ADDRESSING THE LEXICAL QUALITY HYPOTHESIS
AND LANGUAGE COMPREHENSION IN FIRST AND
SECOND LANGUAGE LEARNERS

by

Megan O'Connor

A thesis submitted in conformity with the requirements
for the degree of Master of Arts
Department of Applied Psychology and Human Development
Ontario Institute for Studies in Education
University of Toronto

© Copyright by Megan O'Connor (2016)
ADDRESSING THE LEXICAL QUALITY HYPOTHESIS AND LANGUAGE COMPREHENSION IN FIRST AND SECOND LANGUAGE LEARNERS

Megan O'Connor

Master of Arts

Department of Applied Psychology and Human Development
University of Toronto

2016

Abstract

This study investigated whether there was evidence to support both the lexical quality hypothesis and simple view of reading in a grade 5 diverse sample of 94 English monolingual speakers (EL1) and 178 English Language Learners (ELL). Latent profile analyses conducted with the language groups together and separately revealed clusters consisting of good comprehenders, poor language, and poor word-level skills, with a fourth cluster of average comprehenders for the ELL group. Support was found for the lexical quality hypothesis as good comprehenders showed strong performance across phonological, orthographic and semantic component skills, while poor comprehenders showed profiles consistent with those of dyslexia or language impairment. In support of the simple view of reading, language comprehension emerged as a distinct deficit for those in the poor language group. Finally, more ELLs were identified as having poor language in the whole sample analysis than when compared only to other ELLs.
Acknowledgments

Firstly, I would like to thank my supervisor Dr. Esther Geva for her invaluable guidance and support throughout the process of developing and writing my thesis. Without her expertise and knowledge in the field of reading research, this thesis would not have been possible. Thank you also to Emiko Koyama for her help and input with the statistical analysis during the multiple stages of my data analysis. Additionally, I would like to thank Dr. Becky Chen-Bumgardner, my second reader, for reviewing this thesis.

Finally, I would like to thank both my friends and family for their unending support throughout my Masters. My SCCP cohort consists of some of the most supportive, intelligent, compassionate, and hard-working women that I have ever met, and I am so thankful to have them along for the ride. A special thanks is also due to Ilana, my library buddy, for keeping me company until the very end. Most importantly, I would like to thank my family for their endless support, and for instilling in me a belief that I can do whatever I set my mind to.
Table of Contents

Abstract.................................................................................................................................................. ii

Acknowledgements ................................................................................................................................. iii

Table of Contents ........................................................................................................................................ iv

List of Tables ........................................................................................................................................ vi

List of Figures ........................................................................................................................................ vii

Introduction ........................................................................................................................................... 1

Methodology .......................................................................................................................................... 12

Participants ........................................................................................................................................... 12

Measures .............................................................................................................................................. 13

Results .................................................................................................................................................. 15

Preliminary Analysis ............................................................................................................................... 15

The Relation Among Reading and Language Variables Across Language Groups........ 17

Determining the Models of Best Fit for Latent Profile Analysis .................................................... 18

ELL Latent Profile Analysis Clusters.................................................................................................. 20

Latent Class Group Differences for ELLs ........................................................................................... 22

EL1 Latent Profile Analysis Clusters.................................................................................................... 24

Latent Class Group Differences for EL1s ............................................................................................ 26

Whole Sample Latent Profile Analysis ................................................................................................. 28

Discussion ............................................................................................................................................ 32

Differences Between the ELL and EL1 Groups on Reading and Language Measures .......... 32

Latent Profiles and Support for the Lexical Quality Hypothesis....................................................... 33

Thinking of Vocabulary and Language Comprehension .................................................................... 36

Do Identical Latent Profile Groups Exist for the ELL and EL1 Groups? ............................................ 37

Comparing Analysis of the Whole Sample vs. by Language Status .................................................. 38
Limitations and Future Research ................................................................. 39
Research Implications .................................................................................... 40
References ........................................................................................................... 42
List of Tables

Table 1. Descriptive Statistics and Group Comparisons for English Monolingual Speakers (EL1) and English Language Learners (ELL) ................................................................. 17

Table 2. Correlations Among the Reading and Language Variables for the EL1 and ELL groups ................................................................................................................. 18

Table 3. Latent Class Analysis Fit Statistics for the ELL and EL1 groups ...................... 20

Table 4. Descriptive Statistics and MANOVA Summary Statistics for the ELL Group ... 22

Table 5. Descriptive Statistics and MANOVA Summary Statistics for the EL1 Group ... 26

Table 6. Latent Class Analysis Fit Statistics for the ELL and EL1 Language Groups Together .................................................................................................................. 29
List of Figures

Figure 1 a-d. Mean Z-scores and 95% Confidence Intervals for the ELL Latent Profile Groups ................................................................. 21

Figure 2 a-c. Mean Z-scores and 95% Confidence Intervals for the EL1 Latent Profile Groups ................................................................. 25

Figure 3 a-c. Mean Z-scores and 95% Confidence Intervals for the Whole Sample Latent Profile Groups ................................................................ 30

Figure 4 a-b. The Number of ELLs and EL1s in Each Latent Class Group When Analyzed Together Versus Separately .............................................................. 31
Introduction

Reading comprehension is an essential academic skill that can impact both the educational and career trajectories of students (Savolainen et al., 2008). English Language Learners (ELLs), or students who begin school speaking a language other than English, are often at a disadvantage compared to their English as a first language (EL1) peers with regards to reading comprehension (August, Shanahan, & Escamilla, 2009). While ELLs are likely to reach equal levels to EL1s on skills such as decoding and spelling (Nakamoto, Lindsey & Manis, 2007), ELLs often continue to lag behind on vocabulary (Farnia & Geva, 2011) and text-level skills such as reading comprehension (Farnia & Geva, 2013); this is illustrated by the finding that ELLs are more likely to have difficulties comprehending grade-level texts than EL1s (Geva & Farnia, 2012; August et al., 2005). Studies have shown that increasing the vocabulary of ELLs can contribute to improved reading comprehension (Carlo et al., 2004; Lesaux et al., 2010), but the exact processes that underlie how experience with individual words contributes to reading comprehension in ELLs warrants further study.

Theories of Reading Comprehension

One theory that argues for the importance of our experiences with individual words in relation to reading comprehension is the lexical quality hypothesis (Perfetti & Hart, 2002). The lexical quality hypothesis espouses that in order to have successful reading comprehension, high-quality representations of individual words are necessary. A word representation is considered to be high quality if it integrates phonological, orthographic and semantic information. If any of these three constituents is lacking, the word representation is of a lower quality and it takes more effort to retrieve the representation, leaving less cognitive resources available for comprehension. Relatedly, a word representation is more likely to become high quality when there have been multiple encounters with the word.
Perfetti and Hart found support for the lexical quality hypothesis in an English-speaking undergraduate sample by dividing the students into less skilled, average, and high skilled readers on the basis of their reading comprehension score, and analyzing the factor structure of the groups. They found evidence of more coherent lexical structures for the high skilled group, with higher inter-correlations among measures of phonological, orthographic and semantic knowledge for the high skilled group than the less skilled comprehenders (Perfetti & Hart, 2002). They also found that high skilled readers had a greater bank of high quality word representations than less skilled readers, as evidenced by faster decisions on timed word choice tasks, for both control words and homophones (Perfetti & Hart, 2002).

In parallel to the lexical quality hypothesis, a similar conceptual framework, the simple view of reading (SVR), was proposed in New Zealand by Gough and Tunmer (1986). According to the SVR, in order to comprehend the texts they read, individuals need to have fluent word decoding skills and well-developed language comprehension (Hoover & Tunmer, 1993). The simple view of reading predicts that different reading difficulties arise depending on where a reading deficit lies; for example, individuals who are poor decoders have difficulty correctly mapping phonemes onto graphemes, which interferes with word reading, thus impacting comprehension. In contrast, individuals with poor language comprehension can decode the words on the page, but have difficulty understanding what they read, which also impedes reading comprehension.

Like the lexical quality hypothesis, the SVR framework has been used in numerous studies targeting reading comprehension. There is ample research evidence supporting the SVR as a useful framework for understanding key causes of poor reading comprehension in English monolingual speakers (e.g., Bast & Resitsma, 1998; Carver, 1993; Joshi & Aaron, 2000; Savage & Wolfforth, 2007). More recently, studies have been conducted to investigate whether the
simple view of reading model can also explain the reading comprehension of second language learners. For example, Gottardo and Mueller (2009) found support for the SVR in that both word reading and English oral language proficiency were the strongest predictors of reading comprehension for grade 2 Spanish-speaking ELLs. In another study of Spanish-speaking ELLs, Nakamoto, Lindsey, and Manis (2007) found that both decoding and oral language ability (measured by two tasks – sentence recall and expressive vocabulary) significantly predicted reading comprehension. These studies suggest that the SVR can adequately explain individual differences in ELL reading comprehension, and that the consideration of language comprehension in relation to reading is important.

