Multisyllabic Word Reading in Grades 4 and 5: Accuracy, Errors, and Associated Child-Level Skills

By

Lindsay Heggie

A dissertation submitted to the Graduate Program in Education in conformity with the requirements for the Degree of Doctor of Philosophy

Queen’s University

Kingston, Ontario, Canada

April, 2017

Copyright © Lindsay Heggie, 2017
Abstract

The vast majority of English words are multisyllabic (Balota et al., 2007; Baayen, Piepenbrock, & Gulikers, 1995), constituting an increasingly large proportion of the words students encounter in print beginning as early as Grade 3 (Archer, Gleason, & Vachon, 2003; Nagy & Anderson, 1984; Zeno, Ivens, Millard, & Duvvuri, 1995). Students who continue to struggle with word reading through elementary and into secondary school often have a particular difficulty with longer words (Archer, Gleason, & Vachon, 2003; Moats, 1998). This multi-manuscript dissertation reports on two studies of multisyllabic word reading in typically achieving, English-speaking students in grades 4 and 5. The first study examined children’s multisyllabic word reading ability and the relative contributions of five variables of interest: phonological awareness, suprasegmental phonological awareness, morphological awareness, vocabulary, and naming speed. Regression analyses showed that the contribution of phonological awareness (at both the segmental and suprasegmental levels) was fully mediated by that of morphological awareness and that naming speed and morphological awareness predicted children’s multisyllabic word reading accuracy at this stage of reading development. The second study used error analysis to describe the types of errors students made when reading multisyllabic words. Error productions were examined according to reading ability, based on a six-point coding scheme developed to assess decoding and lexical stress errors; good readers made proportionally fewer errors, but were more likely to make errors involving shifts in lexical stress. Error productions were also described in terms of lexicalization (reading the item as a real but erroneous word) and stress regularization (stress shift to the first syllable, in accordance with the majority of English trisyllabic words). Both morphological awareness and vocabulary were positively
related to the incidence of lexicalized errors, while both reading ability and suprasegmental phonological awareness were positively related to participants’ tendency to regularize the primary stress of multisyllabic words. Results of both studies are discussed in terms of contributions to theory, instruction, and future research in multisyllabic word reading.
Co-Authorship

Chapters 1 and 4 were written by Lindsay Heggie and use the first person to reflect that authorship. Chapters 2 and 3 were written somewhat more in consultation with Dr. Lesly Wade-Woolley (University of South Carolina), Dr. John Kirby (Queen’s University), and Dr. Linda Jarmulowicz (University of Memphis) and as such use the first person plural to reflect this co-authorship. These two manuscripts will be submitted for publication listing all four authors, with Lindsay Heggie as first author on both papers.
Acknowledgements

First and foremost, thank you to my supervisor, Dr. Lesly Wade-Woolley. I am lucky to have had a supervisor who is both a mentor and a friend. Lesly provided countless opportunities for me to learn about research and reading development, to extend and apply my learning, and to participate in a range of professional and academic activities. If the completion of the Ph.D. is considered the beginning of one’s career, I am well positioned to “start as I mean to go on.”

Thank you to the members of my committee, Dr. John Kirby and Dr. Linda Jarmulowicz. John’s feedback and comments on drafts were invaluable, not only for their thoughtfulness and detail but for their swiftness; thank you for helping me to improve the clarity of my thinking and writing. Linda, thank you for sticking with me despite a difficult year and for musing, discussing, and working through error analysis with me. Thank you as well to Drs. Elizabeth Lee, Elizabeth Kelley, and Robert Savage, for their interest, attention, and comments on my dissertation.

The Language & Literacy Research Lab was a constant source of learning, fun, and mentorship, for myself and for all of Lesly’s graduate and honours students. Between data collection, lab meetings, stimulus development, journal club, conferences, and our many extra-curricular events, the LLRLab community added to and enriched the supportive foundation that Lesly set down. I’m so glad it was part of my time at Queen’s. Thank you in particular to my fellow LLRLab Ph.D. student Jessica S. Chan, one of the sunniest and most hardworking people I’ve had the pleasure to know. You’re next!

There are a number of faculty and staff members at Queen’s who have contributed to making the Ph.D. experience one of valuable growth, scholarship, and opportunity. I
couldn’t possibly name all those who have had a positive influence on my time here – truly, it would involve mentioning nearly everyone. But I would be remiss if I did not specifically mention the indispensible Marlene Sayers. When I began graduate work at the Faculty of Education, I was told that if you need help, you should always ask Marlene because she knows everything. It’s true.

Thank you to my dad, who has always supported my endeavors, big and small, academic and otherwise, with the greatest enthusiasm. Special thanks to my mom and mother-in-law, without whom the writing process would have been so much more challenging; their tag team of childcare support was instrumental in the final year of my degree, and I am extremely grateful for their generosity and selflessness.

To my daughters: I love you both so much and can’t wait to see who and what you become. (But don’t grow up too fast, please.) Finally, Vince: your patience and support through this degree were constant, and constantly felt. Thank you for everything you do. I love you.
# TABLE OF CONTENTS

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

Co-Authorship ........................................................................................................................................... iv

Acknowledgements .................................................................................................................................. v

List of Tables .......................................................................................................................................... xi

List of Figures ......................................................................................................................................... xii

CHAPTER ONE: Introduction .................................................................................................................... 1

The Importance of Reading ..................................................................................................................... 1

  Brief History of Reading Development Research .................................................................................. 3

Word Reading .......................................................................................................................................... 4

  Models of Word Reading Development ................................................................................................. 5

Multisyllabic Word Reading ................................................................................................................... 7

  Multisyllabic Word Recognition ........................................................................................................... 8

  Modeling Multisyllabic Word Reading .................................................................................................. 9

Constructs Associated with Improvement in Word Reading Ability ................................................. 10

  Phonological Awareness ......................................................................................................................... 11

  Suprasegmental Phonological Awareness ............................................................................................... 13

  Morphological Awareness ....................................................................................................................... 14

  Additional Predictors ............................................................................................................................... 16

  Error Analysis of Word Reading ............................................................................................................ 17

Chapter One: A Brief Review ................................................................................................................ 18

Dissertation Advance Organizer .......................................................................................................... 18
Error Analysis in Action

Considerations When Applying Error Analysis

Lexical Stress in Trisyllabic Words

The Current Study

Method

Morphological Awareness

Phonological Awareness

Suprasegmental Phonological Awareness

Vocabulary

Naming Speed

Reading Measures

Results

Error Productions in General and by Reading Ability

Lexicalized Error Productions

Regularizing Trisyllabic Stress

Discussion

Error Analysis of Multisyllabic Word Reading

Coding Scheme for MSWR

Limitations and Future Directions

Conclusion

CHAPTER FOUR: General Discussion

Multisyllabic Word Reading: Implications for Theory
The Spreading Activation of an Interconnected Network .......................................................... 101

Multisyllabic Word Reading: Implications for Instruction .................................................... 105

Multisyllabic Word Reading: Implications for Future Research ........................................... 107

Overall Conclusion .................................................................................................................. 109

References ............................................................................................................................... 111

Appendix A: Correlations Between Measures ....................................................................... 152

Appendix B: MSWR Items and Attributes ............................................................................ 153

Appendix C: Approval Letter from General Research Ethics Board .................................. 155
List of Tables

Table 2.1  Descriptive Statistics.................................................................30
Table 2.2  Pearson Correlations Between Measures........................................31
Table 2.3  Multiple Regression Predicting Multisyllabic Word Reading (n=93)........32
Table 2.4  Multiple Regression Predicting Multisyllabic Word Reading (n=93) ..........36
Table 3.1  Scoring Scheme for Multisyllabic Word Reading Task...............................66
Table 3.2  Descriptive Statistics..........................................................................67
Table 3.3  Distribution of Readings Across Codes for MSWR (54 items)..................68
Table 3.4  Proportion of Errors Made by Reading Quartile.....................................69
Table 3.5  Mean Differences Between Reading Quartiles on Error Proportions ..........72
Table 3.6  Correlations Between Error Types (n=93)..............................................73
Table 3.7  Correlations Between Child-Level Factors and Number of Errors of Each Type 74
Table 3.8  Participants With the Most Stress Only Errors and Corresponding Items ..........75
List of Figures

Figure 2.1a. Indirect effect of PA on MSWR through MA .................................................34

Figure 2.1b. Indirect effect of SSPA on MSWR through MA ..............................................34

Figure 3.1. Distribution of error types by reading quartile.....................................................70

Figure 3.2. Mean proportion of error types by reading quartile ...........................................71

Figure 3.3. Number of error productions by participants in each quartile..............................77

Figure 3.4. Percentage lexicalized error productions by quartile ........................................77

Figure 3.5. Percentage of error productions by length (number of syllables).......................78

Figure 3.6. Error productions involving a shift in primary stress........................................80

Figure 3.7. Location of primary stress on three-syllable error productions involving stress
- shifts, by reading quartile ......................................................................................................81
CHAPTER ONE: Introduction

There are several reasons why multisyllabic words are hypothesized to be more difficult to read than monosyllabic words. In fact, they are more difficult just by virtue of being longer (Jared & Seidenberg, 1990; New, Ferrand, Pallier, & Brysbaert, 2006). Additionally, there are complexities that come with increased length: syllable boundaries (Perry, Ziegler, & Zorzi, 2010), linguistic stress and vowel reduction (Chomsky, 1970; Seva, Monaghan, & Arciuli, 2009), vowel pronunciation ambiguities (Venezky, 1999), less predictable grapheme-phoneme correspondences (Berninger, 1994; Venezky, 1999), and morphological complexity (Nagy, Anderson, Schommer, Scott, & Stallman, 1989; Carlisle & Stone, 2005; Nagy, Berninger, & Abbott, 2006; Nagy, Berninger, Abbott, Vaughn, & Vermeulen, 2003). Herein, I focus on multisyllabic words and some of the child-level factors hypothesized to be important for their decoding: prosodic awareness (the awareness of and ability to manipulate lexical stress), phonological awareness, morphological awareness, vocabulary, and naming speed.

The Importance of Reading

Literacy is now so deeply rooted in our culture “that we often take for granted the complex cognitive abilities that are required to read effortlessly” (Norton & Wolf, 2012, p. 428). Indeed, for those of us with a high level of literacy – especially if our fluency with the written word came without a protracted effort – it can be hard to recall a time when we did not read, let alone the steps we took to learn to read. It can be easy to overlook how fundamental, foundational, and essential a component of our education learning to read is, especially once it has been incorporated as an effortless aspect of our adult lives. Over the last half-century (and beyond), an immense amount of research time, effort, and funding
has been directed toward reading development, in the name of children’s well-being and the promotion of the best possible literacy instruction (Snow, Burns, & Griffin, 1998). Society has evolved such that literacy is one of the primary tools with which we learn, communicate, and interact with the world around us. Low literacy has academic, social, and financial consequences for individuals and for society at large (e.g., The Conference Board of Canada, 2016; Howard & Gagnon, 2006; OECD, 2011; Snow, Burns, & Griffin, 1998).

The long-term consequences of illiteracy are socially and personally devastating (Geary, 2013). From the ability to understand essential everyday information (e.g., nutritional, financial, medical) to employment rate and quality to health and wellbeing, literacy impacts adults’ quality of life in myriad ways. Literacy is not a distal concern, relegated to developing countries; according to the Organization for Economic Co-Operation and Development’s (OECD) Survey of Adult Skills (PIAAC, 2015), 16.4% of adults in Canada have low literacy; those with low literacy frequently do not graduate from high school and have difficulty remaining employed in adulthood (Baer, Kutner, Sabatini, & White, 2009). This outcome is relevant for researchers of reading development because low literacy begins in childhood: children who struggle to master fundamental literacy skills (e.g., knowledge of letters and sounds) are more likely to have difficulty learning to read, to be classified as having a reading disability (Gallagher, Frith, & Snowling, 2000; O’Connor & Jenkins, 1999), and to fall behind their peers in other areas such as vocabulary, fluency, and comprehension (Stanovich, 1986; Torgesen, 2002). Low levels of literacy achievement among children and adolescents continue to be of concern (Deshler & Hock, 2007); by Grade 4, 34% of public school students are still reading below a basic level (Lee, Grigg, & Donahue, 2007).
Brief History of Reading Development Research

Research has resoundingly demonstrated the critical importance of oral language for the development of reading: proficiency in written language is built upon a foundation of oral language (Snow, 1991; Snow, Burns, & Griffin, 1998; Shanahan, MacArthur, Graham, & Fitzgerald, 2006). We are able to make connections to meaning in written language thanks to our ability to understand and engage with oral language. That said, the two processes are very different; inherent in the difference between “acquisition” and “learning” is the notion that acquiring language is a natural process while learning to read is an effortful, time-intensive process. All that is typically needed to acquire language in infancy is exposure: children are born wired for language (e.g., Chomsky, 1965; Pesetsky, 1999); in contrast, “print is an optional accessory that must be painstakingly bolted on” (Pinker, 1997, p. ix). Although decoding words begins as a slow, laborious process it finishes as a skill so ingrained, so automatized, that as adults we process text instantaneously, seemingly without effort.

The National Reading Panel’s influential meta-analysis (National Institute of Child Health and Human Development [NICHD], 2000) identified five pillars of beginning reading: phonological awareness, phonics, fluency, vocabulary, and comprehension. These five pillars provide the foundation for literacy development and their efficacy has been validated for young readers. Depending on the age and prior experience of the students, some components should be emphasized more than others, and additional components (e.g., morphological awareness) may also be included. Learning to read is a multifaceted skill that is gradually learned with instruction and practice; to build skilled reading, children must integrate increasingly strategic components of language comprehension (e.g.,
language structures, background knowledge) and with increasingly automatic word recognition skills (e.g., phonological awareness, decoding, sight word recognition) (Scarborough, 2001). Then, as they transition from learning to read to reading to learn, children can use their accumulated decoding abilities for the purpose of comprehension, devoting cognitive resources to the acquisition of new knowledge and information (Chall, 1983).

There are significant individual differences that contribute to the need for high quality, focused, explicit instruction in reading (Bradley & Bryant, 1983; Burns & Kidd, 2010; Dickinson, McCabe, Anastasopoulos, Peisner-Feinberg, & Poe, 2003; Scarborough, 2001); despite adequate instruction, intelligence, and effort, about 10% of children will continue to experience difficulty learning to read (Lyon, Shaywitz, & Shaywitz, 2003).

**Word Reading**

Although comprehension and learning are widely recognized as the ultimate purposes of reading, virtually all theories of reading development recognize that a fundamental component of this process is the ability to read individual words (Kirby & Savage, 2008). Reading individual words quickly and accurately in isolation (as well as in context) is the foundation of reading (Stanovich, 1980) and constitutes the focus of the two studies herein. If readers are unable to decode words fluently and efficiently, they will continue to expend the majority of their limited cognitive resources painstakingly decoding, leaving little to none for the demanding task of comprehension (LaBerge & Samuels, 1974; Samuels, 1994; Snowling & Hulme, 2005). Such slow, painstaking reading is unenjoyable, frustrating, and difficult, leading to less practice and less automaticity (e.g., Lyon, Shaywitz, & Shaywitz, 2003; Nagy & Anderson, 1984; Verhoeven & Schreuder, 2011). Indeed, the
inability to read words lies at the heart of most reading difficulties (Shaywitz & Shaywitz, 2008; Stanovich, 2003). How we learn to read words has been a major research focus over the last half-century and even, as Snowling and Hulme (2005) noted, constitutes one of the oldest areas of research in the whole of experimental psychology (e.g., Cattell, 1886). Over the decades, through the advent of cognitive theories, educational psychology, linguistics, and the reading "wars," we have accrued a great deal of information about how children learn to read words.

Models of Word Reading Development

Two models, one developmental and the other cognitive, are particularly relevant for my study of children’s multisyllabic word reading. First, Chall’s stages of reading development (1983, 1996) conceptualize learning to read as a process that changes as the reader matures, changes, and acquires knowledge. From prereading (Stage 0) to the most mature, skilled level of reading (Stage 5), children progress through learning to read (Stages 1 and 2) in the primary grades to reading to learn (Stages 3, 4, and 5). Beginning around Grade 4, children enter Stage 3: Confirmation, fluency, and ungluing from print. In order to begin to use reading as a tool for learning with the increasingly demanding, complex texts, children must first build fluency in decoding, accumulating sight words that are accessible from long term memory after multiple exposures (Ehri, 1994; Wolf & Katzir-Cohen, 2001). In addition to automatization and internalization of the information learned, Chall hypothesized that there is a transition in this stage to texts that presents more orthographically complex patterns, including more advanced multisyllabic words: during Stage 3, children develop their knowledge of higher level graphophonics, especially the morphological or orthographic patterns found in complex words (Fitzgerald & Shanahan,
A limitation of Chall’s stage theory is that, like other developmental theories of reading (e.g., Adams, 1990; Ehri, 1994), it underestimates the extensive morphological knowledge children have in oral language in (and even before) the primary grades (e.g., Brown, 1972; Clark, 1982). Cognitive models of reading take a more integrated approach to the myriad components involved in reading development, over simultaneously activated pathways.

