.

Another limitation of the study was that all data were collected with students within one school district in Georgia, and all ELLs spoke Spanish as their first language, thus limiting the generalizability of the findings. However, similarities amongst participants also allowed for less
error in data analyses. Additionally, the researcher’s understanding of participants’ home language environment was limited in that data were gathered via parent report on the home language survey. Therefore, the descriptive statistics for the sample may be unreliable.

A final limitation of the current study is that the two OLSR measures may not have been equivalent in level of difficulty. Although identical procedures were used as part of both the English and Spanish OLSR, students’ performance on the two tasks could not be compared due to the forms not being equivalent in number of items and the inability to equate sentences.

**Summary and Implications**

Results from the current study provide support for intervention and measurement in the areas of decoding, oral language, working memory, and ELP for young ELLs, given the close relationship that each of these variables demonstrated with reading comprehension outcomes. Present findings provided further support for the SVR, seeing as decoding and oral language together explained 58.6% of the variance in reading comprehension outcomes. However, results suggest that further research is needed because an alternative model may more comprehensively explain variation in reading comprehension performance for ELLs. In the present study, adding working memory and ELP (ACCESS Overall Composite) to the SVR model explained an additional 11.6% of variance in reading comprehension performance; however, closer examination of this full model indicated that decoding and working memory no longer added uniquely in the model because ACCESS and the English OLSR were stronger predictors. It may however be that the ACCESS score employed included measures of basic reading and comprehension. Similarly, the English OLSR, which is closely related to working memory, may have rendered working memory insignificant.
Past research has largely discounted the importance of ELP to reading comprehension outcomes, seeing as few prior studies have desegregated ELL groups by ELP level, or have done so using crude measures such as qualification for English language services. Findings from the present study suggest that more sophisticated measures of ELP can add value to understanding ELLs’ reading comprehension performance. For this reason, it is important to consider the ELP level of ELLs and adjust expectations and teaching strategies based on this knowledge.

The present study more closely examined what the OLSR was measuring, as well as the value it may have in predicting reading outcomes. Although the Spanish OLSR demonstrated value in predicting reading outcomes, the relationship was stronger for the English OLSR. Results indicated that performance on the English OLSR was closely linked to reading comprehension performance, as well as to a whole host of other variables that correlate with reading comprehension (e.g., working memory, oral language, decoding, ELP). This suggests that the OLSR is a valuable tool for use in schools, especially given the ease with which it can be administered. By making use of the OLSR and other predictive measures highlighted in this study, school personnel can gain a more complete understanding of how their ELL students are performing and skill areas in which they may be lacking. Subsequently, intervention efforts can be targeted to these specific areas of weakness, and English reading comprehension outcomes can be improved.
References