Other research has suggested that the nature of the relationships between the different variables of the SVR changes from the primary to middle school grades (Sticht and James, 1984). In the early primary grades, both decoding and language comprehension correlate with reading comprehension, but the relationship between decoding and reading comprehension is stronger probably because as Chall (1986) has argued, children are still in the “learning to read” stage. As children enter the middle school grades their word related skills, including accuracy and fluency in word reading, and the ability to decode unfamiliar words become well developed, and there is less individual differences on these skills. Concomitantly, language comprehension components typically show a stronger relationship with reading comprehension. In other words, the relationship between the constructs forming the SVR are dynamic and changing (Tilstra et al., 2009); by the middle school grades word reading related skills of typically achieving children are not an issue, but they continue to vary in their command of various aspects of language comprehension, and these continue to play a particularly important role in reading comprehension.
Both the SVR and lexical quality hypothesis attempt to explain reading comprehension by identifying the underlying skills involved. Both frameworks acknowledge the importance of decoding in comprehension; however, where they appear to diverge is with the other components of the frameworks. The SVR focuses on the broader concept of language comprehension, which includes how components such as syntax (Swanson et al., 2008; Gottardo & Mueller, 2009), morphology (Siegel, 2008), listening comprehension (Proctor et al., 2005), and vocabulary (Farnia & Geva, 2013) impact reading comprehension. This is in contrast to the lexical quality hypothesis, which focuses more specifically on semantic knowledge or vocabulary (Perfetti & Hart, 2002). Thus, while both theories include components of language comprehension, they differ in which aspects of language comprehension they focus on.

One study that provides support for both the lexical quality hypothesis and the SVR in a population of L1 children was conducted by Verhoeven and Leeuwe (2008). They tested the lexical quality hypothesis and SVR frameworks in a longitudinal study with Dutch children from grades 1 to 6 and found support for the lexical quality hypothesis in that both knowledge of word forms (i.e. decoding) and vocabulary predicted the development of reading comprehension. They also found support for the SVR in that both decoding and other aspects of language beyond vocabulary knowledge, captured by a task that evaluated listening comprehension, also predicted reading comprehension. Decoding exerted a greater impact on reading comprehension in the early grades, while listening comprehension showed a reciprocal relationship with reading comprehension in later grades. While the majority of studies assessing the lexical quality hypothesis have centered around the reading abilities of adults, this study suggests that the lexical quality hypothesis can be applied to a population of children, even though their reading skills are still developing. Given that this study was conducted with first language (L1) children, it highlights the need for a study that simultaneously compares both vocabulary and additional
language comprehension skills in relation to reading comprehension with a second language (L2) sample.

To the best of my knowledge, no study has yet to examine whether the lexical quality hypothesis also applies to an ELL schoolchildren population, or has compared the lexical quality hypothesis with the SVR in an ELL sample. There is, however, a considerable amount of research that has evaluated the individual roles of the three lexical quality hypothesis constituents (phonological awareness, orthographic processing and semantic knowledge) and language comprehension in relation to reading comprehension for both EL1 and ELL children, which is discussed below.

**Phonological awareness.** Phonological awareness (PA) is commonly defined as the ability to identify and manipulate the sound units of spoken words (Stanovich, 1992). The important role that PA plays in learning how to decode in an alphabetic orthography is well established in the EL1 literature (e.g., Catts, 1991; Vellutino & Scanlon, 1991; Bus & van IJzendoorn, 1999). Research has consistently found that individuals who struggle with PA often have difficulties learning how to decode (Bryant, 1991; Catts, 1991; Elbro, 1996). Research also suggests that a similar relationship exists for ELLs; ELLs who struggle on measures of English PA also have difficulties decoding (Gottardo et al., 2001; Lesaux & Geva, 2006; Lindsey, Manis & Bailey, 2003). There is also some research to suggest that PA plays a direct role in reading comprehension for both ELLs and EL1s; this is supported by the finding that PA in grade 2 (along with vocabulary) predicted reading comprehension in grade 5 for both ELLs and EL1s (Chen, Geva & Schwartz, 2012).

Studies of cross-language transfer of phonological skills have found that PA in the first language predicted performance on English word and pseudoword reading (Durgunoğlu, Nagy, & Hancin-Bhatt, 1993), and that L1 and L2 phonological skills are highly correlated (Geva,
This supports the notion that cross-language transfer occurs for phonological skills. Research has also shown that ELLs often perform similarly to their EL1 peers on basic cognitive skills such as PA (Geva, Yaghoub-Zadeh, & Schuster, 2000), and that even those ELLs who begin school with lower PA than their EL1 peers catch up within a few years (Lesaux, Rupp, & Siegel, 2007). Morrow et al. (2014) found that similar to EL1s, the phonological skills of ELLs also increase over time, and that for ELLs this increase occurs as a function of the age at which English learning began. This body of research suggests that for typically developing ELLs who begin school speaking a different first language, we can expect that their phonological skills will be on levels similar to their EL1 peers after a short amount of time.

Orthographic processing. Orthographic processing, defined as “the ability to form, store and access orthographic representations” (Stanovich & West, 1989, p.404), is an important component of both written expression (Russak, S., & Kahn-Horwitz, 2015), and reading (Barker, Torgesen, & Wagner, 1992). While a considerable amount of attention has been paid to the development of PA in ELL children, much less research has been devoted to the study of ELL orthographic skills or spelling (Geva, 2006).

In EL1 populations, research suggests that orthographic ability is an important contributor to word reading. For example, one study with a sample of EL1 third-grade children found that orthographic ability accounted for 10.2% of the variance on a standardized word-reading task after controlling for age, nonverbal ability and PA (Cunningham & Stanovich, 1991). This indicates that orthographic skills contribute uniquely to word reading beyond that of PA. Additionally, orthographic ability of EL1s as measured in preschool was found to significantly predict reading comprehension in both grade 3 and grade 7, with an even greater proportion of variance explained with increasing grade level (Badian, 2001).
these studies suggest that EL1 orthographic ability contributes directly to both word reading and reading comprehension. Research has also shown that reading and orthographic processing are highly correlated for EL1s, developing in tandem (Conrad, 2008).

Similar to EL1s, research suggests that word reading and spelling are highly correlated for ELLs as well, and that this is true for ELLs of different first language backgrounds (e.g., Fender, 2008; Kahn-Horwitz, Sparks, & Goldstein, 2012; Wang & Geva, 2003). Studies comparing the spelling performance of primary level ELLs to their EL1 peers generally find that ELLs perform similarly to EL1s on measures of spelling even in the early primary grades. For example, Wade-Woolley and Siegel (1997) found that there was no difference in the spelling abilities of grade 2 ELLs and EL1s, and that for both groups, spelling was similarly correlated with PA and pseudoword decoding. Additionally, the report of the National Literacy Panel on Language Minority Children and Youth found that the word-level skills (including spelling) of ELLs are significantly more likely to be at equal levels to EL1s compared to text-level skills such as reading comprehension (August & Shanahan, 2006). This shows that ELLs and EL1s do not differ on orthographic ability after the early primary years, and that one can expect typically developing ELLs to perform similarly to EL1s on measures of spelling.

**Semantic knowledge.** It is well established in the EL1 literature that semantic knowledge, or more specifically vocabulary, plays a critical role in reading comprehension (e.g., Biancarosa & Snow, 2004; National Institute of Child Heath and Human Development, 2000; Ouellette, 2006). Research suggests that by the middle elementary school grades, semantic and syntactic knowledge explain considerably more variance in reading comprehension than do decoding skills (Vellutina & Denckla, 1991). A considerable body of research also supports a significant relationship between vocabulary and reading comprehension for ELLs. For example, in two studies involving Spanish speaking L1, both Proctor et al. (2012) and Manis, Lindsey, and
Bailey (2004) found that the vocabulary of primary school age ELLs significantly predicted English reading comprehension, and Carlo et al. (2004) found that systematic English vocabulary instruction improved ELL reading comprehension to levels equal to those of an EL1 control group. Similar findings have been reported in other language contexts, such as with Hong Kong Chinese children learning English (Li et al., 2012), and with Turkish and Moroccan children learning Dutch (Droop & Verhoeven, 2003). Clearly, vocabulary plays an important role in reading comprehension for both first and second language learners.