There are several cognitive models of reading (e.g., Dual-Route (DRC) model, Coltheart, 2005; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Parallel Distributed Processing (PDP) model, Seidenberg & McClelland, 1989) that make an effort to provide insight into reading acquisition, normal skilled reading, and developmental/acquired deficits; the “triangle model,” first articulated by Seidenberg & McClelland (1989), employs connectionist architecture to examine word reading accuracy as learning to map among orthographic, phonological, and semantic representations and the simultaneous, spreading activation of phonological and semantic pathways during word reading. That is, it is a multi-componential view that incorporates the three major facets on which successful word reading development depends: phonology (sounds), orthography (writing), and semantics (meaning) are integrated, for entire words in parallel (Harm & Seidenberg, 2004; Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989) and perhaps also for morphemes (Kirby & Bowers, in press). While it does not deny the existence of sequential processes, the triangle model’s architecture emphasizes the parallel interactions among central types of lexical information (sound and meaning; Plaut, 2005). In their full implementation of the model using 6,103 monosyllabic words, Harm & Seidenberg (2004) demonstrated that the network’s connections change based on developmental differences
in word reading accuracy: it relied heavily on phonological mediation (orthography → phonology → semantics) in the early stages of reading acquisition but gradually shifted towards increased reliance on the direct mapping (orthography-semantics) as reading skill improved. In addition to Chall’s stages of reading development, the triangle model is another way to frame the development and application of the relevant linguistic (and metalinguistic) variables of interest, as well as their inter-relationships, in the studies discussed herein.

Multisyllabic Word Reading

Comparatively little research exists to describe which skills children use to read multisyllabic words in English (Roberts, Christo, & Shefelbine, 2011), despite the fact that they constitute the vast majority – over 90% – of English words (Baayen, Piepenbrock, & Gulikers, 1995). Studies of English word recognition, for example, tend to rely far more heavily on monosyllabic than multisyllabic words as stimulus items (Roberts, Christo, & Shefelbine, 2011). Although some work has been conducted on issues related to longer words such as syllabification (e.g., Treiman & Zukowski, 1990, 1996), researchers have noted that processes related to multisyllabic word reading (MSWR), such as demonstrations of how beginning readers identify syllables and morphemes or how lexical stress is assigned when reading multisyllabic words, is still insufficient (Verhoeven & Perfetti, 2003).

This imbalance has direct consequences for the classroom: students who continue to struggle with reading into secondary school may be able to decode monosyllabic words but have difficulty with multisyllabic words, for which they have neither a systematic approach nor the confidence to persevere (Archer, Gleason, & Vachon, 2003; Carlisle & Katz, 2006;
Moats, 1998). Further, if they are using a more laborious grapheme-by-grapheme decoding strategy, it will consume readers’ cognitive resources and impinge upon their working memory capacity. This is compounded by the fact that multisyllabic words become both increasingly frequent and relevant for readers over elementary and into secondary school. Studies estimate that the “orthographic avalanche” (Share, 1995; p. 153) begins around Grade 5, when students encounter approximately 10,000 new words in print per year, the majority of which are multisyllabic (Archer, Gleason, & Vachon, 2003; Nagy & Anderson, 1984). Other researchers have estimated that the yearly increase is even more dramatic, with as many as 20,000 new multisyllabic words in print per year, beginning in Grades 3 and 4 (Hiebert, Martin, & Menon, 2005; Kearns et al., 2016; Zeno, et al., 1995).

**Multisyllabic Word Recognition**

Much of the available research on how readers approach the complexities of multisyllabic words comes from the literature on word recognition. For example, just being longer makes words more challenging to read. Using the English Lexicon Project, an online database of 40,481 words (http://elexicon.wustl.edu; Balota et al., 2002), New and colleagues highlighted an effect of word length (defined by number of letters) on both lexical decision and reaction times in English for words 8 letters or longer (New, Ferrand, Pallier, & Brysbaert, 2006), echoing earlier work that showed a length effect specifically for low frequency words (e.g., Jared & Seidenberg, 1990). More recently, Yap & Balota (2009) examined the influence of major psycholinguistic variables on the recognition of monomorphemic, multisyllabic (2-6 syllable) English words. After an analysis of a large-scale database (6115 items), the authors concluded that the processing of monosyllabic and multisyllabic words may not be “radically different,” in terms of what models of visual word
recognition need to accommodate. Additional skills may be required to accurately decode longer words, however. Yap & Balota focused only on monomorphemic items (which constitute the minority of multisyllabic words). Combined with the fact that students with reading difficulties tend to struggle specifically with multisyllabic words (Archer, Gleason, & Vachon, 2003; Carlisle & Katz, 2006; Moats, 1998), there is reason to believe that Yap & Balota’s findings may not necessarily lend themselves directly to individual differences in multisyllabic word reading accuracy. In fact, Yap and Balota noted that their study was limited in terms of variables (e.g., missing both morphological and prosodic attributes of multisyllabic words) and that these psycholinguistic variables are not always related to reaction times.

**Modeling Multisyllabic Word Reading**

There has been some attempt to build comprehensive models of multisyllabic word reading with a combination of item-, word- and child-level variables (Kearns, 2015; Kearns et al., 2016). Kearns (2015) focused on polysyllabic, polymorphemic (PSPM) word reading accuracy in Grades 3 and 4, keeping a focus on morphological processing. His findings suggest that readers rely strongly on morphological units to read unknown words, including their knowledge of root words (both item-specific and in general). Kearns and colleagues (2016) modeled individual differences in multisyllabic word reading in a population of Grade 5 students oversampled for reading disability. They found that multiple sources influence students’ multisyllabic word reading accuracy, including item-specific vocabulary knowledge, word frequency, morphological and orthographic knowledge, and reading disability status. Kearns and colleagues have yet to include lexical stress as a variable in their work, despite having noted its relevance for PSPM words (Kearns, 2015).
Thus studies of multisyllabic word reading suggest that the factors that are involved in monosyllabic word reading are also important for multisyllabic word reading; however, there are additional factors that are critical, making MSWR even more complex. In the next section, I discuss the specific constructs that will be included as predictors of MSWR and the evidence for their importance.

**Constructs Associated with Improvement in Word Reading Ability**

As the elements of the triangle model suggest, learning to read is fundamentally metalinguistic (Mattingly, 1984; Nagy & Anderson, 1999). In general, the simultaneous application of phonology and semantics to orthography allows readers to decode and understand words as they are read; in fact, proficient reading requires the simultaneous coordination of several component skills (Hudson, Pullen, Lane, & Torgesen, 2009; McCardle, Scarborough, & Catts, 2001). Oral language skills are one of the most consistent predictors of reading ability (e.g., Dickinson, Golinkoff, & Hirsh-Pasek, 2010; Scarborough, 1998) but it is their application to written language that supports children’s reading development. For example, the concept of the alphabetic principle – that graphemes stand for sounds – is rooted in an understanding that spoken words are made up of phonemes (e.g., “cat” is /k/-/a/-/t/). Tasks such as phoneme awareness, which emphasize the skills that rely on our implicit knowledge of oral language, are particularly predictive of reading (e.g., Carlisle, 2000; Deacon & Kirby, 2004; Nunes, Bryant, & Bindman, 1997; Singson, Mahony, & Mann, 2000; Wagner & Torgesen, 1987).

Herein, I do not intend not include all the possible predictors that may be important for word reading. Instead, I focus on a subset of metalinguistic skills (phonological awareness (PA), suprasegmental phonological awareness (SSPA), and morphological
awareness (MA)) and two additional predictors (vocabulary, naming speed) that I hypothesize will be key for MSWR. Both PA and MA are fundamental to literacy development (Carlisle, 2000; Rubin, Patterson, & Kantor, 1991; Shankweiler et al., 1995), and there has been some evidence that SSPA may also be important for word reading, especially for long words (Holliman et al., 2016; Wade-Woolley, 2016). I will consider each of these three metalinguistic skills in turn, followed by my two additional predictors.

**Phonological Awareness**

Over the last three decades, an extensive research base has been built in support of one such aspect of oral language: phonological awareness (PA; e.g., Adams, 1990; Ehri, Nunes, Stahl, & Willows, 2001). PA is defined as one’s knowledge of and ability to orally manipulate parts of words and is an umbrella term that includes syllable awareness, onset-rime awareness, and phoneme awareness. PA is one of the most important foundational skills for reading success (e.g., Bradley & Bryant, 1983; Burns, & Griffin, 1998; National Reading Panel, 2000; Goswami & Bryant, 1990; Goswami, 2000; Ehri, Nunes, Willows, Schuster, Yaghoub-Zadeh, & Shanahan, 2001).

PA tasks have been designed utilizing each of these sound units for various age groups: larger units (syllables) are easiest and most appropriate for younger children, while smaller units (phonemes) are more difficult, suitable for older children (Kirby, Desrochers, Roth, & Lai, 2008; Anthony & Lonigan, 2004). All PA tasks involve the perception and manipulation of the sound units of words. Importantly, “awareness” is not merely passive but an active and deliberate process (Castles & Coltheart, 2004); for example, Rime Awareness (e.g., Rime Oddity; “Which of the following words does not rhyme: fin, win, sit?”), and Phoneme Deletion (Bruce, 1964; “Say fan without saying /f/.”).
While PA is vital for successful early reading development, there is some evidence its contribution to reading may be developmentally limited (Griffiths, 1991; Stanovich, 1986), reaching a point at which it no longer explains individual differences in reading skill. Other studies suggest that it maintains a reliable predictive relationship in the elementary grades (Parrila, Kirby & McQuarrie, 2004; Wagner, Torgeson, Rashotte, Hecht, Barker, Burgess Donohoe & Garon, 1997) and even through high school (MacDonald & Cornwall, 1995). Because PA is focused at the sublexical level, it addresses the more fine-grained phonological information such as individual phonemes – what they are and how they are linked – but does not extend to or beyond the lexical level. Thus, one reason why PA’s importance might appear to decrease is that once they are competent readers, children have reached a sufficiently proficient degree of PA, after which improvement no longer makes a substantial contribution to word decoding, and the application of other skills are required to refine and improve performance. As words become longer, there are new complexities with which the reader must contend, such as syllable boundaries (Perry, Ziegler, & Zorzi, 2010), linguistic stress and vowel reduction (Chomsky, 1970; Seva, Monaghan, & Arciuli, 2009), vowel pronunciation ambiguities (Venezky, 1999), less predictable grapheme-phoneme correspondences (Berninger, 1994; Venezky, 1999), and morphological complexity (Nagy, Anderson, Schommer, Scott, & Stallman, 1989; Carlisle & Stone, 2005; Nagy, Berninger, & Abbott, 2006; Nagy, Berninger, Abbott, Vaughn, & Vermeulen, 2003).

Researchers continue to explore factors that add explanatory value to models of reading development, and in recent years there has been increased activity around two areas in particular: suprasegmental phonological (prosodic) awareness and morphological
awareness.

**Suprasegmental Phonological Awareness**

Prosody, the linguistic stress and intonation patterns of a language, is a complex construct best defined by the relative distribution of auditory indexes such as syllable duration, intensity, and frequency. Phonology exists in a hierarchical structure (Selkirk, 1978, 1980; Nespor & Vogel, 1982); like PA, prosody is a phonological feature, one that is applied at different levels of the prosodic hierarchy (e.g., Hayes, 1995). Traditional definitions of PA focus at the sublexical level, on component parts of words (syllable, onset-rime, phoneme); prosody operates at the suprasegmental level, at or above the level of the word. In English, prosodic awareness is often defined in terms of the alternation of strong (or stressed) and weak ( unstressed) syllables (e.g., within words (“CArrot”), across utterances (“She ATE the PURple CArrot.”)). Stressed syllables are distributed at even intervals, creating a distinct rhythm (e.g., “THIS is the RHYthm of WORDS.”), and their distribution also plays an important role in meaning (e.g., “I said SHE might consider it.” versus “I said she MIGHT consider it.”). Thus prosodic awareness refers to the ability to explicitly attend to and manipulate the linguistic rhythm of oral language. Research has shown that PA and prosodic awareness are related, but separate skills (typically $r = .3$ or $.4$, ranging between $r = .26$ (DEEdee; Whalley & Hansen, 2006) and $r = .74$ (Mispronunciations Task; Holliman, Wood, & Sheehy, 2008); see Wade-Woolley & Heggie (2016) for a review); herein, I use the term suprasegmental phonological awareness (SSPA) to refer to prosodic awareness (which is also known as prosodic sensitivity, linguistic stress, and linguistic rhythm) to highlight the relationship between PA and SSPA.
Reading researchers who focus on the relationship between word reading and SSPA (e.g., Goswami et al., 2002; Holliman, Wood, & Sheehy, 2008) have hypothesized that this relationship exists partially because of the processing of multisyllabic words (e.g., Corriveau, Pasquini, & Goswami, 2007; David, Wade-Woolley, Kirby, & Smithrim, 2007; Duncan & Seymour, 2003; Mundy, 2011; Wade-Woolley & Heggie, 2015), even noting specifically that prosodic awareness may "be an ability that is crucial for the transition from reading monosyllabic words to multisyllabic words" (Harrison & Wood, 2016, p. 88).

Although monosyllabic content words in English are stress-bearing, lexical stress is a more distinct feature in multisyllabic words due to stress contrasts between strong and weak syllables (e.g., CA-rrot), which typically result in vowel reduction (e.g., the different pronunciations of “o” in the first and third syllables in toMAto). Likewise, lexical stress is more variable in multisyllabic words and changes to a word’s stress pattern can affect its meaning (e.g., DEsert, deSSERT) or its grammatical category (e.g., REcord (noun), reCORD (verb)). The accurate decoding of multisyllabic words necessarily includes lexical stress assignment.

**Morphological Awareness**

Morphology is the study of morphemes, the semantic building blocks of words. Morphemes include stem and root words and affixes (prefixes, suffixes); for example, the word unlikely has three morphemes: un-, like, -ly. As a metalinguistic skill, morphological awareness (MA) extends beyond our implicit ability to comprehend and produce morphologically complex words to our explicit ability to reflect on and manipulate the rules of morphology (Carlisle, 1995, 2003; Kuo & Anderson, 2006). An understanding of morphology begins as a tacit part of our oral language acquisition (e.g., Tunmer &
Herriman, 1984); later, a more explicit reflection on the meaningful units of words benefits reading instruction because of the morphophonemic structure of English, where morphemes both transmit semantic information and provide productive orthographic cues (Carlisle & Kearns, *in press*).

Psycholinguistic studies of adults have consistently shown that morphological information is utilized when processing complex words (e.g., root word frequency; for a review, see Clahsen, Sonnenstuhl, & Blevins, 2003). As a morphophonemic language, English (like other alphabetic languages) adheres to the *isomorphism principle*, which states that morphemes are to be given priority over phonology, preserving orthographic consistency (Verhoeven, Schreuder, & Baayen, 2003); this can be a contributing factor to multisyllabic words’ complexity because grapheme-phoneme correspondences become more opaque, making decoding less straightforward. This increased complexity is one of the reasons that concern for students with reading disabilities has centered on their (in)ability to read multisyllabic and multimorphemic words (Archer, Gleason, & Vachon, 2003; Carlisle & Katz, 2006; Moats, 1998; Nunes & Bryant, 2006).

MA is thought to be particularly important for multisyllabic words because longer words are frequently morphologically complex (i.e., they contain two or more morphemes). It has been hypothesized that children in middle and upper elementary school with better morphological awareness may be better able to read and retain multisyllabic, morphologically complex words (e.g., Carlisle & Feldman, 1995; Mahoney et al., 2000; Singson et al., 2000). Nagy and colleagues demonstrated that morphological awareness made a significant unique contribution to decoding rate in Grades 4 through 9 (Nagy, Berninger, & Abbott, 2006). Deacon and colleagues provided evidence that children’s speed
and accuracy on multisyllabic, morphologically complex words was affected by morphological structure, even when the items varied in base frequency and morphological transparency (Deacon, Whalen, & Kirby, 2011). As the influence of PA diminishes over time, some research suggests that MA makes an increasing contribution to decoding (Shankweiler & Fowler, 2004; Singson et al., 2000) while others have found a more consistent role across grades (Roman, Kirby, Wade-Woolley, & Deacon, 2009). For example, Berninger and colleagues conducted a six-year longitudinal study examining growth in linguistic awareness between Grades 1 and 6. Using growth curve analysis, they demonstrated that the most growth in PA takes place prior to Grade 3, and that although MA begins to develop early in schooling, it continues to develop over the middle and late elementary years (Berninger, Abbott, Nagy, & Carlisle, 2010).

**Additional Predictors**

**Vocabulary.** There is ample evidence that vocabulary plays an important role in reading achievement (Biemiller, 2007; Biemiller & Boote, 2006; National Reading Panel, 2000), including studies that have suggested that vocabulary is related to word reading. Oral familiarity with words has been shown to increase the likelihood they will be read correctly (Nation & Cocksey, 2009), and researchers have posited that vocabulary knowledge may facilitate word identification because children will recognize words when they are sounding them out (Kirby, Desrochers, Roth, & Lai, 2008). Several studies have also demonstrated vocabulary’s relationship to exception word reading in English (Bishop & Snowling, 2004; Bowey, 2001; Ouellette, 2006; Nation & Snowling, 2004; Ricketts, Nation, & Bishop, 2007), which may be particularly relevant for MSWR because of the inconsistencies
inherent in spelling multisyllabic items, when both phonological and morphological factors are involved.

**Naming speed.** Naming speed, as measured by rapid automatized naming (RAN; Denckla & Rudel, 1976), is also strongly related to word reading (Cutting & Denckla, 2001; Kirby, Parrila, & Pfeiffer, 2003; Wolf & Bowers, 1999; Wolf, O'Rourke, Gidney, Lovett, Cirino, & Morris, 2002), a contribution that increases with age across the elementary school years (Kirby, Parrila, & Pfeiffer, 2003; Scarborough, 1998; Van den Bos, Zijlstra, & Spelberg, 2002). This increase in importance suggests that naming speed may be particularly relevant for a study of multisyllabic word reading, as these words increase in terms of semantic content and frequency of occurrence over the elementary years (Nagy & Anderson, 1984). The strength of naming speed’s relationship with reading outcomes goes beyond the particularities of English, however; it has been described as one of the “most universal, longitudinal, and concurrent predictors of reading ability” (Araújo, Reis, Petersson, & Faisca, 2015, p. 869).

**Error Analysis of Word Reading**

After a more general examination of multisyllabic word reading (Chapter 2), my second study (Chapter 3) breaks down the errors that children make when reading multisyllabic words and examines how different kinds of errors are related to their reading ability and to various reading-related skills (e.g., PA, MA). This more qualitative assessment of error type and rate expands on a more traditional, dichotomous treatment of word reading accuracy and has the potential to provide valuable information about the strategies children use when attempting to decode multisyllabic words (e.g., Biemiller, 1970; Chiappe & Siegel, 1999; Stuart & Coltheart, 1988). Children’s word reading error patterns provide
rich information not only about the course of development in typically achieving readers but also about those who are struggling (e.g., Bruck, 1993; Liberman, Rubin, Duques, & Carlisle, 1985; Worthy & Viise, 1996). The value of observing young children’s errors is evident in its foundational importance for theory building, where it constituted the starting point for the development of several major theories of literacy acquisition (Ehri, 1986; Frith, 1985; see Snowling, 2000).

**Chapter One: A Brief Review**

In this chapter, I have discussed the foundational role that reading plays in our academic and social success and provided a brief history of the research on reading development. I detailed the models of word reading development that are especially pertinent for our focus on single word reading. In anticipation of our two upcoming studies, I set forth the context for an examination of multisyllabic word reading and discussed the rationale for the child-level skills we have selected to measure.

**Dissertation Advance Organizer**

This dissertation is composed of two empirical studies of middle school students’ multisyllabic word reading. Study 1 (Chapter 2) investigates individual differences in Grade 4 and 5 students’ multisyllabic word reading and the extent to which five child-level factors (PA, SSPA, MA, vocabulary, and naming speed) account for individual differences in students’ multisyllabic word reading accuracy. Study 2 (Chapter 3) analyzes the errors made by Grade 4 and 5 students as they attempted to read multisyllabic words at and above their grade level. Discussions of the findings accompany each study, followed by a general discussion in Chapter 4.
CHAPTER TWO: Multisyllabic Word Reading in Grades 4 and 5: Contributions of Phonological Awareness, Morphological Awareness, and Naming Speed

Abstract

Although over 90% of English words are multisyllabic, relatively little research has focused explicitly on long words and the skills children use to read them, perhaps because of the layers of complexity inherent in this more advanced stage of word reading (e.g., morphology, lexical stress assignment). Readers who struggle through elementary and into secondary school tend to have difficulty specifically with multisyllabic words, a problem enhanced by the fact that these items constitute an increasingly large proportion of the texts they read and also carry a majority of the semantic information vital for reading comprehension. Our aim was to investigate individual differences in multisyllabic word reading and the extent to which three metalinguistic skills (phonological awareness, prosodic awareness, and morphological awareness) account for individual differences in students’ multisyllabic word reading accuracy. English-speaking students in Grades 4 and 5 ($n=93$) were recruited for this study. Participants were administered measures of metalinguistic skills (phonological awareness, suprasegmental phonological awareness, morphological awareness), control measures (nonverbal reasoning, vocabulary, naming speed), and word reading ability (Word Identification and an experimental measure of multisyllabic word reading). Regression analyses showed that the contribution of PA (both segmental and suprasegmental) was fully mediated by that of morphological awareness. Naming speed and morphological awareness predicted children’s multisyllabic word reading accuracy at this stage of reading development.

Keywords: Multisyllabic; Word Reading; Morphological Awareness; Prosodic Awareness; Prosody
The majority of English words are multisyllabic (Balota et al., 2007; Baayen, Piepenbrock, & Gulikers, 1995) and constitute an increasingly large proportion of the words students encounter in print, beginning as early as Grade 3 (Archer, Gleason, & Vachon, 2003; Nagy & Anderson, 1984; Zeno, Ivens, Millard, & Duvvuri, 1995). A large number of multisyllabic words is a problem for students who continue to struggle with reading through elementary and into secondary school because these students tend to have particular difficulty with longer words (Archer, Gleason, & Vachon, 2003; Moats, 1998). Compared to studies involving monosyllabic words, multisyllabic word reading (MSWR) receives relatively little attention. In the current study, we investigate the extent to which three metalinguistic skills (phonological awareness (PA), prosodic awareness (or suprasegmental phonological awareness; SSPA), and morphological awareness (MA)) account for individual differences in students’ multisyllabic word reading accuracy.

**Multisyllabic Words and Related Skills**

English orthography is considered irregular or opaque because the grapheme-phoneme correspondences are inconsistent (e.g., the grapheme *ea* in *head* and *heal*). Much of this inconsistency comes from the fact that English orthography is morphophonemic: how words are pronounced depends both on phonology and on morphology, with the preservation of morphological information at the expense of phonology (Chomsky & Halle, 1968; Venezky, 1999). Because multisyllabic words are most often morphologically complex, reading them accurately necessitates both phonological and morphological information. The current study focuses on the relative contributions of three oral language skills (two levels of phonological awareness (PA and SSPA) and MA), in addition to control variables (vocabulary, naming speed), to children’s accuracy in multisyllabic word reading.
Individual Differences in MSWR

In addition to the literature discussed in Chapter 1, there is some recent research that is specifically relevant to the current study. One of the oral language skills of interest in the current study is SSPA. While it is now undisputed that PA contributes to decoding, with a growing body of research supporting MA as a fundamental component of both word reading and reading comprehension (e.g., Carlisle, 2000; Deacon & Kirby, 2004; Mahony, Singson, & Mann, 2000; see Chapter 1), SSPA has received relatively less attention. As a result, we have aimed to review those studies that include all three variables: PA, SSPA, and MA.

There are several ways that a relationship between SSPA and word reading may manifest itself, whether directly or indirectly (Wood, Wade-Woolley, & Holliman, 2009). A recent study examined these potential pathways in beginning (M age = 6.79 years) readers. Kim & Petscher (2015) used structural equation modeling to systematically test and compare five alternate models of the relation of SSPA to word reading. They tested: (a) an indirect effect of SSPA on word reading via PA; (b) an indirect effect of SSPA on word reading via MA; (c) indirect effects of SSPA on word reading via PA and MA; (d) a direct effect of SSPA on word reading; and (e) a direct and indirect effects model via PA and MA. Kim & Petscher’s analyses led them to select Model 3 (i.e., (c) above) as the best fitting, with significant relationships between MA, PA, letter naming fluency and word reading as well as direct relationships between PA, MA, and SSPA; they concluded that SSPA’s relationship to reading is fully mediated by phonological awareness and morphological awareness. This outcome aligns with another study of SSPA in early readers (M age = 6y2m; Holliman et al., 2014b); Holliman and colleagues’ path analysis revealed that for these young readers,
SSPA’s relationship to word reading was indirect, partially mediated in two paths: 1) vocabulary to MA, and 2) vocabulary to PA (at the level of the onset-rime).

More recently, Holliman and colleagues examined multisyllabic word reading directly, although in an exploratory analysis using disaggregated scores on the British Ability Scales’ word reading measure (Holliman, Gutiérrez Palma, Critten, et al., 2016). Holliman and colleagues’ analysis revealed that young (M age = 5.77 years) readers’ SSPA was able to explain unique variance in MSWR in the context of both PA and MA, predicting an additional 13.5% (as opposed to 3.8% unique variance in monosyllabic word reading). Their findings accord with a recent comparison of mono- and multisyllabic word reading by Wade-Woolley (2016), who tested whether PA and SSPA were differentially related to the reading of short and long words in 10-year-old children. She found that both types of PA were implicated in both short and long word reading, suggesting that the two are complementary but not redundant processes. Unfortunately, a measure of MA was not included.

An important consideration for the current study is that SSPA may manifest its influence on word reading differently in younger readers than in older readers, where the influence of PA may have begun to wane in favour of other skills (i.e., MA, naming speed) and in response to longer, more complex words. The nature of the relationship between SSPA and word reading may also have been constrained by a lack of multisyllabic words in Kim & Petscher’s outcome measures. For example, a cross-sectional study of students in Grades 3, 5, and 7 suggests that the inter-relationship of PA, SSPA, and MA may differ for older students: Clin, Wade-Woolley, and Heggie (2009) found that both SSPA and MA predicted reading ability above the influence of PA, accounting for 2-3% (SSPA) and 6-7% (MA) of the variance in reading ability. This is in keeping with the theory that PA’s influence
on decoding wanes while MA’s influence continues to grow over the elementary years. However,Clin and colleagues used a composite measure of reading (reading comprehension, fluency, decoding, and accuracy) for their outcome, making their findings more difficult to interpret in light of our current focus on single word reading and on multisyllabic words.

The Current Study

This study addresses individual differences in multisyllabic word reading. First, we investigated whether three metalinguistic skills (phonological awareness, morphological awareness, and prosodic awareness) would predict variance in children’s multisyllabic word reading accuracy. Second, we asked whether these skills would retain their predictive power over and above the influence of control measures (nonverbal reasoning, vocabulary, naming speed). We hypothesized that each of the three metalinguistic variables would be important specifically for participants’ accurate readings of multisyllabic words and that their relationship would survive the influence of the control measures. Depending on the interrelationship of the metalinguistic variables, we may conduct exploratory mediation analyses.

Method

Participants

Students in Grades 4 and 5 from elementary schools in Eastern Ontario participated in this study (n=93; Grade 4 n=42, 51 female; M_age = 10 years, 5 months). Participants were screened for language proficiency (only native English speakers were eligible to participate) and for cognitive/developmental typicality (i.e., students with developmental disabilities were excluded). Participants were not screened for reading disabilities or difficulties. With parental consent, students met one-on-one with a trained research
assistant four times for sessions lasting no longer than 30 minutes each. Testing occurred in a quiet room in the student’s school in the late spring of the school year.

**Materials**

*Control measures.*

**Vocabulary.** Participants’ expressive vocabulary was measured using the Vocabulary subtest of the Wechsler Abbreviated Scale of Intelligence (WASI Vocab; Wechsler, 1999). This task consists of 42 items: four pictures followed by 38 words to be orally defined. Participants are instructed to define the word (e.g., “What is DINNER?”). Responses are awarded 2 (e.g., a good synonym, correct figurative use), 1 (e.g., vague or less pertinent synonym), or 0 (e.g., no real understanding shown, poverty of content) points. No strict time limit is enforced but if participants pause for more than 30 seconds on an item, it is given a score of 0 and the next item is administered. Participants at this age group proceed to item 34 or until they make 5 consecutive errors, for a maximum score of 64. The sum of the items correctly defined constitutes the total raw score on WASI Vocab. Reliability is reported in the manual at .92 for this age group.

**Nonverbal reasoning.** The Matrix subtest of the WASI was used to measure participants’ nonverbal reasoning (Wechsler, 1999). The task consists of 35 incomplete patterns grouped by strategy: pattern completion, classification, analogy, and serial reasoning. Participants must select, from five available options, the image that completes the pattern for each item. Items are scored correct (1) or incorrect (0) with a stop rule of 4 consecutive incorrect responses. Reliability is reported in the manual at .89 for this age group.

**Naming speed.** The Digits subtest of Rapid Automatized Naming (RAN Numbers;
Wagner, Torgesen, & Rashotte, 1999) was included to control for naming speed. This task consists of two arrays of single digit numbers (2, 6, 9, 4, 7), arranged in four rows of 10 digits over two pages (Form A, Form B). Participants are asked to name the digits as quickly and accurately as possible, and the time taken to name all the digits on a page is recorded. Both Form A and Form B are administered unless the participant makes more than four errors on Form A. Alternate-form reliability is reported in the manual at .90 for this age group.

**Phonological awareness.**

**Phonological awareness.** Phonological awareness was measured using two subtasks of the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999), Blending Words and Elision. Blending Words consists of 20 prerecorded items and is presented over headphones. Participants are told that they will “hear some words in small parts, one at a time” and that they should “put these parts together to make a whole word.” After practice items (e.g., “What word do these sounds make?” “can” – “dy”), feedback can be given for all practice items and the first three test items. Items are scored as correct/incorrect and testing proceeds until three consecutive incorrect responses have been awarded. Cronbach’s alpha was reported at .87 in the manual for this age group.

The Elision subtask consists of 20 items and six practice items, beginning with compound words (e.g., toothbrush, airplane) through 2-syllable (e.g., spider, powder) and 1-syllable (e.g., cup, meet, snail) words. Participants are asked to pronounce items without saying increasingly small parts of the words (word, syllable, phoneme) in increasingly difficult locations (initial phoneme, final phoneme, medial phoneme). For example, “Say 
doughnut. Now say doughnut without saying dough.” and “Say tiger. Now say tiger without
saying /g/.” Responses are scored as correct/incorrect and testing is discontinued after three consecutive failed items. Cronbach’s alpha was reported in the manual at .91 for this age group.

**Suprasegmental phonological awareness.** An experimental measure of SSPA, Aural Stress Assignment (ASA; Wade-Woolley, 2016), was administered orally to participants using prerecorded items over headphones. The task consisted of thirty-two multisyllabic English words between two (e.g., *answer*) and five (e.g., *organization*) syllables in length. Participants were asked to locate the syllable carrying the word’s primary stress (e.g., in the word *opportunity*, the primary stress is on the third syllable, *o-*por-TU-ni-*ty*). Blank tokens representing the number of syllables in each word were provided to reduce cognitive load. For their responses, participants were encouraged to physically manipulate (by sliding forward) the tokens according to the relative emphasis of the word’s pronunciation. For example, three tokens were placed on the table for the word “decision,” and participants nudged the tokens forward with force relative to the syllable’s strength as they repeated the word: “de-CI-sion.” Token manipulation was offered as a way to support participants’ thinking, but was not mandatory. In order to provide an answer for each item, the participant was asked to indicate which syllable had the strongest beat. ASA consists of two practice items and 30 test items; participants worked with the two practice items, and other nouns such as their first name, until they reported understanding the task. Cronbach’s alpha was computed at .64 for this sample.

**Morphological awareness.** Two measures of morphological awareness, both suffix choice tasks, were included. The first is a well-known completion task based on Carlisle’s (1988) work, in which children are asked to orally produce the derived form of a word
based on a given root using carrier sentences as prompts (e.g., Prompt: “The word is “farm.” My uncle is a ______.” Response: “Farmer.”). Cronbach’s alpha was computed at .89 for the current sample.

The second morphological awareness task is the Nonword Suffix Choice task (Muse, 2005), an 18-item task presenting English sentences to be completed by one of four pseudoword options (same stem, completed by different suffixes). Participants must use the suffix to determine which nonword is the appropriate choice to complete the sentence. For example, “The man is a great ________,” with four options: tranter, tranting, trantious, or trantiful. To reduce the influence of both decoding ability and working memory, the sentence and each option is read aloud by the research assistant as participants follow along; participants then circle their answer before the research assistant moves on to the next item. Items on this task were constructed by Muse (2005), based on previous work (Mahony, 1994; Singson, Mahony, & Mann, 2000; Nagy, Diakidoy, & Anderson, 1993; Tyler & Nagy, 1989, 1990) and later adjusted, using the same paradigm (M. Kieffer, personal communication, April 12, 2014). In previous work, Cronbach’s alpha = .78 for native English speakers in Grade 6 (Kieffer & Box, 2013); alpha was computed at .82 for the current sample.

**Reading measures.**

*Word reading.* Participants’ word reading was tested using the Word Identification (Word ID) subtest of the Woodcock Reading Mastery Tests (WRMT; Woodcock, 1998). Word ID consists of 106 English words of increasing difficulty; participants begin at an entry point dependent on their age (e.g., students in Grade 5 begin at Item #56). Items are scored as correct or incorrect, and testing continues until the participant makes six
consecutive errors. Split half reliability is reported in the manual at .91 for this age group.

*Multisyllabic word reading.* An experimental measure of multisyllabic word reading was created for this study, consisting of 54 trisyllabic words (see Appendix B). Using both the MRC Psycholinguistics Database (an online corpus of 150,837 words with 26 linguistic and psycholinguistic attributes for each; Coltheart, 1981) and the Educator's Word Frequency Guide (Zeno, et al., 1995) items were compiled based on: stress pattern, number of morphemes, age of acquisition, standard frequency index (SFI), and word length (defined both as number of letters and number of phonemes). Only words that appeared at least once in either the Grade 4, Grade 5, or Grade 6 corpora were included in the final list of 54 items. Participants were asked to read each word aloud, giving their best attempt even if they were unsure; participants were not timed and responses were scored correct/incorrect at the time of administration, based on the participant’s final answer. Items were presented over three testing periods to avoid tiring participants. Test items were recorded by a female native speaker of English in a sound-attenuating booth at CD quality (44.1 kh) sampling rate, using Audacity® 2.1.0 recording and editing software, and were presented to participants over headphones. Responses were recorded for subsequent confirmation using a Sony Digital Voice Recorder. See Appendix B for a full table of multisyllabic items and their attributes. Cronbach’s alpha was computed at .97.