Another consistent and important finding is that ELLs often lag behind EL1s on semantic knowledge or vocabulary, including deficits in both breadth (August et al., 1999) and depth of vocabulary (Jean & Geva, 2009; Verhallen & Schoonen, 1993). This gap has been found in the primary school grades (Umbel et al., 1992), and evidence suggests that this gap does not close even by middle school (Farnia & Geva, 2011; Hutchinson, Whiteley, Smith, & Connors, 2003).

In a longitudinal study comparing the vocabulary and reading comprehension of first and second language learners from grade 2 to grade 3, Lervåg and Aukrust (2010) found that first language learners had better reading comprehension scores than second language learners at both time points, as well as faster growth over time. While both decoding and vocabulary predicted reading comprehension at grade 2, only vocabulary predicted the later growth in reading comprehension. Interestingly, vocabulary was a stronger predictor of growth in reading comprehension for second language learners than first language learners, which suggests that there are differences between the two groups with regards to the impact of vocabulary on reading comprehension.

In an investigation of ELL reading performance up to the middle school years, a longitudinal study by Farnia and Geva (2011) followed ELLs from grades 1 to 6 and found that even though over time the gap between EL1 and ELLs’ vocabulary gradually diminished, it continued to be significantly poorer than that of their EL1 peers across all grade levels. In other
words, even after six years of schooling in English, ELLs have still not reached a command of English vocabulary that is comparable to that of their EL1 peers. When readers are faced with too many unknown words in a text, comprehension becomes extremely difficult (Carver, 1994); this places ELLs at significantly greater risk than EL1s of lacking the English vocabulary they need to understand grade-appropriate texts (August et al., 2005). These studies indicate that vocabulary is an important contributor to reading comprehension, and that ELLs are at a significantly greater risk than EL1s of having poor vocabulary. This research also suggests that the relationship between vocabulary and reading comprehension may not be identical for both language groups.

**Language comprehension.** In addition to the phonological, orthographic and semantic constituents highlighted in the lexical quality hypothesis, a significant body of research has emphasized the importance of language comprehension in relation to reading comprehension for EL1s as well as ELLs. As discussed earlier, the construct of language comprehension includes many different components of oral language proficiency, including syntax, morphology, and listening comprehension. For EL1 populations, numerous studies have found that after the primary grades, listening comprehension is an important contributor to reading comprehension (e.g., Catts et al., 2002; Hoover & Gough, 1990; Savage, 2001; Tilstra et al., 2009). Catts and Weismer (2006) also found that language comprehension (including measures of receptive vocabulary, syntax and listening comprehension) contributed significantly to reading comprehension in a middle-school age sample.

The role of language comprehension in reading comprehension has also been shown for ELLs. A study by Royer and Carlo (1991) with a Spanish-speaking ELL sample reported positive correlations between English listening comprehension and reading comprehension, with listening comprehension in grade 5 being one of the strongest predictors of grade 6 reading
comprehension. Another study with an adolescent Spanish-speaking ELL population found that English listening comprehension and vocabulary definitions together explained 50% of the variance in reading comprehension scores (Carlisle, Beeman, & Shah, 1996). A study by Proctor et al. (2005) using a structural equation model of reading comprehension found that for grade 4 ELLs, oral language proficiency (measured by both vocabulary and listening comprehension) was a significantly better predictor than decoding skills. Finally, in a longitudinal study of Spanish-speaking ELLs from kindergarten to grade 8, Kieffer (2012) found English vocabulary to better predict reading comprehension in the later grades than that of language comprehension (consisting of both listening comprehension and story retelling). This study would suggest that it is not always the case that language comprehension explains reading comprehension over and above vocabulary. However, Farnia & Geva (2013) showed that vocabulary was an early predictor of later reading comprehension, but with an increase in language proficiency in later years more individual differences emerged, and then syntax and listening comprehension became additional predictors of reading comprehension.

A recent meta-analysis of studies analyzing reading comprehension and the underlying processes of phonological awareness, decoding, and language comprehension (with language comprehension including both vocabulary and listening comprehension) in L1 and L2 children found that the greatest gap between the groups existed on language comprehension (Melby-Lervåg & Lervåg, 2014). Given that only small differences between the L1 and L2 groups existed for PA or decoding, the authors concluded that reading interventions aimed at improving second language reading comprehension should focus on language comprehension, unless specific decoding problems were found. Similarly, a meta-analysis of variables correlated with second language reading comprehension found one of the strongest correlations ($r = .77$) to exist between second language listening comprehension and reading comprehension (Jeon &
Yamashita, 2014). Together, these two meta-analyses provide further support for the role of language comprehension in second language reading comprehension.

Taken together, these studies suggest that by the middle school years, the language-based variables that contribute most significantly to reading comprehension have changed from the early primary years. This body of research supports the idea that lexical knowledge often measured with vocabulary tasks continues to play a pivotal role in reading comprehension. At the same time, other aspects of language comprehension that address skills such as syntactic knowledge and listening comprehension also contribute to reading comprehension among the middle school grades; in order to have a full picture of the component skills underlying reading comprehension, language comprehension should not be overlooked.

**The Current Study**

So far the lexical quality hypothesis has not been tested with ELL and monolingual children in middle school, at a pivotal time when they are expected to rely on their existing language and word reading skills to learn from texts. The current study was designed to address this dearth of research and investigate whether there is support for the lexical quality hypothesis as an explanatory framework of reading comprehension of ELLs and EL1s with different reading comprehension profiles, and whether a better understanding of reading comprehension can be reached by considering, in addition, language comprehension. Latent profile analysis was used to determine the underlying group profiles that exist among ELLs and EL1s. The following research questions were addressed:

(1) Is there evidence to support the lexical quality hypothesis in an ELL and EL1 grade 5 population? More specifically, how well do ELL and EL1 poor comprehenders and good comprehenders do on measures of phonological, orthographic, and semantic skills?
(2) Is language comprehension also important for differentiating poor comprehenders from good comprehenders?

(3) Are the emerging profiles similar among ELLs versus EL1s?

(4) Are similar latent groups identified when ELLs and EL1s are analyzed together versus separately, and is the proportion of ELLs and EL1s identified in each group the same?

Methodology

Participants

The data used for the current study were taken from a larger, multi-cohort longitudinal study examining the development of reading in both ELLs and EL1s. The study began in 1996, recruiting grade 1 students from 14 different schools across 4 school boards in the Greater Toronto Area. The communities comprising each of the school boards were ethnically and linguistically diverse. The majority of the schools were located in neighbourhoods of low to middle socioeconomic status. Students were then followed over the next six years, completing a significant test-battery of various reading, language and cognitive measures. Four cohorts of students were recruited, with a final overall sample of 614 participants. Students with any known developmental disability or impairment were identified by the school and excluded from the study. Only students who had lived in an English-speaking country for at least four months at the onset of the study were allowed to participate, in order to ensure a minimum level of English oral language proficiency necessary to understand task instructions. Parental consent was obtained for all participants.

Data collected when students were in Grade 5 were used for the current study. The overall sample size was 272 students, consisting of both EL1 students \( n = 94 \) and ELL students \( n = 178 \) from diverse ethnic and linguistic backgrounds. Home languages for the ELL students included Punjabi \( n = 70 \), Portuguese \( n = 59 \), Tamil \( n = 28 \), Urdu \( n = 6 \), Hindi \( n = 4 \),
Gujarati \((n = 3)\), Mandarin \((n = 2)\), Cantonese \((n = 3)\), and Other \((n = 3)\). The average age at the time of testing was 10 years, 8 months \((SD= 3.72 \text{ months})\) for EL1s, and 10 years, 7 months \((SD=3.68 \text{ months})\) for ELLs \(t(270)=3.28, p<.01\).

**Measures**

Measures relevant to the current study were selected from the total battery of tests administered. Tasks were selected that assessed reading comprehension and listening comprehension. Tasks were also selected that assessed each of the components of the lexical quality hypothesis: phonological awareness, orthographic processing, and semantic knowledge. However, as discussed by Perfetti and Hart (2002), it is very difficult to find measures that isolate each of these constituents, and most of the tasks involve an overlap of knowledge from multiple constituents. While none of the tasks used in the current study can be considered “pure” measures of the individual constituents, it is possible to conclude that they require to some extent knowledge from one or more constituents.