**Results**

Descriptive statistics for raw scores on all measures are reported in Table 1, including skewness and kurtosis statistics. To characterize the sample, we calculated the mean standard score on Word ID (M = 99.97) and the T-scores on WASI Vocabulary (M = 47.74) and WASI Matrix (M = 47.14); means were similar to the norming sample in all cases.
Outliers were defined as scores 2.5 SDs above or below the mean for each variable; four outliers were identified on two tasks (one on RAN Numbers and three on Aural Stress Assignment) and these scores were Winsorized to minimize data loss. Only Elision was significantly skewed (Tabachnick & Fidell, 1996); scores were corrected using a reflected square root transformation. Raw scores are reported in Table 2.1, but transformed scores were used for Elision in all analyses.
Table 2.1

Descriptive Statistics (n=93)

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>Max</th>
<th>M</th>
<th>SD</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>WASI Vocabulary (/64)</td>
<td>17</td>
<td>56</td>
<td>36.40</td>
<td>9.59</td>
<td>-.012</td>
<td>-.861</td>
</tr>
<tr>
<td>WASI Matrix (/35)</td>
<td>3</td>
<td>30</td>
<td>19.44</td>
<td>6.83</td>
<td>-.560</td>
<td>-1.005</td>
</tr>
<tr>
<td>RAN Numbers</td>
<td>21</td>
<td>54</td>
<td>34.22</td>
<td>7.00</td>
<td>-.301</td>
<td>3.244</td>
</tr>
<tr>
<td>Blending Words (/20)</td>
<td>6</td>
<td>20</td>
<td>12.72</td>
<td>3.69</td>
<td>.141</td>
<td>-1.121</td>
</tr>
<tr>
<td>Elision (/20)</td>
<td>4</td>
<td>20</td>
<td>14.77</td>
<td>4.49</td>
<td>-.927</td>
<td>-.560</td>
</tr>
<tr>
<td>Aural Stress Assignment (/30)</td>
<td>5</td>
<td>23</td>
<td>12.59</td>
<td>4.22</td>
<td>.624</td>
<td>-.115</td>
</tr>
<tr>
<td>Suffix Completion (/44)</td>
<td>7</td>
<td>40</td>
<td>25.92</td>
<td>7.38</td>
<td>-.401</td>
<td>-.237</td>
</tr>
<tr>
<td>Nonword Suffix Choice (/18)</td>
<td>2</td>
<td>18</td>
<td>10.87</td>
<td>4.34</td>
<td>-.114</td>
<td>-1.096</td>
</tr>
<tr>
<td>Word ID (/106)</td>
<td>38</td>
<td>94</td>
<td>69.58</td>
<td>12.01</td>
<td>-.374</td>
<td>-.458</td>
</tr>
<tr>
<td>Multisyllabic Word Reading (/54)</td>
<td>0</td>
<td>54</td>
<td>33.72</td>
<td>14.88</td>
<td>-.540</td>
<td>-.862</td>
</tr>
</tbody>
</table>

*Note. Standard Error of Skewness = .250; Standard Error of Kurtosis = .495*

Correlations between variables, including the MA composite, are available in Table 2.2; a full table of correlations is in Appendix A. To increase power and reduce error, a composite score was calculated by averaging z-scores for Morphological Awareness (Suffix Completion and Nonword Suffix Choice; \( r = .706, p < .001 \)). The correlation between the two phonological awareness measures \( r = .328, p < .01 \) was deemed insufficiently strong to create a composite; each PA measure will be considered separately in subsequent models.
Table 2.2

*Pearson Correlations Between Measures*

<table>
<thead>
<tr>
<th></th>
<th>Vocab</th>
<th>Matrix</th>
<th>RAN</th>
<th>PA: Blending</th>
<th>PA: Elision</th>
<th>SSPA</th>
<th>MA Comp</th>
<th>Word ID</th>
<th>MSWR</th>
</tr>
</thead>
<tbody>
<tr>
<td>WASI Vocab</td>
<td>—</td>
<td>.245*</td>
<td>—</td>
<td>.373***</td>
<td>.421***</td>
<td>.254*</td>
<td>.769***</td>
<td>.612***</td>
<td>.611***</td>
</tr>
<tr>
<td>WASI Matrix</td>
<td>—</td>
<td>.060</td>
<td>.250*</td>
<td>.345**</td>
<td>.069</td>
<td>.454***</td>
<td>.314**</td>
<td>.311**</td>
<td></td>
</tr>
<tr>
<td>RAN Numbers</td>
<td>—</td>
<td>.001</td>
<td>—</td>
<td>-.152</td>
<td>-.036</td>
<td>-.316**</td>
<td>-.420***</td>
<td>-.429***</td>
<td></td>
</tr>
<tr>
<td>PA: Blending</td>
<td></td>
<td>.328**</td>
<td>.136</td>
<td>.431***</td>
<td>.368***</td>
<td>.381***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA: Elision</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.199</td>
<td>.494***</td>
<td>.472***</td>
<td>.466***</td>
</tr>
<tr>
<td>SSPA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>—</td>
<td>.276**</td>
<td>.358***</td>
<td>.318**</td>
</tr>
<tr>
<td>MA Composite</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>—</td>
<td>.805***</td>
<td>.824***</td>
<td></td>
</tr>
<tr>
<td>Word ID</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>—</td>
<td></td>
<td></td>
<td>.938***</td>
</tr>
</tbody>
</table>

*Note.* *p < .05, **p < .01, ***p < .001.

As Table 2.2 shows, all of the variables were significantly correlated with the multisyllabic word reading (MSWR) measure. Some variables were highly correlated (*r > .80*); we addressed this by examining multicollinearity statistics. Variance inflation factors (VIFs) for the independent variables ranged from 1.088 to 3.602, well under the critical value of 5.0, indicating that multicollinearity is unlikely to be a source of concern for linear regression (Bowerman & O'Connell, 1990; Tabachnick & Fidell, 2007).

**PA, SSPA, and MA: Predictors of MSWR?**

All three variables of interest were related to multisyllabic word reading; zero order correlations were significant, ranging from .32 (SSPA) to .79 (MA) (see Table 2.2). Multiple regression was used to determine whether the three variables of interest predicted multisyllabic word reading: PA (Blending Words and Elision, considered separately), SSPA (Aural Stress Assignment), and MA (MA Composite) were entered as independent variables.
and predicted multisyllabic word reading as the dependent variable (see Table 2.3). PA was entered first, due to its foundational importance to word reading. Following this, we added the other phonological measure, SSPA, in the second step and MA in the final step.

Table 2.3

*Multiple Regression Predicting Multisyllabic Word Reading (n=93)*

<table>
<thead>
<tr>
<th>Step</th>
<th>$R^2$</th>
<th>$R^2$ change</th>
<th>$B$</th>
<th>$SE$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA: Blending Words</td>
<td>.275</td>
<td>.275***</td>
<td>1.031</td>
<td>.383</td>
<td>.256**</td>
</tr>
<tr>
<td>PA: Elision</td>
<td></td>
<td></td>
<td>8.571</td>
<td>2.130</td>
<td>.382***</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA: Blending</td>
<td></td>
<td></td>
<td>.961</td>
<td>.374</td>
<td>.238*</td>
</tr>
<tr>
<td>PA: Elision</td>
<td>.320</td>
<td>.045*</td>
<td>7.730</td>
<td>2.104</td>
<td>.345***</td>
</tr>
<tr>
<td>SSPA</td>
<td>.320</td>
<td>.045*</td>
<td>.766</td>
<td>.315</td>
<td>.217*</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA: Blending Words</td>
<td>.692</td>
<td>.372***</td>
<td>1.514</td>
<td>1.547</td>
<td>.068</td>
</tr>
<tr>
<td>PA: Elision</td>
<td>.692</td>
<td>.372***</td>
<td>1.514</td>
<td>1.547</td>
<td>.068</td>
</tr>
<tr>
<td>SSPA</td>
<td>.692</td>
<td>.372***</td>
<td>.330</td>
<td>.218</td>
<td>.093</td>
</tr>
<tr>
<td>MA</td>
<td>.692</td>
<td>.372***</td>
<td>12.185</td>
<td>1.182</td>
<td>.756***</td>
</tr>
</tbody>
</table>

** = $p < .01$, *** = $p < .001$

As Table 2.3 shows, when entered as the sole predictors (Step 1), both PA measures (Blending and Elision) made independent contributions, together accounting for 27.5% of the variance in MSWR. In Step 2, SSPA accounted for an additional 4.5% of multisyllabic word reading. When morphological awareness (MA) was included as a predictor (Step 3), it accounted for an additional 37.2%. Neither PA nor SSPA were significant predictors in the final model; MA’s addition reduced their contributions to essentially zero ($\beta = .020, .068, .093, p = .764, .330, and .133$, respectively), consistent with the interpretation that MA mediates the relationship of phonological awareness (at both the segmental and suprasegmental levels) to multisyllabic word reading. To confirm whether MA was in fact a mediator, we assessed each component of the proposed mediation model (see Figure 2.1)
for paths) using regression analyses (using bootstrapping; see below). Because the
PROCESS macro for SPSS (Hayes, 2013) returns only unstandardized $B$ coefficients, we first
standardized all variables by creating z-scores to improve comparison of coefficients.

PA (Elision$^1$) and SSPA were each positively associated with multisyllabic word
reading (c-path, total effect; $B = .466, t(91) = 5.025, p < .001$ and $B = .318, t(91) = 3.201, p < .01$, respectively); PA and SSPA were also positively associated with MA (a-path; $B = .494, t(91) = 5.415, p < .001$ and $B = .276, t(91) = 2.735, p < .01$, respectively). Lastly, results
indicated that the mediator, MA, was positively associated with multisyllabic word reading
(b-path; $B = .785, t(91) = 11.521, p < .001$ (PA) and $B = .797, t(91) = 13.011, p < .001$
(SSPA)). Because both the a-path and b-path were significant, mediation analyses were
tested using the bootstrapping$^2$ method with bias-corrected confidence estimates (Hayes &
Scharkow, 2013; MacKinnon, Lockwood, & Williams, 2004; Preacher & Hayes, 2004). In the
present study, the 95% confidence interval of the indirect effect was obtained with 5000
bootstrap resamples (Preacher & Hayes, 2008). Results of this analysis confirmed the
mediating role of MA in the relation between both PA and SSPA and MSWR ($B = .388, CI =
.242$ to .539 and $B = .220, CI = .060$ to .370, respectively). In addition, results indicated that
the direct effect of both PA and SSPA became non-significant ($B = .078, t(91) = 1.149, p =
.254$ and $B = .098, t(91) = 1.607, p = .112$, respectively) when controlling for MA, suggesting
full mediation. Sobel’s test and Kappa-squared (Preacher & Kelley, 2011) confirmed this

$^1$ For the sake of parsimony, we have used Elision as the measure of PA for mediation analyses; it
was selected because it was more highly correlated with the outcome. The models using Blending
Words were similar to those presented herein.

$^2$ Bootstrapping is “one of the more valid and powerful methods for testing intervening variables
effects” with the best Type I Error control of the various mediation analyses (Hayes, 2009, p. 412); it
is a nonparametric resampling procedure that repeatedly samples from the data set, with
replacement, to approximate the sampling distribution of the indirect effect ($ab$) and construct
confidence intervals for the indirect effect (Preacher & Hayes, 2008).
interpretation ($Z = 4.89, p < .001, \kappa^2 = .431$ and $Z = 2.67, p < .01, \kappa^2 = .279$, respectively). All results are displayed in Figures 2.1a (PA) and 2.1b (SSPA).

Figure 2.1a.

Figure 2.1b.

Figures 2.1a and 2.1b depict the indirect effect of PA (segmental and suprasegmental, respectively) on MSWR through MA. In these models, PA/SSPA affects MA (a-path), and this effect propagates to MSWR (b-path). The models also depict the total effect of PA/SSPA on MSWR (c-path) and the direct effect of PA on MSWR (c'-path). In the case of full mediation, the introduction of the mediator significantly reduces the c'-path. All paths are quantified by standardized regression coefficients.

Note. For the sake of parsimony, mediation model 2.1a (PA) uses Elision; model was tested using Blending and achieved the same outcome (full mediation).

Note. ** $p < .01$, *** $p < .001$
It is important to note that despite the outcome of these analyses, based on the current study’s cross-sectional design we are unable to truly support causal inferences such as mediation (Kline, 2015) and cannot rule out the possibility that PA, SSPA, and MA are simply correlated. However, the outcome of our analyses does suggest that future work specifically designed to test MA’s role as a mediator of the relationship between both levels of phonological awareness (segmental and suprasegmental) and multisyllabic word reading is warranted.

**PA, SSPA, and MA: Predictive of MSWR above age, nonverbal reasoning, vocabulary, and naming speed?**

To test whether PA, SSPA, and MA would survive the addition of pertinent control variables to the model, we ran a multiple regression with the four control variables in the first step, followed by the variables from our first research question in the same order as above (i.e., the second (PA), third (SSPA), and fourth (MA) steps); see Table 2.4.
### Table 2.4

*Multiple Regression Predicting Multisyllabic Word Reading (n=93)*

<table>
<thead>
<tr>
<th>Step</th>
<th>R²</th>
<th>ΔR²</th>
<th>B</th>
<th>SE</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td>.054</td>
<td>.171</td>
<td>.026</td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td></td>
<td>.501</td>
<td>.501***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nonverbal Reasoning</td>
<td>.474</td>
<td>.173</td>
<td>.215**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vocabulary</td>
<td>.734</td>
<td>.133</td>
<td>.472***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RAN</td>
<td>- .699</td>
<td>.168</td>
<td>- .327***</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td>.145</td>
<td>.167</td>
<td>.068</td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td></td>
<td>.325</td>
<td>.173</td>
<td>.148</td>
</tr>
<tr>
<td></td>
<td>Nonverbal Reasoning</td>
<td>.533</td>
<td>.143</td>
<td>.343***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vocabulary</td>
<td>.552</td>
<td>.050*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RAN</td>
<td>.577</td>
<td>.026*</td>
<td>- .703</td>
<td>.159</td>
</tr>
<tr>
<td></td>
<td>PA: Blending Words</td>
<td>.625</td>
<td>.330</td>
<td>.155</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PA: Elision</td>
<td>4.064</td>
<td>1.960</td>
<td>.176*</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>.183</td>
<td>.164</td>
<td>.086</td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td></td>
<td>.334</td>
<td>.169</td>
<td>.152</td>
</tr>
<tr>
<td></td>
<td>Nonverbal Reasoning</td>
<td>.466</td>
<td>.143</td>
<td>.300**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vocabulary</td>
<td>.577</td>
<td>.026*</td>
<td>- .703</td>
<td>.159</td>
</tr>
<tr>
<td></td>
<td>RAN</td>
<td>.733</td>
<td>.156***</td>
<td>- .411</td>
<td>.134</td>
</tr>
<tr>
<td></td>
<td>PA: Blending Words</td>
<td>.234</td>
<td>.263</td>
<td>.058</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PA: Elision</td>
<td>1.958</td>
<td>1.556</td>
<td>.085</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SSPA</td>
<td>.369</td>
<td>.212</td>
<td>.105</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MA</td>
<td>12.214</td>
<td>1.752</td>
<td>.756***</td>
<td></td>
</tr>
</tbody>
</table>

* *p < .05, ** p < .01, *** p < .001

In the first step, the control variables account for 50% of the variance in multisyllabic word reading with nonverbal reasoning, vocabulary, and naming speed contributing significantly.
When the two PA measures are added in the fourth step, they contribute an additional 5% of the variance, carried by Elision; in the fifth step, SSPA contributes an additional 2.6% to multisyllabic word reading. When added in the final step, MA contributes an additional 15.6%; only MA and naming speed remain unique predictors of multisyllabic word reading.

**Discussion**

The current study was designed to examine individual differences in multisyllabic word reading, guided by two specific research questions. First, we wished to determine whether three child-level variables of interest (PA, SSPA, and MA) would predict accuracy in multisyllabic word reading (MSWR). Second, we examined these variables in the context of other predictors, namely age, nonverbal reasoning, vocabulary, and naming speed. We hypothesized that these three initial variables would be important for participants’ accurate readings of multisyllabic words, and that their relationship with MSWR would persist in the context of these other variables. We begin by discussing the relationship between PA, SSPA, MA and MSWR, including MA’s role as a potential mediator. Following this, we discuss the relationship between these variables in the context of the control variables. Finally, we discuss the limitations of the current study and how they might inform future work on MSWR.