**Phonological awareness.** Phonological awareness was assessed using the Auditory Analysis Skills Task (Rosner & Simon 1971). This test was a precursor to the CTOPP elision subtest, which was not yet available at the time this data was collected. The Auditory Analysis Skills Task is an elision task consisting of 25 items, increasing in difficulty. Participants are required to delete syllables or phonemes from spoken words. The test was terminated when the child made five consecutive errors. The first set of items involved deletion of a syllable from either the beginning or ending of a word (e.g. ‘Say “baseball”; Now say it again but don’t say “base”’). The second set of items involved deletion of the beginning or ending phoneme from one-syllable words (e.g. ‘Say “meat”; Now say it again but don’t say /m/’). The last set involved deletion of a phoneme from a consonant blend in a word (e.g. ‘Say “stop”; Now say it again but don’t say /s/’). This test was considered to primarily assess phonological knowledge. The
Cronbach’s alpha for this task is .90.

**Orthographic processing.** Orthographic processing was assessed using the Wide Range Achievement Test-Revised (WRAT-R) Spelling subtest (Wilkinson, 1993). This task was administered in a group format, and required participants to write down the correct spelling of words that were read out loud to them. The test consisted of 40 items, increasing in difficulty. Each word was read out loud, given in a sentence, and then repeated again. All 40 items were administered to each participant. This test was considered to primarily assess orthographic knowledge. The Cronbach’s alpha for this task is .88.

**Semantic knowledge.** Participants’ understanding of the meaning of individual words was assessed using a shortened version of Biemiller’s Written Root Word Inventory task (Biemiller & Slonim, 2001). Participants were given a sheet containing individual sentences, with the target word in bold. Participants were instructed to read the whole sentence, and then write a sentence that explains what the word in bold means. The task consisted of 30 items, and participants completed all the items. Target words ranged from easy (e.g. ‘dropped’) to more difficult (e.g. ‘trawl’). The definitions were scores by trained graduate students and research assistants. The Cronbach’s alpha for this task is .71.

**Reading comprehension.** The Gates-MacGinitie Reading Test (MacGinitie & MacGinitie, 1992) Level D5/6 form 4 for grade 5 was used to assess participants’ reading comprehension ability. Participants independently read short passages that consisted of both expository and narrative texts, and answered a series of multiple-choice questions to assess their understanding. The entire test consisted of 48 questions. The test was administered in a group format, and participants were given 35 minutes to complete all of the questions. The Cronbach’s alpha for this task is .88.

**Listening comprehension.** The Listening to Paragraphs subtest from the Clinical
Evaluation of Language Fundamentals- Third Edition (CELF-3; Semel, Wigg, & Secord, 1995) was administered to assess students’ listening comprehension ability. Four passages were presented orally to the children. After listening to each passage, the children answered five questions that assessed their ability to identify the main idea and details, understand the sequence of events, and make inferences and predictions. Total scores could range from 0-20. The Cronbach’s alpha for this task is .83

**Nonverbal ability.** Raven’s Standard Progressive Matrices was used to assess overall nonverbal reasoning ability in both the ELL and EL1 samples (Raven, Raven, & Court, 1998). This measure was included to ensure that no differences existed between the language groups on a measure of general nonverbal ability. Raven’s is a widely used measure of nonverbal reasoning that is relatively free of cultural and linguistic influences, and has been used with children from diverse cultural and linguistic backgrounds (Raven, Raven, & Court, 1998). The test consists of 60 multiple-choice questions in which students are shown a visual puzzle and asked to choose the picture that best completes the puzzle. The test is divided into five sets, with twelve items for each set. Children are asked to point to the correct picture, thus involving very little language demands apart from understanding the initial instructions.

**Results**

**Preliminary Analysis**

Given that the norms used to standardize the measures used in this study were based on EL1 samples, raw scores rather than standardized published normed scores were used in subsequent analyses. Where relevant (e.g., to compare scores across different measures) raw scores were converted into standardized z-scores separately for ELLs and EL1s.

The primary approach to addressing the research questions was latent profile analysis. It was first necessary to establish whether the ELL and EL1 groups should be analyzed separately
or merged. To this end the correlation matrices of the ELL and EL1 groups were compared using the Box’s M test. This was supplemented by comparing individual correlations using Fisher’s exact test, and comparing raw scores of the ELL and EL1 groups across each of the measures. In comparing the two correlation matrices, an overall trend appeared to show slightly higher correlations across measures for ELLs than EL1s (see Table 2). Box’s M test showed that overall, the ELL and EL1 correlation matrices did not differ significantly from each other (p=.16). However, Fisher’s exact test showed a significantly higher correlation (p < .005) between orthographic processing and semantic knowledge for ELLs (r=.597) than EL1s (r=.304).

Table 1 summarizes the descriptive statistics (means, standard deviations, and group differences) of the reading, language and cognitive measures. Group comparisons showed that EL1 and ELL students did not differ significantly on the measures of phonological awareness or orthographic processing. The test of group differences for listening comprehension was close to significance (p = .06), indicating that a difference may exist. There was also no significant difference between the language groups on nonverbal ability, indicating that no difference in overall cognitive ability exists between the ELL and EL1 groups. However, as expected, EL1s outperformed ELLs on the measures of semantic knowledge and reading comprehension. That is, by grade 5, a gap still exists between EL1 and ELL students with regards to vocabulary and reading comprehension, and thus confirms the ELL designation. Given that differences in English language proficiency still existed between the groups, and that the relationship between orthographic processing and semantic knowledge was significantly stronger for ELLs than EL1s, subsequent analyses were performed separately for each language group.
Table 1. Mean Raw Scores, Standard Deviations, and Group Comparisons for English Monolingual Speakers (EL1) \((n = 94)\) and English Language Learners (ELL) \((n = 178)\).

<table>
<thead>
<tr>
<th>Measures</th>
<th>EL1</th>
<th>ELL</th>
<th>Test of Group Differences</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonverbal Ability</td>
<td>35.52 (8.90)</td>
<td>33.90 (8.43)</td>
<td>1.48</td>
<td>.18</td>
</tr>
<tr>
<td>Phonological Awareness</td>
<td>19.15 (5.54)</td>
<td>19.99 (5.08)</td>
<td>-1.27</td>
<td>-.15</td>
</tr>
<tr>
<td>Orthographic Processing</td>
<td>15.80 (5.02)</td>
<td>16.76 (5.45)</td>
<td>-0.14</td>
<td>-.17</td>
</tr>
<tr>
<td>Semantic Knowledge</td>
<td>10.86 (4.46)</td>
<td>9.61 (3.74)</td>
<td>2.41*</td>
<td>.30</td>
</tr>
<tr>
<td>Listening Comprehension</td>
<td>11.91 (4.86)</td>
<td>10.81 (4.21)</td>
<td>1.87</td>
<td>.24</td>
</tr>
<tr>
<td>Reading Comprehension</td>
<td>25.21 (8.51)</td>
<td>23.02 (8.11)</td>
<td>2.08*</td>
<td>.25</td>
</tr>
</tbody>
</table>

*Note. The test of group differences for listening comprehension was \(p = .06\), thus close to significance. *: \(t\) test significant at the .05 level.

The Relations Among Reading and Language Variables Across Language Groups

Table 2 displays the correlations among phonological awareness (PA), orthographic processing, semantic knowledge, listening comprehension, and reading comprehension variables for the EL1s \((n = 94)\) above the diagonal, and the ELLs \((n = 178)\) below the diagonal. For EL1s, all of the variables were significantly correlated with each other. For ELLs, most of the variables were correlated with each other, while listening comprehension and orthographic processing were not correlated with each other. In general, most of the variables correlated moderately with each other, with a tendency for slightly higher correlations in the ELL group. The strongest correlations for both groups existed between PA and orthographic processing, as well as between semantic knowledge and reading comprehension. Of note are also the moderate correlations between semantic knowledge and listening comprehension for both language groups, suggesting that the semantic knowledge and listening comprehension measures assessed different component skills.
Table 2. Bivariate Correlations Among the Component Skills and Reading Comprehension Within Each Language Group (EL1s above and ELLs below the diagonal).

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Phonological Awareness</td>
<td><strong>1.00</strong></td>
<td>.555**</td>
<td>.279**</td>
<td>.269*</td>
<td>.418**</td>
</tr>
<tr>
<td>2. Orthographic Processing</td>
<td>.630**</td>
<td><strong>1.00</strong></td>
<td>.304**</td>
<td>.236*</td>
<td>.396**</td>
</tr>
<tr>
<td>3. Semantic Knowledge</td>
<td>.460**</td>
<td>.597**</td>
<td><strong>1.00</strong></td>
<td>.399**</td>
<td>.606**</td>
</tr>
<tr>
<td>4. Listening Comprehension</td>
<td>.214**</td>
<td>.116</td>
<td>.368**</td>
<td><strong>1.00</strong></td>
<td>.546**</td>
</tr>
<tr>
<td>5. Reading Comprehension</td>
<td>.411**</td>
<td>.514**</td>
<td>.612**</td>
<td>.484**</td>
<td><strong>1.00</strong></td>
</tr>
</tbody>
</table>

* Correlation is significant at the p < 0.05 level (2-tailed).
** Correlation is significant at the p < 0.01 level (2-tailed).