**Metalinguistic Skills & MSWR**

Our first research question asked whether three metalinguistic skills (PA, SSPA, MA) would predict variance in children’s multisyllabic word reading accuracy. MA’s emergence as a strong, robust predictor of MSWR is in agreement with a considerable amount of evidence showing that MA makes a unique contribution to various reading abilities, even after controlling for factors such as intelligence, working memory, vocabulary, and PA (e.g.,
Bryant, Nunes, & Bindman, 1998; Carlisle, 1995, 2003; Carlisle & Stone, 2003; Kirby, et al., 2012). Although both PA and SSPA initially predicted variance in MSWR, their impact was greatly reduced with the addition of MA. This finding is supported by previous research suggesting that MA makes an independent contribution to word reading beyond PA at this stage of reading development (e.g., Mahony, Singson, & Mann, 2000; Roman, Kirby, Parrila, Wade-Woolley, & Deacon, 2008; Singson, Mahony, & Mann, 2000). Morphemes are strongly associated with phonological units like phonemes and syllables (McBride-Chang, Wagner, et al., 2005); phonology is “an essential part” of morphology, as it is the basis for initial recognition of morphemes (Carlisle, 2003, p. 307); but, morphology is understood to be more than an extension of phonology (Nunes & Bryant, 2009).

The picture is much the same when SSPA is included. Studies have shown that SSPA makes an independent contribution to word reading in the context of both PA and MA. In the current study, with multisyllabic (often morphologically complex) items constituting the word reading outcome, MA made an independent contribution of 15.6%, substantially greater than contributions to word reading typically reported (~5%) (e.g., Mahoney et al., 2000; Roman et al., 2009); the increased impact of MA in the current study may have resulted from a focus on multisyllabic words, for which phonological decoding is necessary but not sufficient for word reading in an opaque, morphophonemic orthography; longer words have an internal structure that involves grapheme-phoneme connections requiring other decoding methods, including morphological analysis (Verhoeven & Perfetti, 2011; Kirby & Bowers, in press).

**MA as a mediator of PA/SSPA and MSWR.** In terms of the mediating relationship of MA, an indirect relationship between SSPA and MSWR was one of the pathways
hypothesized by Wood, Wade-Woolley, & Holliman (2009). In Wood and colleagues’ proposed model, SSPA should make both a direct contribution to reading development, independent of its contribution via vocabulary development and PA, as well as an indirect contribution via MA. Wood and colleagues’ model has been specifically tested with beginning readers (5-7 years old) (Holliman, Critten, Lawrence, Harrison, Wood, & Hughes, 2014). At this early stage of reading development, the authors’ path analysis saw both vocabulary and MA surface as key mediators between PA and SSPA and word reading: SSPA was mediated by both vocabulary and MA and by PA (at the rime level); PA (at the rime level) linked directly to word reading, while PA (at the phoneme level) linked only through MA. Holliman and colleagues suggested that these younger children might have been making use of the linguistic rhythm of oral language (e.g., linguistic stress) to isolate and acquire words from the speech stream (Cutler & Mehler, 1993) before making connections to orthographic information. Their findings were similar to those of Kim & Petscher (2015), who showed that SSPA’s relationship to word reading was fully mediated by both MA and PA. Results of Holliman and colleagues’ and Kim and Petscher’s studies, as well as the current study’s, suggest that MA continues to act as a mediator of both PA and SSPA in more advanced readers.

The relationship between SSPA and vocabulary may change from that reported by Holliman and colleagues in more advanced readers, as MA and vocabulary become more closely linked in upper elementary grades (e.g., Wagner, Muse, & Tannenbaum, 2007; Wagner et al., 2007). There is compelling evidence to suggest that MA and vocabulary constitute one dimension in later elementary school, a relationship not subject to developmental change between Grades 4 and 8 (Spencer, et al., 2015). The inclusion of
vocabulary as a covariate in subsequent studies designed to test MA as a mediator between PA/SSPA and MSWR could provide further evidence for the dimensionality of MA/vocabulary; if the two are indeed one dimension, we would expect to find that vocabulary participates in a similar mediating relationship between PA/SSPA and MSWR. Finally, Holliman and colleagues (2014) did not include a measure of naming speed, which surfaced in the current study as a significant predictor of MSWR in the context of a number of relevant variables; its inclusion in future studies is important for a more complete picture of the factors influencing MSWR.

Given what we know about the relative contributions of PA and MA in beginning and more advanced readers (e.g., Berninger, Abbott, Nagy, & Carlisle, 2010; McCutchen, Green, & Abbott, 2008), it is likely that the differences between these findings and the current results lie in developmental differences in how PA and MA influence single word reading, including MSWR specifically. The current results suggest that both PA and SSPA’s initial (in the case of beginning readers) and direct influence on word reading may shift to an indirect one, via MA, as the reader’s skill develops and the type of words to be read become longer and more complex.

**In the Context of Other Predictors**

Our second research question focused on whether the relationship between PA, SSPA, MA and MSWR would survive the contributions of age, nonverbal reasoning, vocabulary, and naming speed. Both PA and SSPA were sufficiently robust predictors of MSWR to make small but significant contributions beyond the four control variables. Once MA was included in the final model, however, only it and naming speed remained as significant predictors of MSWR.
Naming speed and MSWR. Naming speed’s emergence as a unique predictor of MSWR, in the context of a number of other variables, conforms to naming speed’s position as one of the “most universal, longitudinal, and concurrent predictors of reading ability” (Araújo, Reis, Petersson, & Faisca, 2015, p. 869). As Kim & Petscher exemplify when they state that “the role of [SSPA] in word reading primarily involves PA and MA,” as opposed to other variables like vocabulary (Kim & Petscher, 2015, p. 3), naming speed is rarely if ever included in studies involving SSPA. However, a recent meta-analysis by Araújo and colleagues found a persistent effect of naming speed over the course of reading development, although results varied somewhat, depending on type of reading assessment (Araújo, et al., 2015). Although some research has suggested that naming speed should not be a significant factor for readers in Grades 4 and 5 (e.g., Torgesen et al., 1997; Araújo, Inácio, Faisca, Petersson, & Reis, 2011), other work (e.g., Landerl & Wimmer, 2008; Ziegler et al., 2010) has shown a persistent and increasing developmental relationship between naming speed and reading ability over the elementary school years. The current study confirms that naming speed’s influence persists at least through Grade 5 and that it may be specifically important for multisyllabic words.

The three fundamental layers of information for reading are phonology, orthography, and semantics (Seidenberg & McClelland, 1985; Seidenberg, 2005). If we consider the ways in which the variables examined in the current study may be involved in readers’ neural networks, it is possible that PA and SSPA contribute at the phonological level while, for its part, MA may act as a “binding agent,” linking together phonology, orthography, and semantics to facilitate the integration of the mental representations of words (Kirby & Bowers, in press). Naming speed, on the other hand, may influence word reading at the level
of fluency, wherein readers’ application of knowledge becomes increasingly automatic, improving lexical quality (Kirby & Bowers, *in press*; Perfetti, 2007). Naming speed’s importance in the current study reminds us that word reading skill is not only a matter of accuracy, but also of fluency (or automaticity) (LaBerge & Samuels, 1974; Norton & Wolf, 2012; Vellutino et al., 2004). Both inaccuracies and disfluency can act as a bottleneck to the reading process, preventing comprehension (Perfetti & Lesgold, 1979).

**MA and vocabulary.** The current study contributes further evidence of the strong relationship between MA and vocabulary: MA’s addition to the model significantly reduced vocabulary’s contribution, likely due in part to how highly the two are correlated ($r = .769$). Such a strong association between these two constructs has been frequently reported in the literature (e.g., Kieffer & Lesaux, 2012; McBride-Chang, Wagner, et al., 2005), as high as .91 in Grade 4 (Wagner, Muse, & Tannenbaum, 2007). Some researchers have suggested that the two may develop in a reciprocal relationship (McBride-Chang, Wagner, et al., 2005; Muse, 2005; Stahl & Nagy, 2007), while others have postulated that, by the late elementary years, MA and vocabulary overlap (Wagner, Muse, & Tannenbaum, 2007; Wagner et al., 2007) or constitute one dimension (Spencer et al., 2015). More generally, a lack of direct relationship between vocabulary and word reading, after accounting for both PA and MA, converges with recent findings (e.g., Holliman et al., 2014b; Kim, Apel, & Al Otaiba, 2013).

**Limitations & Extensions for Future Work**

Our study was designed using typically achieving participants in Grades 4 and 5 with a focused list of multisyllabic words at or just above their grade level. A larger sample size, over a broader range of grades and reading abilities, would have given us the statistical power and sample diversity for a clearer picture of multisyllabic word reading. Because the
number of multisyllabic words in elementary school texts begins sharply increasing as early as Grade 3, it would have been ideal to include slightly younger participants. However, this would in turn create challenges for item selection on the MSWR task: finding appropriate, balanced items (e.g., number of syllables, lexical stress placement, morphological complexity, suffix type, sufficiently familiar, etc.) that are neither too difficult nor too simple to decode for all participants would be complicated.

**Longitudinal data.** Our study did not set out to examine the development of SSPA or of MSWR across elementary school; however, after concluding the current study and examining the literature closely, many of the questions that we have involve the need for longitudinal data. How do children’s SSPA skills influence MSWR? How does this relationship change over the elementary grades? Does PA’s influence on word reading wane if it is broadened to include suprasegmental phonology as well as segmental? How does MA’s role in MSWR change over the course of elementary school? How does the interplay of MA and vocabulary relate to MSWR? Is naming speed a consistent predictor across grades? And most importantly, as suggested by the current study’s results, what are the various interrelationships of these variables and how might they change as reading skills develop? These are questions that can only be answered with carefully designed, multi-measure longitudinal studies but that have the potential to contribute significantly to theory building and, eventually, intervention support for struggling readers who have particular deficits in MSWR.

**Mediation analysis.** The cross-sectional nature of our data precludes us from confidently stating that MA fully mediates the relationship between PA/SSPA and MSWR accuracy. Despite the statistical support of bootstrapping (a nonparametric resampling
procedure that reduces the possibility of Type I error and approximates the sampling
distribution of the indirect effect (Preacher & Hayes, 2008)), a more appropriate and valid
estimation of mediation would involve time precedence and a longitudinal design wherein
the cause (PA) is measured before the mediator (MA), which is in turn measured before the
outcome (MSWR), emphasizing the transmission of changes from cause to mediator to
outcome (Kline, 2015; Little, 2013). The current study’s results do suggest that such a study
is warranted. A further benefit of a longitudinal design is that it would allow us to tease
apart PA and SSPA’s relationship to reading, which has the potential to add significantly to
our understanding of the nuances of PA’s relationship to word reading across elementary
school as well as to the role SSPA plays at different stages of reading.

**Measurement.** The inclusion of additional measures of PA and SSPA may have
strengthened measurement to clarify their influence on MSWR; although two PA tasks
(Blending Words and Elision) were included in the current study, they did not correlate
strongly enough to justify creating a composite, potentially weakening PA’s predictive
ability. Likewise, multiple SSPA measures with improved reliability would have ameliorated
measurement of that construct.

In addition, Aural Stress Assignment’s reliability was relatively low ($\alpha = .64$) and
participants performed poorly on average ($M = 12.59/30$, 42% correct). This is in keeping
with the literature: it is not uncommon for tasks to have low to fair reliability (e.g., $\alpha = .37$ in
Holliman, Wood, & Sheehy, 2012; $\alpha = .47$ and .45 in Goswami, Gerson, & Astruc, 2010; $\alpha =$
.63 and .57 in Holliman et al., 2014a, 2014b, respectively) or for reliability scores to go
unreported (e.g., Whalley & Hansen, 2006; Gutierrez-Palma, Raya-Garcia, & Palma-Reyes,
2009; Clin, Wade-Woolley, & Heggie, 2009). Measures of SSPA are frequently experimental;
straightforward measures of SSPA (i.e., those that do not involve morphology or rely heavily on working memory) that are at an appropriate level of difficulty for this age group are less common. Aural Stress Assignment has been a reliable measure of adults’ SSPA (e.g., $\alpha = .81$, Chan & Wade-Woolley, 2016; $\alpha = .86$, Heggie & Wade-Woolley, submitted). The items used in the current study were selected to be at an appropriate level for middle school students, but further refinement of Aural Stress Assignment for this age group is necessary to create a more reliable measure of SSPA.

In future work, reflecting MA’s influence on MSWR, the items on the MSWR task should be selected with morphological complexity in mind. In the current study, they were neither all morphologically complex nor all morphologically simple — but neither were they evenly divided between the two. Indeed, depending on how morphological complexity is defined, several of our items are difficult to classify as one or the other. As Jarmulowicz notes, morphological complexity depends on one’s perspective. Lexicographers rely on etymology and historical morphology, examining the Germanic, Nordic, Latinate influences on English; this approach is the most complex, but it is also the most readily available in dictionary definitions. A psychological stance defines morphological complexity as the cognitive reality of morphemes, how they are stored and processed — living language over historical. A developmental perspective is primarily interested in the emergence and sequence of morphemes as they are organized in oral language, such as children’s sequential understanding inflectional and derivational morphology (L. Jarmulowicz, personal communication, April, 2014). We focused primarily on other factors such as stress pattern, frequency, and grade level (see Appendix B); a word list compiled based on morphological complexity, including attention to derivational suffixes given their
relationship with lexical stress patterns, would allow us to tease apart one of the potential reasons that readers’ MA was so strongly related to MSWR.

**Additional measures.** Due to time restrictions, we were unable to include all the potentially important constructs in our study. Notably, there was neither a familiarity rating for the items on the MSWR task nor a measure of participants’ orthographic processing. Although we carefully selected the items on the MSWR task such that they were at or just above participants’ grade level, this does not guarantee that individual students will know or even have seen all of these words. Some researchers have argued that subjective familiarity ratings or estimations of exposure provide a superior index to written frequency (e.g., Balota, Pilotti, & Cortese, 2001; Gernsbacher, 1984; Gilhooly & Logie, 1980).

Research has demonstrated that orthographic processing, or one’s knowledge and use of the rules and conventions concerning the visual and orthographic elements of print, is foundational to word reading (e.g., Cunningham, Perry, & Stanovich, 2001; Roman, Kirby, Parrila, Wade-Woolley & Deacon, 2009; Share, 1995, 1999) by supporting fast, fluent recognition of larger word units, especially those that do not correspond to grapheme-phoneme correspondence (GPC) rules (e.g., complex graphemes (th, ch), short morphemes (-ed, -ing)). Thus, it may be particularly important for multisyllabic word reading, wherein English spelling has preserved morphology at the expense of phonology, rendering GPC rules less consistent. Although it relies both on phonological information and on naming speed, orthographic processing has been shown to contribute uniquely to reading beyond these two constructs (Barker, Torgesen, & Wagner, 1992; Cunningham, Perry, & Stanovich, 2001; Cunningham & Stanovich, 1990; Stanovich & West, 1989) as well as consistently across grades (Roman et al., 2009; Share, 1995). Some researchers have argued that RAN
taps the skill that underlies orthographic processing (e.g., Manis, Doi, & Bhada, 2000; Wolf & Bowers, 1999); naming speed’s importance for MSWR in the current study suggests that a measure of orthographic processing would benefit future studies of multisyllabic word reading.

Conclusion

In sum, a number of variables work together to influence children’s multisyllabic word reading ability. The current study set out to investigate some of the metalinguistic skills that might be related to children’s multisyllabic word reading. Our results suggest that both PA and SSPA are important for accurate multisyllabic word reading, contributing indirectly through MA, with naming speed as an additional predictor. Multisyllabic word reading remains a relatively understudied area, perhaps because of the increased complexity of these items. Kearns and colleagues’ (2015, 2016) thorough and expansive studies highlight the myriad covariates that, with sufficient time and funding, may be included in studies of multisyllabic word reading — such breadth can be daunting and even unfeasible in terms of testing time and resources.

Our inclusion of SSPA as a predictor of MSWR constitutes a significant contribution to the more familiar interplay of variables (e.g., PA, MA) involved in word reading. Very few studies have examined SSPA’s specific role in MSWR, despite its frequent inclusion in theoretical explanations for SSPA’s importance for reading. SSPA is indeed important for MSWR; just how that relationship shifts and changes over reading development, across the various stages of reading development and in concert with other important covariates, remains to be seen.
CHAPTER THREE: Analysis of Error Productions by Grade 4 and 5 Students on Multisyllabic Words

Abstract

The majority of word reading tasks focus on accuracy as a dichotomous state. But because there is more than one way to be wrong, error analysis has the potential to enrich our understanding of word reading difficulties. The aim of the current study was to analyze the errors students made on a list of multisyllabic words at and above their grade level. Our study was guided by four research questions that examined: (a) the types of errors and whether they were related to reading ability; (b) how several child-level skills would be related to error types; (c) whether the incidence of lexicalization errors (i.e., real but erroneous words) would be related to reading ability; and (d) whether stress-shifting errors would conform to a pattern of regularization (shifts to the first syllable, in accordance with the majority of English trisyllabic words). English-speaking students in Grades 4 and 5 ($n=93$; Grade 4 $n=42$) were recruited for this study. Participants were administered measures of metalinguistic skills (segmental and suprasegmental phonological awareness, morphological awareness), control measures (vocabulary, naming speed), and word reading ability (Word Identification and an experimental measure of multisyllabic word reading). Good readers made proportionally fewer errors but were more likely to make errors involving shifts in lexical stress. Both morphological awareness and vocabulary were positively related to the incidence of lexicalized errors, while reading ability and suprasegmental phonological awareness were positively related to stress regularization errors.