Determining the Models of Best Fit for the Latent Profile Analysis

Latent profile analysis (Bartholomew, 1987) was used to identify clusters of children with similar profiles on PA and orthographic processing, the two language proficiency measures (semantic knowledge and listening comprehension), as well as reading comprehension. Latent profile analysis is a model-based cluster analysis, in which continuous indicator variables are used to identify unmeasured class membership (Pastor et al., 2007). One of the greatest advantages of using latent profile analysis over traditional cluster analytic techniques is that it is not based on arbitrary criteria for cutoffs and group classification (Branum-Martin, Fletcher, & Stuebing, 2013). Instead, formal fit statistics are used to determine the best number of groups or clusters found in the dataset (Vermunt & Magidson, 2002). Another benefit is that there are fewer assumptions (such as normality of distributions and moderate correlations amongst indicator variables) that need to be met in order justify the analytic approach (Cohan et al., 2008).

MPlus 7.0 statistical software was used to determine the clusters that existed in the ELL and EL1 groups. Two, three and four-class models were compared to determine the best fitting models for the ELL and EL1 groups. Several indicators were used to determine the best model fit, including the Lo-Mendell-Rubin (LMR) test, the Akaike Information Criterion (AIC; Akaike,
1987), the Bayesian Information Criterion (BIC; Schwarz, 1978) and the Sample-size Adjusted Bayesian Information Criterion (SABIC; Sclove, 1987). Entropy values were also used in order to determine classification accuracy. The model of best fit was determined by choosing the model with a significant LMR test, small AIC, BIC, and SABIC values, as well as a high entropy value (Cohan et al., 2008).

Table 3 summarizes the model fit statistics for the ELLs and EL1s. For the ELL group, a statistically significant LMR test indicated that a four-class solution provided a better fit to the data than a three-class model. AIC, BIC and SABIC values were lowest in the four-class solution, and an entropy value of 0.83 indicated a high degree of classification accuracy. For the ELLs, a four-class solution emerged as the best model fit. For the EL1 group, a non-significant LMR test for a four-class solution indicated that a four-class solution did not provide a better model fit than a three-class solution. The three-class solution produced the second lowest AIC, BIC and SABIC values, as well as a high entropy value. For EL1s, a three-class solution was selected as the best model fit. Since the separate analyses for the ELL and EL1 groups produced different models of best fit, subsequent analyses were conducted separately for the two groups.
Table 3. Latent Class Analysis - Fit Statistics for Two, Three, and Four-class Models for the ELL and EL1 Groups.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Model</th>
<th>AIC</th>
<th>BIC</th>
<th>SABIC</th>
<th>LMR</th>
<th>Entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELL</td>
<td>Two-class</td>
<td>5175</td>
<td>5226</td>
<td>5175</td>
<td>p = .02</td>
<td>.80</td>
</tr>
<tr>
<td></td>
<td>Three-class</td>
<td>5086</td>
<td>5156</td>
<td>5086</td>
<td>p &lt; .01</td>
<td>.84</td>
</tr>
<tr>
<td></td>
<td>Four-class</td>
<td>5052</td>
<td>5140</td>
<td>5052</td>
<td>p = .03</td>
<td>.83</td>
</tr>
<tr>
<td>EL1</td>
<td>Two-class</td>
<td>2804</td>
<td>2845</td>
<td>2794</td>
<td>p &lt; .01</td>
<td>.99</td>
</tr>
<tr>
<td></td>
<td>Three-class</td>
<td>2754</td>
<td>2810</td>
<td>2740</td>
<td>p &lt; .01</td>
<td>.90</td>
</tr>
<tr>
<td></td>
<td>Four-class</td>
<td>2736</td>
<td>2807</td>
<td>2718</td>
<td>p = .10</td>
<td>.92</td>
</tr>
</tbody>
</table>

Note: AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; SABIC = Sample size Adjusted BIC; LMR = Lo-Mendell-Rubin likelihood ratio test.

ELL Latent Profile Analysis Clusters

Analysis of the fit statistics showed that a four-class solution provided the best model for the ELL group. Class 1 comprised 20.2% \((n = 36)\) of the total ELL sample. This group was characterized by high scores on the four component skills as well as reading comprehension; they were labeled the *good comprehenders* group (Figure 1a). Class 2 \((n = 89)\) comprised 50.0% of the total ELL sample. This group was characterized by average reading comprehension scores and generally average performance across the four component skills. This group was named the *average comprehenders* group (Figure 1b). Class 3 \((n = 24)\) represented 13.5% of the total ELL sample. This group consisted of students with average PA, but poor orthographic skills as well as the two language measures – vocabulary and listening comprehension; this group was labeled the *poor language* group (Figure 1c). Class 4 \((n = 29)\) comprised 16.3% of the total ELL sample. This group had difficulties across all four component skills as well as reading comprehension, with particular struggles in PA and orthographic processing. This group was labeled as the *poor word-level skills* group (Figure 1d). Mean standardized scores and standard deviations for each group are summarized in Table 4.
Figures 1 a-d. Mean Z-scores and 95% Confidence Intervals for the ELL Latent Profile Groups - Good Comprehenders (GC), Average Comprehenders, (AC), Poor Language (PL), and Poor Word-level Skills (PWL).

*Note: PA = phonological awareness; OP = orthographic processing; SE = semantic knowledge; LC = listening comprehension; RC = reading comprehension.
Table 4. Descriptive Statistics (Z-scores) and MANOVA Summary Statistics for the ELL Groups Defined by the Latent Profile Analysis.

<table>
<thead>
<tr>
<th></th>
<th>PA</th>
<th>OP</th>
<th>SE</th>
<th>LC</th>
<th>RC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>GC (n=36)</td>
<td>.75</td>
<td>.25</td>
<td>.93</td>
<td>.76</td>
<td>1.07</td>
</tr>
<tr>
<td>AC (n=89)</td>
<td>.33</td>
<td>.57</td>
<td>.21</td>
<td>.73</td>
<td>.17</td>
</tr>
<tr>
<td>PL (n=24)</td>
<td>-.07</td>
<td>.48</td>
<td>-.67</td>
<td>.55</td>
<td>-1.17</td>
</tr>
<tr>
<td>PWL (n=29)</td>
<td>-1.89</td>
<td>.53</td>
<td>-1.24</td>
<td>.63</td>
<td>-.84</td>
</tr>
</tbody>
</table>

F (Groups) 161.13** 60.66** 66.23** 17.10** 114.44**

Post-hoc comparisons
GC > AC, GC > AC, GC > AC, GC > AC, GC > AC, GC > AC, GC > AC, GC > AC, GC > AC, GC > AC, GC > AC, GC > AC, GC > AC

Note: GC = good comprehenders; AC = average comprehenders; PL = poor language; PWL = poor word-level skills; PA = phonological awareness; OP = orthographic processing; SE = semantic knowledge; LC = listening comprehension; RC = reading comprehension.

**p < .01

Latent Class Group Differences for ELLs

A multivariate analysis of variance (MANOVA) was run with the latent profile grouping as the between-subject variable to determine whether the latent groups differed from each other on any of the reading and language measures; raw scores were used. Preliminary assumption checking revealed that most of the data were normally distributed on each of the reading variables for each of the latent profile groups, as assessed by Shapiro-Wilk test (p > .05).

However, PA for the good comprehenders and poor language groups, as well as semantic processing for the poor word-level skills group violated normality. The histograms of the variables that violated the Shapiro-Wilk test were visually inspected, and no severe deviations from normality were observed. Given that MANOVA is less sensitive to violations of normality (Warne et al., 2014), no data transformations were made. Univariate outliers were found for PA in the poor word-level skills group and poor language group, orthographic processing for the poor language group, and semantic knowledge for the good comprehenders group. The mean for
each of the groups was compared with the 5% trimmed mean to determine if the outliers skewed the data. Given that none of the 5% trimmed means changed by more than 0.09 points, the outliers were not removed for the remaining analysis. There were no multivariate outliers, as assessed by Mahalanobis distance ($p > .001$); there were linear relationships, as assessed by scatterplots; and no multicollinearity, as assessed by Pearson correlations. The assumption of homogeneity of variance-covariance matrices was violated, as assessed by Box’s M test, ($p < .001$), so Pillai’s Trace was used instead of Wilk’s Lambda to test statistical significance, given that it is less sensitive to violations of assumptions (Olson, 1976).