Keywords: Multisyllabic; Word Reading; Error Analysis; Reading Ability
As was discussed in Chapter 1, multisyllabic words constitute the majority of the items students encounter in print, beginning in the elementary grades (e.g., Balota et al., 2007; Nagy & Anderson, 1984; Zeno, Ivens, Millard, & Duvvuri, 1995). Not only are multisyllabic words longer, they are more difficult to decode and can pose particular challenges for struggling readers (Archer, Gleason, & Vachon, 2003; Moats, 1998). This swift and simultaneous increase in frequency and difficulty means it is likely that children will make errors as they attempt to decode multisyllabic words, even while their reading skill is also improving. Poor readers, even when they are able to read monosyllabic words, often have a difficult time reading multisyllabic words (Just & Carpenter, 1987; Perfetti, 1986; Samuels, LaBerge, & Bremer, 1978).

The vast majority of word reading tasks score participant readings as either correct or incorrect. But as Jarmulowicz and Hay (2009) have noted, dichotomous scoring is limited: there are many ways to be incorrect, especially when both segmental and suprasegmental information is taken into account. In the current study we examine the types of errors made by Grade 4 and 5 students when reading multisyllabic words, to investigate connections between error type and other child-level factors (e.g., PA, SSPA, MA, reading ability).

We begin by briefly reviewing the history of error analysis, including miscue analysis and its limitations. We then overview some reading theories that lend themselves specifically to an examination of word reading errors. Finally, we review specific studies that applied error analysis to examinations of word reading and discuss the implications for the current study.
Error Analysis: The Broader Context

An examination of reading errors, in conjunction with the measurement of other reading-related skills (e.g., PA, MA), has the potential to provide valuable information about the strategies children apply, consciously or unconsciously, when attempting to decode multisyllabic words (e.g., Biemiller, 1970; Chiappe & Siegal, 1999; Stuart & Coltheart, 1988; Thomson, 1986). Children’s word reading error patterns provide rich information not only about the course of development in typically achieving readers but also about those who are struggling (e.g., Bruck, 1993; Liberman, Rubin, Duques, & Carlisle, 1985; Perin, 1998; Worthy & Viise, 1996). Observation of young children’s errors has also been foundational for theory building, constituting the starting point for the development of several major theories of literacy acquisition (Ehri, 1986; Frith, 1985; see Snowling, 2000). Broadly speaking, when individuals come across a word they do not know, they will either phonologically recode it (i.e., pronounce it as it is spelled) or they will misread it as another real word (Adams & Huggins, 1985). Word reading errors that are phonologically similar to target words are important, as they suggest the use of processes involving grapheme-phoneme correspondences, while those that are dissimilar suggest that other processes were involved (e.g., Chiappe & Siegal, 1999; Stuart & Coltheart, 1988).

Despite its potential utility, relatively few studies have examined word reading using error analysis; even fewer have specifically examined multisyllabic word reading errors. The close and detailed examination of children’s error productions in the current study constitutes a novel point of view on multisyllabic word reading. In this section we briefly discuss the origins of error analysis (miscue analysis of connected text reading), some
limitations to consider, as well as what we can learn about word reading through the application of error analysis.

**Miscue analysis of connected text reading.** Miscue analysis (Goodman, 1969; Goodman, Watson, & Burke, 2005) was first developed in response to evidence that children appeared to identify words more accurately in context than in isolation (Goodman, 1965). From this and other evidence involving children's reading errors, Goodman developed an influential approach that encouraged teachers to examine the errors their students made when reading continuous passages of text and to use that information to guide their reading instruction. Goodman divided errors — dubbing them “miscues” to avoid more pejorative terms such as “errors” and “mistakes” — into three “cueing” systems: syntactic, semantic, and orthographic. Miscues were thought to result when the child relied too strongly on one system or another (McKenna & Picard, 2006). For example, overreliance on syntactic cues may involve insertions or deletions (e.g., *He rode a big brown horse* instead of *He rode a brown horse*); overreliance on the semantic cueing system results in semantically-related errors (e.g., *pony* instead of *horse*); and an overreliance on the orthographic cueing system results in errors that may have nothing to do with the text’s meaning (e.g., *house* instead of *horse*).

**Limitations of miscue analysis.** Goodman’s reasoning placed these three cueing systems on equal footing. Counter to what we now understand about the foundational importance of phonological decoding skills, miscue analysis promoted the idea that contextual cues (semantic and syntactic) were equally as important as orthographic cues, as they were thought to emphasize the fact that children were reading for meaning. This lack of emphasis on phonological skills and decoding is a “major failing” of Goodman’s model.
(Rayner & Pollatsek, 1989, p. 351). Additionally, researchers have failed to replicate the findings of Goodman’s original study (see Nicholson, Lillas, & Rzoska, 1988). Critical reviews of miscue analysis have highlighted several of its shortcomings: a lack of consensus on the error categories, inadequate specification of error categories, lack of theoretical basis for the categories, failure to consider the effects of passage difficulty, and ambiguity regarding the causes of multiple-source errors (see Wixon, 1979; Leu, 1982 for review). Despite these weaknesses, miscues (as well as the Reading Miscue Inventory; Goodman & Burke, 1972) continue to be popular with educators today (e.g., Reading Recovery includes a three-way Venn diagram reminding teachers of the three systems; McKenna & Picard, 2006).

Two of the most serious concerns about miscue analysis are directly applicable to the broader use of error analysis: (a) the interpretation of children’s reading errors is necessarily subjective, reducing the reliability of error analysis as a method of assessment (Moats, 1999), and (b) reading aloud and reading silently do not necessarily draw upon identical cognitive processes and even if they do, it may be to different degrees (Singleton, 2005). Even after considering this additional complexity, oral reading is a pragmatic, efficient method for both researchers and teachers to gain information about students’ reading accuracy and fluency, especially given the foundational importance of phonological skills and decoding ability (National Reading Panel, 2000).

**Theories Underpinning Word Reading Errors**

In Chapter 1, Chall’s (1983, 1996) stages of reading development and Seidenberg & McClelland’s (1989) triangle model of reading were reviewed. An examination of error productions also benefits from these perspectives, in addition to the nuances provided by
two other theories: the Self-Teaching Hypothesis (Share, 1995, 1999) and Ehri’s (1991, 1994) discussion of how words are processed.

**Share’s Self-Teaching Hypothesis.** The Self-Teaching Hypothesis states that decoding acts as a self-teaching mechanism to build readers’ word recognition skill, and as such constitutes a way for children to contend with the thousands of new words introduced yearly (e.g., Archer, Gleason, & Vachon, 2003; Nagy & Anderson, 1984). Share argued that phonological decoding is the mechanism by which beginning readers not only build word-specific knowledge but more general orthographic knowledge: it is a necessary skill for reading unfamiliar words, even when readers have reached higher levels of proficiency (Rastle & Coltheart, 1998; Rey, Ziegler, & Jacobs, 2000). The studies of error productions during word reading (reviewed subsequently) underscore phonological decoding’s importance: children demonstrate their developing proficiency by making scaffolding errors (i.e., errors preserving both the initial and the final phoneme) on CVC words (Savage, Stuart, & Hill, 2001); adults with low literacy skills are less likely to use phonological decoding than children and non-native speakers of English, potentially sabotaging their own progress as readers (Davidson & Strucker, 2002; Greenberg, Ehri, & Perin, 2002); and children with different levels of phonological decoding ability make different patterns of errors (e.g., vowel pronunciation, prefix and suffix accuracy; Shefelbine & Calhoun, 1991).

**Beyond decoding: Processing words.** In terms of the application of strategies to word reading, Ehri identified four ways to read words: by decoding (letter by letter to begin, followed by larger letter units for those with more advanced decoding skill), by sight (retrieving words from long term memory), by analogy (applying information from known sight words to read new words) and by prediction (using context cues to predict words). As
readers attain skill, they learn to read isolated (i.e., without context) words in these first three ways (by decoding, sight, and analogy). These processes should each lead to different types of errors: decoding is slower and more laborious, and does not always work with English, especially with respect to multisyllabic words which involve more variable spellings and irregular pronunciations. With sufficient practice, all words attain sight word status — to be read quickly and automatically, as whole units. Cognitive energy is thus saved for reading comprehension. Words not read by sight require alphabetic knowledge (grapheme-phoneme correspondences). Finally, words can be read by analogy (e.g., reading *thump* using a known word such as *jump*). Reading words using context cues is inefficient, giving the reader only about a one in ten chance they will read it correctly (Nunes & Bryant, 2013); although the fourth and final strategy, by prediction, would also produce certain types of errors, it is irrelevant for a study involving reading words in isolation.

**Error Analysis in Action**

Error analysis of single word reading exemplifies the benefits of qualitative assessments of error type and rate. In this section, we discuss studies of word reading that have employed error analysis; they differ in their goals and participants, and often focus on monosyllabic words, but they provide guidance and considerations for the current study’s design and application of error analysis. Additionally, each study we discuss examined word reading using error analysis specifically to explore the relationship between error types and other reading-related skills, making them particularly relevant for the current study.

**Error analyses of single word reading.** In a longitudinal study of young children’s monosyllabic word reading, Savage, Stuart, and Hill (2001) aimed to determine which type
of error at age six would be related to overall reading ability two years later, as well as what reading abilities were correlated with “scaffolding errors,” (Error (e) below; Laxon, Masterson, & Moran, 1994). The authors presented participants with a list of 24 CVC words and categorized their errors into six subtypes, based on shared phonemes: (a) unrelated errors (no orthographic or phonological overlap; e.g., bean read as room); (b) errors with overlap (at least one letter from the target word, not necessarily in the same position (e.g., chin read as nip) or the same pronunciation (e.g., goat read as log); (c) errors preserving only the initial phoneme (e.g., rain read as road); (d) errors preserving only the final phoneme (e.g., lark read as bike); (e) errors preserving both the initial and the final phoneme (“scaffolding errors;” e.g., bead as bed); and (f) refusals (i.e., “I don’t know.”).

Scaffolding Errors surfaced as the only error class that was strongly positively associated with reading ability at age 8, predicting variance in word reading even above the influence of reading ability at age 6. It was a “necessary but not sufficient determiner,” however: few scaffolding errors co-occurred with later low reading scores, but many scaffolding errors occurred for both high and low scores in later word reading (Savage, Stuart, & Hill, 2001, p. 7). For those readers making few scaffolding errors — less than 25% of their total errors, eight of the 50 participants — Savage and colleagues found that two years later, these children were all about a year behind in reading ability. The use of error analysis in this study highlighted the fact scaffolding errors may be a desirable signal of reading development for this age group; most 6-year-olds should already be able to accurately decode both consonants in CVC words and be making errors of this type (as opposed to errors wherein only the initial or final consonant is decoded accurately). Few
errors of this type may signal the need for explicit instruction in consonant letter-to-sound knowledge.

Greenberg, Ehri, & Perin (2002) examined the word reading errors of low-literate adults and normally developing children who read at grade-equivalent levels (that is, at Grades 3, 4, and 5). Because research suggests that low-literate adults show relative strength in orthographic knowledge but a serious deficit in the phonological domain (e.g., Bruck, 1993; Fawcett & Nicholson, 1995), the authors applied error analysis in order to explore different cognitive-linguistic processes or strategies used by the two groups. Using a list of 50 atypically spelled words (e.g., sugar, aisle, ocean; Adams & Huggins, 1985), Greenberg and colleagues coded participants’ error productions as either real words (e.g., machine for mechanic) or nonwords (e.g., dente for deny), which they hypothesized provided evidence of orthographic or phonological decoding strategies, respectively. Nonword responses were then further analyzed to determine whether adults and children differed in the degree to which they used spelling-sound correspondence rules to produce acceptable decoding attempts (e.g., deef for deaf). Participants also completed the Word Attack subtest of the Woodcock Reading Mastery Test (Woodcock, 1998) and their responses were coded as either real or nonreal words.

Results of the word reading task suggested that adults were both less likely than children to use phonological decoding strategies and more likely to rely on visual or orthographic processes. Adults’ word reading errors were more likely to be real words that they guessed or accessed from memory based on partial letter cues, whereas children produced more mistakes attributable to phonological decoding, resulting in nonwords (Greenberg, Ehri, & Perin, 2002). Upon examination of the nonword reading task, adults
had also lexicalized items more often than the children, giving further evidence to the authors’ hypothesis that adults’ decoding skills were impaired. Inaccurate real word responses decreased with reading skill: proportionally fewer lexicalization errors occurred at higher word-reading grade levels for both groups (i.e. mean lexicalization error rate dropped from 42% of errors to 25% for adults and from 21% to 18% for children). Thus error analysis demonstrated that adult literacy students and children differ not only quantitatively (rate) but qualitatively (type) in their word reading errors, based on adults’ specific deficit with phonological decoding.

Davidson and Strucker (2002) used a similar, but slightly more detailed, coding scheme as Greenberg and colleagues’ to examine the decoding errors of two subgroups of adult basic education students: native speakers of English (NSE) and nonnative speakers of English3 (NNSE). They applied error analysis to determine in what ways the two groups would display different types of word reading errors. Participants read graded word lists (Diagnostic Assessments of Reading [DAR]; Roswell & Chall, 1991) until they reached ceiling on a list (less than 70% correct); the items on the last list read (10 items) were then analyzed and coded into one of four categories: 1) correct readings; 2) phonetically plausible substitutions, wherein all phonemes and syllables were pronounced in accordance with English phonetics (e.g., stress errors (e.g., soliTary, monoTONy) and vowel-sound substitutions (e.g., man-aige for manage)); 3) phonetically implausible substitutions (i.e., virtually all other errors types that did not result in real words; e.g., omission or addition of syllables or phonemes (e.g., imagative for imaginative), native language pronunciations for English cognates (e.g., French trahnkee for tranquil); and 4)

3 Authors only note that NNSEs were from “diverse linguistic and educational backgrounds.” Unfortunately, no specifics on participants’ first languages were provided.
substitutions of real English words, regardless of the number of syllables (e.g., collaborate for celebrate). Items not attempted by participants were not included in analyses.

The benefit of error analysis was evident in the parsing of word reading error types. Although group means and distributions on Word Attack were virtually identical, NSEs and NNSEs showed different patterns of errors: NSEs made more (2.1 times as many) lexicalized errors, while NNSEs made more (2.7 times as many) phonetically plausible errors. The majority of both groups’ errors (59.57% of NSE errors; 65.33% of NNSE errors) were classified as phonetically implausible. Both Davidson and Strucker’s and Greenberg and colleagues’ studies suggest that readers with phonological deficits make more lexicalized errors. Meanwhile, the NNSEs “resemble normally developing young readers” in that they do not have phonologically-based reading problems and use decoding as their preferred approach to word reading, particularly if they have skills to transfer from their native language (provided it was alphabetic) (Davidson & Strucker, 2002, p. 309).

Finally, based on previous work showing that approximately 15-20% of Grade 4 and 6 students had a great deal of difficulty with multisyllabic words (Shefelbine, 1990), Shefelbine & Calhoun (1991; in Cunningham, 1998) used a multisyllabic nonword reading task to classify Grade 6 students as high, moderate, or low decoders. Readers in each group were then given a list of real multisyllabic words to read aloud; error patterns were analyzed on the first ten words missed by each participant. High decoders pronounced vowels correctly 80% of the time, rarely omitted any phonemes, and had equal facility with beginning and ending syllables, accurately pronouncing prefixes and suffixes 75% of the time. Moderate decoders were equally successful with beginning syllables and prefixes (usually, but not always, the same thing) but read final syllables and suffixes correctly only
about half the time. Low decoders read beginning and ending syllables correctly one third of the time: they could read 75% of the prefixes but less than half of the suffixes.

Furthermore, both moderate and low decoders were 2-4 times more likely to omit syllables than the high decoder group. The authors described the process of developing proficiency in identifying multisyllabic words as “a vicious circle in which students need to read many polysyllabic words successfully to learn likely letter patterns but knowing these same patterns is necessary for reading the words in the first place” (Shefelbine & Calhoun, 1991, p. 176).

**Considerations When Applying Error Analysis**

A simultaneous strength and weakness of error analysis is its flexibility: depending on the goal of the study, including the participants and the items selected, error analysis permits the researcher to shine a light on a subset of a broad range of possible elements (e.g., decoding errors at various grain sizes (syllable, grapheme), insertions, deletions, stress placement, morphological analysis.). Error analysis has the potential to extend descriptions of group differences beyond the quantitative (error rate) to the qualitative (error type), highlighting the respective patterns of strengths and weaknesses (e.g., low literacy adults may have deficits in phonology but strengths in orthography) and describing the patterns of errors made by various groups (e.g., non-native speakers of English, children with more or less decoding ability). In the current study, we apply error analysis to elucidate some of the reasons why multisyllabic words pose particular challenges.

Before proceeding, however, it is important to discuss another challenge of error analysis: whether to apply it to participants’ raw number or proportion of errors. There does not appear to be a consensus in the literature; while some of the studies reviewed
above calculated the proportions for each error type of interest (Greenberg, Ehri, & Perin, 2002; Savage, Stuart, & Hill, 2001), others (Davidson & Strucker, 2002; Shefelbine, 1990; Shefelbine & Calhoun, 1991) used participants’ raw number of errors. Each method provides valuable information, but there are also downsides to each. When comparing error type X, for example, proportions equate the performance of a participant who made two errors (one of which was type X) with a participant who made fifty errors, twenty-five of which were type A: .50 on error type X. Raw numbers of errors limit group comparisons, especially when those groups are divided by reading ability: better readers make fewer errors by definition, and groups compared on reading ability (e.g., using quartiles) may not have the same number of participants. In order to make comparisons across groups and error types, we will use proportions of errors in the majority of our analyses.

**Lexical Stress in Trisyllabic Words**

Before turning to a description of the current study, it is necessary to provide some information as to the nature of the items selected for the multisyllabic word reading (MSWR) task. In our analysis of participants’ error productions, we focused on one aspect of multisyllabic words: the placement of primary stress, specifically relating to the statistical properties of lexical stress. All of the multisyllabic words in the current study were three syllables in length, and in English, the majority (about 60%) of trisyllabic words are stressed on the first syllable (Teschner & Whitely, 2004). In general, primary stress is typically either on the first or second syllable: which one depends on the nature of the third syllable (i.e., whether it is strong4 or weak) because long English words “generally follow an alternating stress rule (or “skip-a-syllable” rule)” wherein weak stress is assigned to every

---

4 The rule of thumb is that “all stressed syllables are strong, and all weak syllables are unstressed.” (Fear, Cutler, & Butterfield, 1994, p. 1893)
other syllable (Teschner & Whiteley, 2004, p. 18). Thus, if a trisyllabic word’s third syllable is strong, primary stress will fall on the first syllable; if it is weak, then the stress will either land on the second syllable (if that syllable is strong; e.g., *projectile*) or on the first (if the second syllable is weak; e.g., *Domicle*) (Fudge, 2015). Strong final syllables are identified orthographically by −CC5 (e.g., *asterisk*), -VV (e.g., *jubilee*), -VVC (e.g., *parakeet*), or −VCe (e.g., *antelope*) and most of the exceptions to the rules are for words with final syllable stress, which tend to have certain endings that attract stress (−een as in *smithereen*) (Fudge, 2015). Our hypotheses regarding lexical stress placement stem from these statistical properties of trisyllabic words.

**The Current Study**

Our analyses of students’ error productions on multisyllabic words is organized around four research questions. We began by asking what types of errors students in Grades 4 and 5 would make on a list of multisyllabic words that were at or just above their grade level and whether these error types would be related to reading ability. We devised and applied a six-point scheme to code error productions concurrently for decoding errors and for stress placement errors (see Table 3.1 below). After establishing a general picture of participants’ errors, we asked how five associated skills (morphological awareness, phonological awareness, prosodic awareness, vocabulary, and naming speed; see Chapter 2) would be related to the types of errors made by readers at various levels of reading ability. We hypothesized that good readers would make fewer errors overall, but more specifically that they would make fewer errors of each type than poorer readers. With respect to our second research question, we expected that children’s abilities on the five

---

5 Where C = consonant and V = vowel.
associated skills would be positively correlated with their ability to read words correctly and negatively correlated with the various error production subtypes in our scheme. However, due to the lack of previous research, we did not have a specific hypothesis as to how these five skills would be related to the individual error types.

Our third research question involved the extent to which participants’ error productions would consist of lexicalized errors, and whether this type of error would be related to participants’ reading ability status. True “lexicalization” errors are readings of nonwords that result in real English words (e.g., Treiman, Goswami, & Bruck, 1990); we have broadened the term to include errors resulting in erroneous, but real, words (e.g., reading origin as organ). Based on previous work suggesting that poor readers with phonological decoding difficulties make a higher rate of lexicalized errors (Greenberg, Ehri, & Perin, 2002; Davidson & Strucker, 2002), we hypothesized that participants’ phonological awareness would be negatively correlated with their proportion of real word error productions.

Finally, we asked whether error productions involving a shift in lexical stress would conform to a pattern of regularization, based on the fact that primary stress is most frequently on the first syllable of English trisyllabic words (Teschner & Whitely, 2004). We hypothesized that error productions involving a shift in primary stress would most frequently involve the relocation of second- or third-syllable stress to the item’s first syllable and that a) better readers and b) readers with better suprasegmental phonological awareness would be more likely to make this type of error (i.e., errors in accordance with the probabilistic cues of English).
Method

Students in Grades 4 and 5 from elementary schools in Ontario, Canada, participated in this study (n=93, 51 female; M age = 10 years, 5 months; n=42 in Grade 4). Participants were screened for language proficiency (only native English speakers were eligible to participate) and for cognitive/developmental typicality (students with developmental disabilities were excluded). Participants were not screened for reading disabilities or difficulties. Testing occurred in a quiet room in the student's school in the late spring of the school year. See Chapter 2 for more detailed descriptions of each task.

Morphological Awareness

Morphological awareness was measured using the Suffix Completion Task (Carlisle, 1988); in this expressive task, children are asked to produce the derived form of a word based on a given root, using carrier sentences as prompts (e.g., "The word is farm. My uncle is a ______.") ("Farmer"). Cronbach’s alpha was computed at .89 for the current sample.

Phonological Awareness

The Elision subtask of the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999) was used to measure participants' phonological awareness. Elision consists of 20 items and six practice items, beginning with compound words (e.g., toothbrush, airplane) through 2-syllable (e.g., spider, powder) and 1-syllable (e.g., cup, meet, snail) words. Participants must remove increasingly smaller parts of the words (word, syllable, phoneme) in increasingly difficult locations (initial, final, medial positions). Cronbach’s alpha was reported in the manual at .91 for this age group.
Suprasegmental Phonological Awareness

Aural Stress Assignment (ASA; Wade-Woolley, 2016) measures participants’ prosodic awareness of spoken words. The task consists of 32 multisyllabic words between two and five syllables in length. Participants are asked to locate the syllable carrying the word’s primary stress (e.g., opportunity's stress is on the third syllable, o-ppor-TU-ni-ty). Participants are asked to indicate which syllable has the strongest beat. ASA consists of two practice items and 30 test items; participants worked with the two practice items, and other nouns such as their first name, until they report understanding the task. Cronbach’s alpha was computed at .64 for this sample.

Vocabulary

Participants’ expressive vocabulary was measured using the Vocabulary subtest of the Wechsler Abbreviated Scale of Intelligence (WASI Vocab; Wechsler, 1999). This task consists of 42 items: four pictures followed by 38 words to be orally defined. Responses are awarded 2 (e.g., a good synonym, correct figurative use), 1 (e.g., vague or less pertinent synonym, example using the word itself), or 0 (e.g., no real understanding shown, poverty of content) points. The sum of the items correctly defined constitutes the total raw score on WASI Vocab. Reliability is reported in the manual at .92 for this age group.

Naming Speed

The Digits subtest of Rapid Automatized Naming (RAN Numbers; Wagner, Torgesen, & Rashotte, 1999) was included to control for naming speed. This task consists of two arrays of single digit numbers (2, 6, 9, 4, 7), arranged in four rows of 10 digits per page (Form A, Form B). Participants are asked to name the digits as quickly and accurately as possible, and the time taken to name all the digits on a page is recorded. Both Form A and Form B are
administered unless the participant makes more than four errors on Form A. Alternate-form reliability is reported in the manual at .90 for this age group.

**Reading Measures**

**Word reading.** Participants’ word reading was tested using the Word Identification (Word ID) subtest of the Woodcock Reading Mastery Tests (Woodcock, 1998). Word ID consists of 106 English words of increasing difficulty; participants begin at an entry point dependent on their age (e.g., students in Grade 5 begin at Item #56). Items are scored as correct or incorrect, and testing continues until the participant makes six consecutive errors. Split half reliability is reported in the manual at .91 for this age group.

**Multisyllabic word reading (MSWR).** Participants were presented with a list of 54 trisyllabic words at or slightly above their grade level (Zeno et al., 1995), compiled to balance stress pattern, standard frequency index (SFI), and word length (number of letters and number of phonemes) (Balota et al., 2007; Coltheart, 1981; Zeno et al., 1995). The complete list of items is in Appendix B, including selection criteria. The items were presented in a fixed random order, divided over three sessions to avoid tiring participants. Children were instructed to read each item aloud, giving it their best try, even if they were unsure of its pronunciation. Participants’ readings were digitally recorded using a Sony Digital Voice Recorder and uploaded for later scoring.

Upon review of the recorded responses, participants’ responses were coded according to the six-point scheme (see Table 3.1), roughly transcribing all error productions (e.g., “PE-ril-mint” for parliament). Categories were considered mutually exclusive and each participant’s reading was awarded a single code. Note that for each error type within the scheme, an error was considered a deviation from the correct pronunciation, regardless of
plausibility. Inter-rater reliability was calculated at .87 using a second rater with training in speech-language pathology and phonetic transcription, using a random selection (10%) of participant error productions.

Table 3.1

Scoring Scheme For Multisyllabic Word Reading Task

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>Example(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓</td>
<td>Correct stress, accurate decoding</td>
<td>CON-se-quence</td>
</tr>
<tr>
<td>A</td>
<td>Correct stress, decoding error</td>
<td>CON-se-kence</td>
</tr>
<tr>
<td>B</td>
<td>Incorrect stress, accurate decoding</td>
<td>con-SEE-quence</td>
</tr>
<tr>
<td>C</td>
<td>Incorrect stress, decoding error</td>
<td>con-SE-kence</td>
</tr>
<tr>
<td>D</td>
<td>Equal stress</td>
<td>con... se... quence...</td>
</tr>
<tr>
<td>E</td>
<td>Incorrect number of syllables (+/-)</td>
<td>cons-kence, con-se-cue-ence</td>
</tr>
</tbody>
</table>

Note. Letters separated by dashes denote syllables. An ellipsis represents a pause.

Results

Descriptive statistics for raw scores on all measures are reported in Table 3.2, including skewness and kurtosis statistics. We began by characterizing the sample; participants’ mean Standard Score on Word Identification ($M = 99.97$) and mean T-scores on WASI Vocabulary ($M = 47.74$) were both similar to the norming samples. Outliers were defined as scores 2.5 SDs above or below the mean for each variable; four outliers were identified on two tasks (one on RAN Numbers and three on Aural Stress Assignment) and these scores were Winsorized to minimize data loss. Only Elision was significantly skewed (Tabachnick & Fidell, 1996); scores were corrected using a reflected square root transformation. Raw scores are reported in Table 3.2, but transformed scores were used for Elision in all analyses. Proportions follow in Table 3.3.
Table 3.2

Descriptive Statistics (n=93)

<table>
<thead>
<tr>
<th>Test Description</th>
<th>Min</th>
<th>Max</th>
<th>M</th>
<th>SD</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suffix Completion (/44)</td>
<td>7</td>
<td>40</td>
<td>25.92</td>
<td>7.38</td>
<td>-.401</td>
<td>-.237</td>
</tr>
<tr>
<td>Elision (/20)</td>
<td>4</td>
<td>20</td>
<td>14.77</td>
<td>4.49</td>
<td>-.927</td>
<td>-.560</td>
</tr>
<tr>
<td>Aural Stress Assignment (/30)</td>
<td>5</td>
<td>23</td>
<td>12.59</td>
<td>4.22</td>
<td>.624</td>
<td>-.115</td>
</tr>
<tr>
<td>WASI Vocabulary (/64)</td>
<td>17</td>
<td>56</td>
<td>36.40</td>
<td>9.59</td>
<td>-.012</td>
<td>-.861</td>
</tr>
<tr>
<td>RAN Numbers</td>
<td>21</td>
<td>54</td>
<td>34.22</td>
<td>7.00</td>
<td>.685</td>
<td>.280</td>
</tr>
<tr>
<td>Word ID (/106)</td>
<td>38</td>
<td>94</td>
<td>69.58</td>
<td>12.01</td>
<td>-.374</td>
<td>-.458</td>
</tr>
<tr>
<td>Multisyllabic Word Reading (/54)</td>
<td>0</td>
<td>54</td>
<td>33.63</td>
<td>13.81</td>
<td>-.516</td>
<td>-.696</td>
</tr>
</tbody>
</table>

Note. Standard Error of Skewness = .250; Standard Error of Kurtosis = .495

Error Productions in General and by Reading Ability

We first asked what types of errors students would make on a list of multisyllabic words and whether error type would be related to reading ability. A descriptive summary of the participants' readings (frequency) and as proportions (i.e., # errors of each type / # total errors made for each participant) is in Table 3.3. Note that while Maximum scores are included in the table, Minimum scores were omitted because in each case, the value was zero.
Table 3.3

*Distribution of Readings Across Codes for MSWR (54 items)*

<table>
<thead>
<tr>
<th></th>
<th>Max</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#</td>
<td>Prop (SD)</td>
</tr>
<tr>
<td>✓ Correct Readings (No Error)</td>
<td>54</td>
<td>1.00</td>
</tr>
<tr>
<td>Error Codes:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A Correct Stress, Decoding Error</td>
<td>19</td>
<td>.71</td>
</tr>
<tr>
<td>B Stress Error, Correct Decoding</td>
<td>5</td>
<td>.50</td>
</tr>
<tr>
<td>C Stress Error, Decoding Error</td>
<td>12</td>
<td>.50</td>
</tr>
<tr>
<td>D Equal Stress</td>
<td>12</td>
<td>.25</td>
</tr>
<tr>
<td>E Incorrect Number of Syllables</td>
<td>28</td>
<td>.75</td>
</tr>
<tr>
<td>Errors Total</td>
<td>54</td>
<td>1.00</td>
</tr>
</tbody>
</table>

*Note.* Errors are listed as mean raw number (#) and as proportions *(prop).* Mean proportions do not sum to 1.0 due to uncategorized errors (i.e., non-responses, n=12).

For most error codes, the standard deviations are relatively high, indicating a broad range in instances of errors across participants. To examine these individual differences, we began by grouping relatively stronger and weaker readers by transforming participants’ raw scores on Word ID into z-scores to group them into reading ability quartiles. Z-scores ranged from -2.629 to 2.033; the Lower Quartile (-.714), Second Quartile (.118), and Upper Quartile (.701) cutoffs were calculated to determine the four groups based on reading ability. See Table 3.4 for participants’ mean number and proportion of errors by reading quartile.
Table 3.4

Proportion of Errors Made by Reading Quartile

<table>
<thead>
<tr>
<th></th>
<th>1st Quartile (Q1)</th>
<th>2nd Quartile (Q2)</th>
<th>3rd Quartile (Q3)</th>
<th>4th Quartile (Q4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n = 21)</td>
<td>(n = 25)</td>
<td>(n = 21)</td>
<td>(n = 26)</td>
</tr>
<tr>
<td><strong>M (SD)</strong></td>
<td><strong>M prop</strong></td>
<td><strong>M (SD)</strong></td>
<td><strong>M prop</strong></td>
<td><strong>M (SD)</strong></td>
</tr>
<tr>
<td>A Decoding Errors</td>
<td>11.00 (3.46)</td>
<td>.284</td>
<td>10.12 (4.13)</td>
<td>.406</td>
</tr>
<tr>
<td>B Stress Errors</td>
<td>.67 (.91)</td>
<td>.019</td>
<td>0.24 (.44)</td>
<td>.008</td>
</tr>
<tr>
<td>C Stress &amp; Decoding Errors</td>
<td>5.38 (2.99)</td>
<td>.135</td>
<td>3.40 (2.24)</td>
<td>.131</td>
</tr>
<tr>
<td>D Equal Stress</td>
<td>4.14 (3.61)</td>
<td>.096</td>
<td>1.20 (1.41)</td>
<td>.053</td>
</tr>
<tr>
<td>E Wrong # Syllables</td>
<td>16.90 (5.18)</td>
<td>.416</td>
<td>10.20 (4.43)</td>
<td>.403</td>
</tr>
<tr>
<td>Errors Overall</td>
<td>39.29 (7.64)</td>
<td>.728</td>
<td>25.16 (6.71)</td>
<td>.466</td>
</tr>
</tbody>
</table>

**Note.** Mean (i.e., sum of errors, by type), mean proportion (e.g., sum B / sum of errors overall). Proportions appear in bold to improve the readability of the table.

For additional clarity, the distribution of errors made by participants in each reading quartile, shown as percentages of the total number of errors, is presented graphically in Figure 3.1. Note that the information presented here corresponds to the mean proportions by quartile in Table 3.4.
The total number of errors made by participants in each group decreases with quartiles’ increasing reading skill: Q1 made a total of 825 errors, Q2 made 629 errors, Q3 made 281 errors, and Q4 made 159 errors overall. However, each quartile group has a different number of members; thus, mean error proportion was also calculated for each group to confirm this pattern: Q1 $M = .73$, Q2 $M = .47$, Q3 $M = .25$, and Q4 $M = .11$. Thus the first part of our hypothesis is confirmed: good readers make fewer errors overall. The second part of our initial research question addressed error types and reading ability; as we hypothesized, good readers made fewer errors of some types (i.e., Equal Stress, Incorrect Number of Syllables; see Figure 3.2). However, good readers (both Q3 and Q4) were more likely to
make errors involving a stress shift: (a) Stress Only errors \( (M = .095 \text{ (Q3)} \) and \( .087 \text{ (Q4)}, \) while \( M = .019 \text{ (Q1)} \) and \( .008 \text{ (Q2)}); and (b) Stress & Decoding errors\(^6\) \( (M = .174 \text{ (Q3)} \) and \( .209 \text{ (Q4)} \) while \( M = .135 \text{ (Q1)} \) and \( .131 \text{ (Q2)}).\)

![Graph of Error Types by Reading Quartile](image)

**Figure 3.2.** Mean proportion of error types by reading quartile (Q1 – Q4).

To determine whether the types of errors children made were in fact related to their reading status, we ran a 4 (reading quartile) x 5 (error type) repeated measures ANOVA on participants' proportion of errors by type. Mauchly's test indicated that the assumption of sphericity had been violated \( (\chi^2(9) = 77.95, p < .001) \), and so the degrees of freedom were corrected using Huynh-Feldt estimates of sphericity \( (\epsilon = .75) \). A main effect of error type \( (F(2.48) = 153.77, p < .001) \) indicated that the proportion of errors of each type varied

\(^6\) Trending at \( p < .053.\)
significantly; a significant interaction between error type and reading quartile ($F(7.43) = 3.82, p < .001$) indicated that participants tended to commit different types of errors based on reading quartile membership (see Figure 3.2).