The differences between the latent class groups on the combined dependent variables was statistically significant, $F(15, 477) = 29.614, p < .005$; Pillai’s Trace = 1.447; partial $\eta^2 = .482$. Follow-up univariate ANOVAs showed significant (latent class) group differences on PA ($F(3, 161) = 161.134, p < .0005$; partial $\eta^2 = .750$), orthographic processing ($F(3, 161) = 60.656, p < .0005$; partial $\eta^2 = .531$), semantic knowledge ($F(3, 161) = 66.247, p < .0005$; partial $\eta^2 = .552$), listening comprehension ($F(3, 161) = 17.100, p < .0005$; partial $\eta^2 = .242$), and reading comprehension ($F(3, 161) = 114.444, p < .0005$; partial $\eta^2 = .681$). Tukey post-hoc tests showed that on PA, orthographic processing and reading comprehension, all four groups were significantly different from each other. As for semantic knowledge, all groups were significantly different from each other, but the poor word-level skills group and the average comprehenders group did not differ from each other. On the listening comprehension task the good comprehenders group outperformed the other three groups, but the poor language, poor word-level skills and average comprehenders groups did not differ significantly from each other. The univariate ANOVAs and post-hoc comparisons for ELLs are summarized in Table 4.
EL1 Latent Profile Analysis Clusters

Analysis of the fit statistics showed that a three-class solution provided the best model for the EL1s (see table 3). Class 1 comprised 38.3% \((n = 36)\) of the total EL1 sample. This group was characterized by high scores on the four component skills as well as reading comprehension; this latent class was labeled the *good comprehenders* group (Figure 2a). Class 2 \((n = 30)\) comprised 31.9% of the total EL1 sample. This group consisted of students with strong PA and orthographic processing, and relatively low semantic knowledge and listening comprehension; this latent group was named the *poor language* group (Figure 2b). Class 3 \((n = 28)\) comprised the remaining 29.8% of the total EL1 sample. This latent group had difficulties across all four component skills as well as reading comprehension, with particular struggles in PA and orthographic processing; they were labeled the *poor word-level skills* group (Figure 2c). Mean standardized scores and standard deviations for each group are summarized in Table 5.
Figures 2 a-c. Mean Z-scores and 95% Confidence Intervals for the EL1 Latent Profile Groups - Good Comprehenders (GC), Poor Language (PL), and Poor Word-level Skills (PWL).

Note: PA = phonological awareness; OP = orthographic processing; SE = semantic knowledge; LC = listening comprehension; RC = reading comprehension.
Table 5. Descriptive Statistics (Z-scores) and MANOVA Summary Statistics for the EL1 Groups Defined by the Latent Profile Analysis.

<table>
<thead>
<tr>
<th></th>
<th>PA</th>
<th>OP</th>
<th>SE</th>
<th>LC</th>
<th>RC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>GC (n = 36)</td>
<td>.66</td>
<td>.37</td>
<td>.47</td>
<td>.83</td>
<td>.63</td>
</tr>
<tr>
<td>PL (n = 30)</td>
<td>.51</td>
<td>.38</td>
<td>.09</td>
<td>.92</td>
<td>-.43</td>
</tr>
<tr>
<td>PWL (n = 28)</td>
<td>-1.40</td>
<td>.45</td>
<td>-.70</td>
<td>.91</td>
<td>-.41</td>
</tr>
<tr>
<td>F (Groups)</td>
<td>220.91**</td>
<td>13.79**</td>
<td>14.42**</td>
<td>20.66**</td>
<td>74.28**</td>
</tr>
<tr>
<td>Post-hoc comparisons</td>
<td>GC &gt; PWL, GC &gt; PWL, GC &gt; PWL, GC &gt; PWL</td>
<td>GC &gt; PWL, GC &gt; PWL, GC &gt; PWL, GC &gt; PWL</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: GC = good comprehenders; PL = poor language; PWL = poor word-level skills; PA = phonological awareness; OP = orthographic processing; SE = semantic knowledge; LC = listening comprehension; RC = reading comprehension.

**p < .01

Latent Class Group Differences for EL1s

A MANOVA was run for the EL1s with the latent profile grouping as the between-subject variable. Preliminary assumption checking revealed that most of the data were normally distributed on each of the reading variables for each of the latent profile groups, as assessed by Shapiro-Wilk test (p > .05). However, PA for the good comprehenders group and poor language group, as well as listening comprehension for the good comprehenders group, violated normality. Visual inspection of the histograms revealed no severe deviations from normality, so no transformations were performed. Two univariate outliers were found for PA in the poor word-level skills group, as well as a single outlier for both semantic knowledge and listening comprehension in the good comprehenders group. None of the group means were significantly skewed by the outliers, so they were kept in the remaining analysis. There were no multivariate outliers, as assessed by Mahalanobis distance (p > .001); there were approximately linear relationships, as assessed by scatterplots; no multicollinearity, as assessed by Pearson
correlations; and there was homogeneity of variance-covariance matrices, as assessed by Box’s M test, (p = .119).

A MANOVA for EL1s using the latent profile groups as the independent variable was overall statistically significant, $F(10, 156) = 50.052, p < .0005$; Wilks’ $\Lambda = .056$; partial $\eta^2 = .762$. Follow-up univariate ANOVAs showed that there were significant (latent profile) group differences on PA ($F(2, 82) = 220.910, p < .0005$; partial $\eta^2 = .252$), orthographic processing ($F(2, 82) = 13.788, p < .0005$; partial $\eta^2 = .531$), semantic knowledge ($F(2, 82) = 14.423, p < .0005$; partial $\eta^2 = .260$), listening comprehension ($F(2, 82) = 20.661, p < .0005$; partial $\eta^2 = .335$), and reading comprehension ($F(2, 82) = 74.276, p < .0005$; partial $\eta^2 = .644$). Tukey post-hoc tests showed that on PA, most (latent profile) groups were significantly different from each other, however, as might be expected, the good comprehenders group and poor language group did not differ from each other. Likewise, on orthographic processing, Tukey post-hoc tests showed that most groups were significantly different from each other, though as might be expected, the good comprehenders and poor language groups did not differ from each other on this measure either. As for semantic knowledge, most groups were different from each other. However, the poor word-level skills group and poor language group did not differ from each other on the measure of semantic knowledge. On listening comprehension, most latent class group means were different from each other, but the poor word-level skills group and poor language group again were not different from each other. Finally, on reading comprehension most groups were different from each other, but the poor word-level skills group and poor language latent class group did not differ from each other. The univariate ANOVAs and post-hoc comparisons for EL1s are summarized in Table 5.
Whole Sample Latent Profile Analysis

In addition to the latent profile analysis conducted separately for each language group, a latent profile analysis was also conducted with both language groups together. This was done for two main reasons: first, to see whether similar latent groups were identified whether the language groups were analyzed together or separately, and second, to see if the proportion of ELLs and EL1s identified as good or poor readers changed when the sample was analyzed together versus separately.

Analysis of the fit statistics showed that a three-class solution provided the best model for the whole sample together (see table 6). Class 1 comprised 31.6% \((n = 86)\) of the total sample. This group was characterized by high scores on the four component skills as well as reading comprehension, and was labeled the *good comprehenders* group (Figure 3a). Class 2 \((n = 65)\) comprised 23.9% of the total sample. This group consisted of students with adequate PA and orthographic processing, and relatively low semantic knowledge and listening comprehension. This group also had below average reading comprehension. This latent group was named the *poor language* group (Figure 3b). Class 3 \((n = 121)\) comprised the remaining 44.5% of the total sample. This latent group had difficulties across all four component skills as well as reading comprehension, with particular struggles in PA and orthographic processing; they were labeled the *poor word-level skills* group (Figure 3c).
Table 6. Latent Class Analysis - Fit Statistics for Two, Three, and Four-class Models for the ELL and EL1 Language Groups Together.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Two-class</th>
<th>Three-class</th>
<th>Four-class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>AIC</td>
<td>BIC</td>
<td>SABIC</td>
</tr>
<tr>
<td>Two-class</td>
<td>8007</td>
<td>8064</td>
<td>8014</td>
</tr>
<tr>
<td>Three-class</td>
<td>7850</td>
<td>7929</td>
<td>7859</td>
</tr>
<tr>
<td>Four-class</td>
<td>7798</td>
<td>7898</td>
<td>7810</td>
</tr>
</tbody>
</table>

Note: AIC = Akaike Information Critrion; BIC = Bayesian Information Criterion; SABIC = Sample size Adjusted BIC; LMR = Lo-Mendell-Rubin likelihood ratio test.
Figures 3 a-c. Mean Z-scores and 95% Confidence Intervals for the Whole Sample Latent Profile Groups - Good Comprehenders (GC), Poor Language (PL), and Poor Word-level Skills (PWL).

Note: PA = phonological awareness; OP = orthographic processing; SE = semantic knowledge; LC = listening comprehension; RC = reading comprehension.
Figures 4a and 4b summarize the number of ELLs and EL1s classified into each latent profile group when the analysis was conducted with the ELLs and EL1s together versus separately. No difference in the classification of ELLs and EL1s in the good comprehenders group emerged whether the sample was analyzed together or separately ($\chi^2 (1) = 1.37, p = .24$). Similarly, no difference emerged for the classification of ELLs into the poor word-level skills profile whether their ELL data was analyzed separately or jointly with the EL1s ($\chi^2 (1) = .45, p = .50$). However, a significant difference emerged for the poor language group ($\chi^2 (1) = 14.74, p < .001$). Specifically, when the latent profile analysis was run with both language groups together, a significantly greater proportion of ELLs (50.6%) were identified as having poor language than when compared with only other ELLs (13.5%). In other words, ELLs were significantly more likely to be classified as having a poor language profile when they were compared with EL1s in the same analysis. As for EL1s, group classification remained almost unchanged regardless of whether they were classified with the whole sample or only with other EL1s.

![Number of ELLs in Each Latent Group](image1)

**Figure 4a.** The number of ELLs in each latent profile group when the ELL sample was analyzed separately or jointly with the EL1 sample. * indicates a significant chi square.

![Number of EL1s in Each Latent Group](image2)

**Figure 4b.** The number of EL1s in each latent profile group when the EL1 sample was analyzed separately or jointly with the ELL sample.
Discussion

Differences Between the ELL and EL1 Groups on Reading and Language Measures

Comparison of the overall group means on each of the measures administered revealed that by grade 5, no differences existed between the ELL and EL1 groups on word-level skills (i.e., phonological awareness and orthographic processing). This is consistent with findings from previous research that suggests that ELLs who have attended school in English since grade one perform similarly to EL1s on measures of PA, spelling, and word reading skills (Geva, Yaghoub-Zadeh, & Schuster, 2000; Wade-Woolley & Siegel, 1997; Lesaux & Siegel, 2003). However, differences between the groups were still present on vocabulary and reading comprehension, and a similar trend concerning listening comprehension (the test of group differences was close to significance, indicating that a difference may exist). These ELL-EL1 comparisons are consistent with previous research concerning vocabulary, listening comprehension and reading comprehension comparisons among ELLs and EL1s (Au-Yeung et al., 2015; Hutchinson, Whitely, Smith, & Connors, 2003; Farnia & Geva, 2011; Farnia & Geva, 2013; Kieffer, 2012; Proctor, Carlo, August, & Snow, 2005). Overall these results are consistent with previous research showing that by grade 5, ELLs have caught up to EL1s on essential skills associated with literacy acquisition such as PA and orthographic skills.

The fact that the groups did not differ from each other on non-verbal ability supports the notion that essentially the two samples are similar in terms of their overall ability. Yet, the ELLs do continue to lag behind their EL1 peers on vocabulary and reading comprehension despite more than four years of schooling in English, and there are indications that they may also be behind their EL1 peers on listening comprehension. Having established overall ELL-EL1 group differences, the discussion shifts now to the key issue addressed in this study, namely the reading and language profiles that emerged for the ELL and EL1 groups.
Latent Profiles and Support for the Lexical Quality Hypothesis

The latent profile analyses yielded different groupings in the ELL and EL1 groups. In the case of the EL1s it yielded a three-class solution, comprising a good comprehenders group, a poor language group, and a poor word-level skills group. For ELLs it yielded a four-class solution consisting of the same three groups, as well as an additional average comprehenders group. The good comprehenders group, comprising about a fifth (20.2%) of the ELL sample and more than a third (38.3%) of the EL1 sample, stuck out as the group who, in comparison with their respective language group, had above average reading comprehension skills, as well as above average skills pertaining to word reading (PA, orthographic processing) and to language skills (vocabulary and listening comprehension skills). Similarly, the average comprehenders profile of the ELLs was characterized by average scores on the component skills commensurate with their average reading comprehension. Half of the children in the ELL sample fit this profile.

In line with Perfetti and Hart’s (2002) lexical quality hypothesis, each of the other two groups for both ELLs and EL1s had difficulties with reading comprehension, and struggled with associated reading and language skills. The poor word-level skills group (16.3% for ELLs and 29.8% for EL1s) struggled primarily with skills associated with dyslexia, performing significantly worse than the other groups on both PA and orthographic skills. The poor language group (13.5% for ELLs and 31.9% for EL1s) struggled primarily with skills associated with language impairment, having adequate word-level skills, but poor language comprehension, poor lexical skills, and poor reading comprehension. While there was no significant difference between the poor word-level skills group and poor language group on vocabulary or listening comprehension for either ELLs or EL1s, the overall trend showed the poor language group had lower mean scores on both measures. Given the small size of each of the groups, there may not have been enough power to detect a group difference. The general trend in the means suggests
that the poor language group struggled more significantly with the two components of language comprehension, and had commensurate difficulties with reading comprehension.

On the whole, over two thirds of grade 5 ELLs and more than a third of EL1s achieved at least average scores on reading comprehension as well as on the associated component reading and language skills essential for reading comprehension, while the remaining students struggle with reading comprehension and with the associated component skills. The struggling readers, making up about a third of the ELL sample consists of ELLs whose difficulties are not associated only with their ELL status, and reflect instead deficits in specific skills that are essential for reading comprehension. For EL1s, this group is similarly comprised of students who are performing below that of their typically-developing EL1 peers. For both ELLs and EL1s two such profiles emerged, one of children who struggled primarily with word-level skills and another who struggled primarily with language comprehension.

These results suggest that the lexical quality hypothesis can help to explain the reading comprehension difficulties of ELLs and EL1s. To recapitulate, the lexical quality hypothesis proposes that those with good reading comprehension have a large repertoire of high-quality word representations that are well specified with phonological, orthographic and semantic components (Perfetti & Hart, 2002). Accordingly one would expect that readers with strong reading comprehension should also perform well on measures of these three component skills. Conversely, readers who are poor comprehenders should struggle with at least one of the component variables, and those with difficulty in multiple component areas should have the greatest difficulty with comprehension. In general, results of the latent profile analysis provided support for the lexical quality hypothesis in both the ELL and EL1 samples, as the group with the highest reading comprehension scores had the highest scores on each of the component variables, and those who struggled with reading comprehension had difficulties with phonological and
orthographic skills (poor word-level skills group), or with semantic knowledge and listening comprehension (poor language group). Good comprehenders have good, well-specified word representations, evidenced by strong performance on measures of the three component variables, while poor comprehenders struggle on one or more of the components.

At the same time, one slight deviation from this theory was that according to the lexical quality hypothesis, the poor-word level skills group should have had the lowest reading comprehension score, given that they struggled on each of the component skills, while the poor language group performed significantly better on both PA and orthographic knowledge. This was not the case for ELLs, as the poor word-level skills group actually had higher reading comprehension scores than the poor language group. It appears that at least in the case of ELL students in middle school, the lowest profile is that of students who struggle with poor listening comprehension and not with decoding. They obtain lower reading comprehension scores than ELLs whose difficulties concern primarily word-level skills. Clearly, other factors such as working memory, executive function, and reading fluency may have to be included in the latent profile analysis to be able to distinguish further these groups. Nevertheless, this study provides evidence that very poor word-level reading skills and poor language skills are associated with poor reading comprehension, as the lexical quality hypothesis would predict.

In general, the latent profiles support the lexical quality hypothesis to the extent that one can identify distinct profiles of ELLs and EL1s whose reading comprehension skills vary as a function of how well they can draw on their phonological, orthographic, and semantic skills. At the same time, the results demonstrate that a subset of the sample for both ELLs and EL1s have difficulties with language comprehension skills beyond that of just semantic knowledge; this issue is discussed in more detail in the next section.
Thinking of Vocabulary and Language Comprehension

The second research question addressed in this study concerned the construct of language comprehension and whether considering not only lexical knowledge but also listening comprehension provides a more nuanced understanding of how good comprehenders and poor comprehenders differ from each other. According to the SVR, difficulties with language comprehension and/or word reading skills should lead to a breakdown in reading comprehension (Hoover & Gough, 1990). Results pertaining to the ELL and EL1 groups provide support for the SVR through the important role of language comprehension in the reading comprehension of monolingual as well as ELL students. In particular, the latent profile analysis showed that the ELLs and EL1s with a poor language profile struggled not only with semantic knowledge but also with listening comprehension, and also showed significant difficulties with reading comprehension.

Some ELLs and EL1s experience difficulties with orthographic skills, remembering how words are spelled, and with phonological skills, being able to identify and manipulate the sound units of words. Difficulties with these word-reading skills lead to a breakdown in reading comprehension. At the same time some ELL and EL1 school children have serious reading comprehension difficulties that are associated with more limited vocabulary knowledge, and with difficulties in the ability to listen to and comprehend discourse more broadly. In the listening comprehension task used in this study, children were required to listen to a series of short stories and answer both factual and inferential questions about them. Accomplishing this involves a multitude of skills, including attending carefully to the stories, holding information in working memory, knowing the meaning of the individual words, understanding the use of grammar and morpho-syntax, integrating prior knowledge, and being able to make inferences and predictions. Listening comprehension tasks tap into a variety of skills beyond familiarity with vocabulary.
Even in comparison with their ELL peers, children with a poor language profile struggled with both vocabulary and listening comprehension, suggesting that this group has an overall deficit in oral language comprehension that goes beyond poor vocabulary and that cannot be attributed only to their ELL status. Vocabulary knowledge is a pivotal aspect of language comprehension, and the two correlate positively and significantly. Vocabulary is often used as an index of language comprehension – this is practical in research, assessment, and teaching contexts. However, researchers, teachers, and practitioners should be mindful of the broad and complex nature of language comprehension and the fact that an examination of vocabulary skills does not provide a full picture of how language skills contribute to and in turn benefit from reading comprehension. Researchers, teachers and clinicians should think of language comprehension broadly as an important and necessary component of reading comprehension that should not be overlooked. Stated differently, both the lexical quality hypothesis and the simple view of reading are needed to fully explain reading comprehension.

**Do Identical Latent Profile Groups Exist for the ELL and EL1 Groups?**

The third research question concerned the extent to which there would be differences between ELLs and EL1s with regards to the groups that emerged. Overall, similar latent profile groups emerged for both ELLs and EL1s, with similar patterns among the four component variables and reading comprehension. Nevertheless, a couple key differences emerged between the language groups, and these are discussed below.

Firstly, the latent profile analysis indicated the best fitting model to be four groups for ELLs and three groups for EL1s. Of significance, the proportion of ELLs and EL1s who were classified in each latent profile group was not identical. This can be at least partially attributed to the existence of a fourth group for ELLs. However, it is important to note that the additional group for ELLs, the average comprehenders, consisted of half of the ELL sample, and when
combined with the good comprehenders group comprised approximately 70% of the ELL sample. This shows that when compared with other ELLs, a high proportion are identified as having at least average reading skills.

Secondly, the ELL and EL1 poor language group profiles were not identical. The EL1 poor language group showed a profile more typical of language impairment, consisting of average or above average phonological and orthographic skills (when compared to other EL1s), with specific deficits in lexical knowledge and listening comprehension. In contrast, the ELL poor language group also had below average orthographic processing (when compared to other ELLs). Given that the ELL sample was significantly larger than the EL1 sample, it is possible that a group with both decoding and language difficulties emerged for the ELLs, but not for the EL1s.

Regardless of language status, students with strong phonological awareness, orthographic processing, and language comprehension (including both vocabulary and listening comprehension) have better reading comprehension than those students who struggle with one or more of these skills. In addition, one can detect among ELLs and EL1s students whose profiles are consistent with dyslexia and language impairment.

**Comparing Analysis of the Whole Sample vs. by Language Status**

An additional measurement question of interest was whether similar latent groups would be identified when the sample was analyzed as a whole, with ELLs and EL1s together, versus when they were analyzed separately, as well as whether the proportion of ELLs and EL1s identified would be the same. The latent profile analysis conducted with the joint ELL and EL1 sample yielded three clusters: a good comprehenders group, a poor word-level skills group, and a poor language group. Most importantly, the profile of average comprehenders disappeared and there was a concomitant shift in the proportion of ELLs associated with each profile, but this
shift did not occur in the EL1 sample. Notably, half of the ELL sample were considered as belonging to the poor language comprehension profile when their performance was based on absolute criteria that did not consider their ELL status. By comparison, only 13.5% appeared to have more serious difficulties with language comprehension that extend over and above what may be a more typical ELL profile when compared with other ELLs. These results suggest that when the performance of ELLs and EL1s are analyzed jointly and no consideration is given to the fact that by definition ELLs have lower language skills than their monolingual counterparts, they are more likely to be considered as underperforming and perhaps as having language and reading comprehension difficulties. By definition ELLs have poorer language skills than their monolingual peers. However, when the focus is on different profiles, the ELLs whose language skills reflect their L2 status are more likely to be lumped together with those EL1s who have more serious language related difficulties when the data are analyzed jointly. These results demonstrate the danger of attributing a very poor language skills profile to ELLs whose profile can be more accurately described as average.

**Limitations and Future Research**

While this study allows for some important conclusions to be drawn regarding ELL and EL1 reading comprehension, there are a few limitations to consider. Firstly, the sample size for the EL1 group was significantly lower than the ELL group, and this presumably reduced the statistical power of the analysis and also made it more difficult to make direct comparisons between the language groups. In the future, it would be beneficial to have more equal sample sizes. Secondly, as with many studies, there are some limitations with regard to the measures used to assess each of the constructs. For example, the semantic knowledge measure used in the current study focused on vocabulary depth and required participants to provide word definitions, whereas Perfetti & Hart (2002) used a multiple-choice procedure. It is an empirical question
whether similar findings would be found when additional measures of the same constructs are used. Another issue is the use of different measures to assess the same constructs. Ideally, one would want to use the same measure across different studies to allow for a more precise comparison of emerging patterns. Lastly, while the data for this study were taken from a longitudinal study, this paper has focused only on performance in grade 5. Future research should examine these patterns longitudinally.

**Research Implications**

Even by grade 5, after more than four full years of immersion in English, ELLs still lagged behind EL1s on their vocabulary and reading comprehension. This suggests that it is important for educators and clinicians to avoid comparing ELLs to their monolingual peers on these skills, as these abilities are still developing. It also was evident, however, that by grade five the ELLs had caught up to the EL1s on the word-level skills of phonological and orthographic processing. This suggests that if ELLs are still displaying difficulties in these skills by grade five, it may be attributable to a reading difficulty rather than their language proficiency.

Another important implication is with regard to the theories of reading comprehension that were tested in this study. For both ELLs and EL1s, it is important to consider phonological, orthographic and semantic skills to obtain an accurate picture of reading comprehension. Importantly, considering both vocabulary and listening comprehension as part of the construct of language comprehension helps to discriminate those with poor language from those with word-level skill difficulties for both ELLs and EL1s. Intervention programs aimed at improving reading comprehension should first identify the underlying deficit, whether it be with word-level skills or with language comprehension, and programs targeting language comprehension should focus on language skills beyond that of just vocabulary knowledge.
Interesting findings emerged when the latent profile groups based on the analysis with the whole sample were compared to those in the separate language groups. Overall, similar profiles of readers were found, pointing to the fact that deficits on similar component skills underlie reading comprehension difficulties for both ELLs and EL1s, and that the variables contributing to reading difficulties do not vary greatly between the groups. Decoding skills are not sensitive to language proficiency and therefore the profiles of ELLs and EL1s with dyslexia are stable whether the data are analyzed separately or jointly. However, this is not the case for those with serious language comprehension difficulties; there is a danger in blurring the subtle but real distinction between average ELLs, who by definition have poorer language skills than their monolingual peers, and ELLs whose profiles are associated with persistent language and reading comprehension difficulties. Some ELLs have poor language skills associated with their L2 status. Others are struggling with a “double whammy” as they are ELL and in addition they struggle with underlying difficulties in acquiring adequate language skills in their L2.
References


processing skills on five different reading tasks. *Reading Research Quarterly*, 335-345.


Lesaux, N., & Geva, E. (2006). Synthesis: Development of Literacy in Language-


Savage, R., & Wolforth, J. (2007). An Additive Simple View of Reading Describes the Performance of Good and Poor Readersin Higher Education. *Exceptionality*