As Figure 3.2 illustrates, there were several potential sources of the significant interaction. We ran a series of one-way ANOVAs to compare the groups’ proportion of errors by type. There were significant differences on all types of errors except Stress & Decoding Errors (trending at $F(3, 89) = 2.66, p = .053$). An examination of mean differences indicated that a number of group differences were responsible for these effects; the four groups have been organized by lower (Q1 and Q2) and higher (Q3 and Q4) reading quartiles (see Table 3.5).

Table 3.5

Mean Differences Between Reading Quartiles on Error Proportions

<table>
<thead>
<tr>
<th>Error Type</th>
<th>Lower Reading</th>
<th>Higher Reading</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Decoding Errors</td>
<td>Q1</td>
<td>Q3</td>
<td>-.146*</td>
</tr>
<tr>
<td>B Stress Errors</td>
<td>Q2</td>
<td>Q3, Q4</td>
<td>-.087*, -.079*</td>
</tr>
<tr>
<td>C Stress &amp; Decoding Errors</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>D Equal Stress</td>
<td>Q1</td>
<td>Q3, Q4</td>
<td>.077**, .081**</td>
</tr>
<tr>
<td>E Incorrect Number of Syllables</td>
<td>Q1</td>
<td>Q3, Q4</td>
<td>.133*, .216**</td>
</tr>
</tbody>
</table>

Note. For example, there is a significant difference between Q1 and Q3 as well as Q1 and Q4 on Equal Stress errors. ** $p < .01$ and * $p < .05$.

These results indicate that there was a significant difference between poorer and better readers on most types of errors: good readers made significantly fewer errors read with equal stress and errors involving an incorrect number of syllables, but significantly more decoding errors and stress misplacement errors.
**Error Productions and Associated Skills**

In our second research question, we asked how five associated skills (MA, PA, SSPA, vocabulary, and naming speed) would be related to the various types of errors made by readers at various levels of reading ability. We began by determining the correlations between both the mean raw number and the mean proportion of errors of each type (below and above the diagonal; see Table 3.6). Using raw scores, participants’ correct responses were significantly negatively correlated with all types of errors except Stress Errors (Code B). All remaining error types were correlated except for Stress Errors, which was not significantly correlated with any other error type. Error proportions displayed a slightly different pattern: participants’ correct responses were positively correlated with Stress Errors, but negatively correlated with Equal Stress and Number of Syllable errors. Error types were intermittently correlated (see Table 3.6).

Table 3.6

*Correlations Between Error Types (n=93)*

<table>
<thead>
<tr>
<th>Correct Readings</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓ Correct Readings</td>
<td>—</td>
<td>.191</td>
<td>.281**</td>
<td>.158</td>
<td>-.447***</td>
</tr>
<tr>
<td>A Decoding Errors</td>
<td>-.748***</td>
<td>—</td>
<td>-.121</td>
<td>.167</td>
<td>-.333**</td>
</tr>
<tr>
<td>B Stress Errors</td>
<td>-.015</td>
<td>.034</td>
<td>—</td>
<td>.025</td>
<td>-.164</td>
</tr>
<tr>
<td>C Stress &amp; Decoding Errors</td>
<td>-.679***</td>
<td>.601***</td>
<td>.005</td>
<td>—</td>
<td>-.308**</td>
</tr>
<tr>
<td>D Equal Stress</td>
<td>-.680***</td>
<td>.258*</td>
<td>-.001</td>
<td>.296**</td>
<td>—</td>
</tr>
<tr>
<td>E Wrong # Syllables</td>
<td>-.907***</td>
<td>.634***</td>
<td>.061</td>
<td>.579***</td>
<td>.465***</td>
</tr>
</tbody>
</table>

*Note.* Raw number of errors appears below the diagonal; proportion of errors appears above. Shading has been added to increase table readability.

Correlations between error production types and child-level skills (vocabulary, MA, PA, SSPA, and naming speed) appear in Table 3.7. As we saw in Chapter 2, all five of the child-
level skills were significantly related to participants’ ability to correctly read multisyllabic words. Herein, we focus on the relationship between each of these skills and the various error types.

Table 3.7

Correlations Between Child-Level Factors and Number of Errors of Each Type

<table>
<thead>
<tr>
<th></th>
<th>MA</th>
<th>PA</th>
<th>SSPA</th>
<th>Vocabulary</th>
<th>RAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓ Correct Readings</td>
<td>.736***</td>
<td>.466***</td>
<td>.318**</td>
<td>.611***</td>
<td>-.429***</td>
</tr>
<tr>
<td>A Decoding Errors</td>
<td>-.582***</td>
<td>-.322**</td>
<td>-.319**</td>
<td>-.550***</td>
<td>.285**</td>
</tr>
<tr>
<td>B Stress Errors</td>
<td>-.028</td>
<td>.120</td>
<td>-.040</td>
<td>.057</td>
<td>-.010</td>
</tr>
<tr>
<td>C Stress &amp; Decoding Errors</td>
<td>-.555***</td>
<td>-.354***</td>
<td>-.152</td>
<td>-.509***</td>
<td>.254*</td>
</tr>
<tr>
<td>D Equal Stress</td>
<td>-.458***</td>
<td>-.283**</td>
<td>-.203†</td>
<td>-.272**</td>
<td>.351**</td>
</tr>
<tr>
<td>E Wrong # Syllables</td>
<td>-.646***</td>
<td>-.436***</td>
<td>-.283**</td>
<td>-.510***</td>
<td>.344**</td>
</tr>
</tbody>
</table>

Note. *** p < .001, ** p < .01, * p < .05. The correlation between Code D and SSPA (†) p = .051

The most striking pattern evident in Table 3.7 is that none of the child-level skills (MA, PA, SSPA, vocabulary, or naming speed) were related to the incidence of Stress Only error productions (Code B). This may have to do with the fact that so few errors of this type were made: only 57 error productions were coded as B (3% of the error productions overall). (As opposed to Decoding Errors (A) and Incorrect Number of Syllables (E), which were the most frequently coded error types (661 (35%) and n=720 (38%), respectively).) Upon examination of those participants who made Stress Only errors (n=38), most made only one or two errors of this type each. Three participants made slightly more: one Grade 4 student made five, and two Grade 5 students made three each. Stress Only errors were made on 11 unique items across participants, 82% of which should have been stressed on the first syllable; otherwise, the stress was shifted to the first syllable (see Table 3.8).
Table 3.8

Participants With the Most Stress Only Errors and Corresponding Items

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>Item</th>
<th>Pronunciation</th>
<th>Item’s Correct Primary Stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>47004</td>
<td>candidate</td>
<td>can-DI-date</td>
<td>1</td>
</tr>
<tr>
<td>(Q3)</td>
<td>compliment</td>
<td>com-PLEH-ment</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>enterprise</td>
<td>en-ter-PRIZE</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>estimate</td>
<td>es-tuh-MATE</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>tendency</td>
<td>ten-DEN-cy</td>
<td>1</td>
</tr>
<tr>
<td>52018</td>
<td>consequence</td>
<td>con-SEE-quence</td>
<td>1</td>
</tr>
<tr>
<td>(Q3)</td>
<td>physician</td>
<td>PHY-si-cian</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>resentment</td>
<td>RE-sent-ment</td>
<td>2</td>
</tr>
<tr>
<td>53002</td>
<td>origin</td>
<td>or-IH-jin</td>
<td>1</td>
</tr>
<tr>
<td>(Q1)</td>
<td>parliament</td>
<td>par-LEM-ent</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>primary</td>
<td>pri-MA-ry</td>
<td>1</td>
</tr>
</tbody>
</table>

Participants’ MA PA, vocabulary, and naming speed were related to all subtypes of errors (except Stress Only errors), with MA and vocabulary generally representing the strongest relationships for each error type. SSPA, on the other hand, was related only to Decoding Errors and to errors resulting in the Incorrect Number of Syllables; SSPA was not related to errors involving stress placement (Stress Errors, Stress & Decoding Errors) or to errors resulting in Equal Stress. This latter category was included to catch those errors involving slow, drawn-out readings of words that sounded almost robotic, with a lack of fluency and obvious primary stress on any one syllable (e.g., reading proportion as “praw... per... shun”). The error productions included as Equal Stress errors often involved more laborious decodings, phoneme-by-phoneme, suggesting poor phonological decoding ability, specifically poor chunking skills (e.g., reading announcement as “an... ah... ow... n... k... em... nt”).

75
Lexicalized Error Productions

For our third research question, we asked to what extent participants’ error productions would consist of lexicalized errors (that is, errors resulting in real English words) and whether this type of error would be related to participants’ reading ability. Participants’ accurate readings of words (n=3070 productions) were set aside, leaving only error productions (n=1889). Each of these error productions was coded either as a lexicalized error (e.g., reading specialist as selection) or as a nonlexicalized error (e.g., reading discussion as di-se-men). Only single word responses were considered real words; two-word responses (e.g., “fizz cone” for physician) were coded as nonwords as it would require a subjective judgement to determine whether the production’s status as a real word was an intentional or coincidental compound. Of the 1889 error productions, 399 (21.1%) were lexicalized. Grade 4 students made 191 (17.2% of 1109 errors total) and Grade 5 students made 208 lexicalized error productions (26.7% of 780 errors); using percentage of overall errors made, Grade 5s’ error productions were real words significantly more often than were Grade 4s’ $t(1887) = -4.981, p < .001$.

To determine whether there was a relationship between proportion of lexicalized errors and reading ability, we first calculated the total number of errors made by participants in each quartile (see Figure 3.3).
Figure 3.3. Number of error productions by participants in each quartile.

Based on these error productions, Figure 3.4 illustrates the percentage of lexicalized error productions for each quartile (e.g., Q1 made 857 errors, of which 18.6\% were lexicalized).

Figure 3.4. Percentage lexicalized error productions by quartile. Note that the total number of error productions differs by quartile (see Figure 3).

The proportion of lexicalized error productions was not significantly correlated with reading quartile ($r = .069, p = .515$), confirming the similarity in percentages of lexicalized
errors across quartiles in Figure 3.4. Of the associated skills, only MA ($r = .214, p < .05$) and vocabulary ($r = .220, p < .05$) were significantly correlated with participants’ proportion of lexicalized errors.

**Regularizing Trisyllabic Stress**

In our final research question, we asked whether error productions involving a shift in lexical stress would conform to a pattern of regularization based on the statistical properties of English trisyllabic words. We began by characterizing participants’ responses in terms of length (number of syllables), by reading ability quartile (see Figure 3.5).

*Figure 3.5. Percentage of error productions by length (number of syllables). Note that the total number of error productions differs by quartile (see Figure 3.4). Some participants in Q1 and Q2 made responses coded as “0-syllable” (e.g., phoneme-by-phoneme decodings), resulting in a sum < 100%.*
An examination of lexical stress shifts on trisyllabic words necessarily relies on error productions accurately reproducing the number of syllables in the target items; thus, only trisyllabic error productions were included in the current analyses (n = 1081, 57.2% of error productions total). Of these, error productions involving a shift in lexical stress placement (n=333, 17.6% of error productions total) were isolated and coded for stress location.

Error productions involving a lexical stress shift included three distinct subtypes: (a) real trisyllabic words with a stress pattern that differs from the target word (e.g., reading cabinet as “clarinet”), (b) nonwords with a different stress pattern than the target word (that is, lexical stress and decoding error(s); e.g., reading parliament as “per-LENT-mint”), and (c) approximations of the target word with lexical stress shifted (that is, lack of decoding error(s) save necessary vowel alterations to accommodate stress shift; e.g., reading consequence as “con-SEE-quence”). Of the 333 error productions involving a shift in primary stress, 44 (13.2%) resulted in real words, 276 (82.9%) included both stress and decoding errors, and 57 (17.1%) were purely stress shifting errors. Because they have previously been discussed, real word productions were removed from analyses (289 error productions remaining; see Figure 3.6).
Figure 3.6. Error productions involving a shift in primary stress. Only nonlexicalized errors were included in stress shift analyses.

Our investigation of lexical stress regularization focused on the fact that approximately 60% of English trisyllabic words are stressed on the first syllable (Teschner & Whitely, 2004). (Note that the items on the MSWR task were originally selected to strike a balance in primary stress location (i.e., primary stress falls on either the first (48%) or second (50%) syllable stress in the vast majority of items) rather than to reflect this statistical distribution of trisyllabic words’ primary stress.) Most (51.9%) error productions involved shifts to the first syllable (e.g., PHYSician); the remainder of the shifts moved primary stress to the second syllable (36.3%; e.g., reading compliment as “comPLETEment”) or to the third syllable (11.8%; e.g., enterPRISE). Of those error productions in which primary stress was placed on the first syllable (n=187), 98.9% were on items that originally had second syllable stress (e.g., depression, resentment) and 80.2% involved both a stress and a decoding error (e.g., “CAWmintee” for committee).
Figure 3.7. Location of primary stress on three-syllable error productions involving stress shifts, by reading quartile. Raw number of error productions in each quartile: Q1 (n=108), Q2 (n=74), Q3 (n=68), and Q4 (n=38).

Figure 3.7 shows that the percentage of participants’ errors involving shifts accordant with the rules of English (i.e., to a trisyllabic word’s first syllable) steadily and significantly increases (from 39.8% in Q1 to 86.8% in Q4), \(F(3, 329) = 7.93, p < .001\). (Note as well that the raw number of stress-shifting error productions decreases, from 108 in Q1 to 38 in Q4). Of the child-level skills measured in the current study, the proportion of participants’ stress shifting errors that regularized to the first syllable was significantly correlated with their SSPA\(^7\) \((r = .264, p < .05)\).

---

\(^7\) Remainder of correlations were not significant: MA \(r = .064, p = .555\); PA \(r = -.120, p = .267\); vocabulary \(r = .132, p = .223\); and naming speed \(r = -.133, p = .298\).
Discussion

We set out to gain more information about the kinds of errors children make when reading multisyllabic words at or just above their grade level, as well as how several child-level skills might be related to these errors. The application of error analysis to children’s word reading provides the opportunity for a more qualitative, in depth analysis of an outcome typically measured dichotomously. However, error analysis also presents a descriptive challenge because error categories are not necessarily mutually exclusive: a participant’s misreading of a word may contain a number of different, overlapping errors (Jarmulowicz & Hay, 2009). We founded our analyses on four research questions, discussed below in turn.

We began by coding and examining the types of errors students made as they read multisyllabic words, and whether these error types were related to reading ability. We designed a coding scheme that concurrently measured phonological decoding and suprasegmental phonological errors, based on research showing that typically-achieving children were likely to rely on phonological decoding strategies (Greenberg, Ehri, & Perin, 2002), as well as on the notion that phonological decoding is a necessary skill, even for readers at advanced levels of reading ability, because we never fully cease to encounter unfamiliar words (Rastle & Coltheart, 1998; Rey, Ziegler, & Jacobs, 2000; Share, 1995, 1999). Results indicate that readers of different skill levels tended to commit different types of errors; specifically, good readers were more likely to commit decoding errors and stress misplacement errors on multisyllabic words, but pronounced them with equal stress across syllables or with an incorrect number of syllables significantly less often than poorer readers. Thus the better the reader, the finer the grain size of their error: decoding errors
were often close approximations of the target (e.g., “a-men-dent” for *amendment*), as were stress placement errors (e.g., “CAW-mi-ttee” for *committee*), while equal stress and incorrect number of syllable errors frequently reflected more basic reading errors (e.g., “seat-see-een-t-em-en-k” for *sentiment*; “ca-nay-dee-ant” for *candidate*).

Next, we asked how five associated skills (MA, PA, SSPA, vocabulary, and naming speed) would be related to the types of errors. We found that four of these five skills were important for (nearly) all error types: MA, PA, vocabulary, and naming speed were related to all subtypes of errors except Stress Only errors, an error type that was related neither to the child-level skills nor to any other error type. Only one child-level skill exhibited a different pattern; SSPA was related to decoding errors and errors resulting in an incorrect number of syllables. SSPA was *not* related to any error subtype that explicitly involved stress misplacement or a lack of stress assignment. It was, however, significantly related to correct readings of multisyllabic words, confirming previous research findings (e.g., Goswami et al., 2002; Holliman, Wood, & Sheehy, 2008). If it were the case that its more distal relationship to MSWR (i.e., mediated by MA; see Chapter 2) was responsible for this pattern, we would also have expected to find that PA was not related to all types of errors. It may have been that SSPA’s relationship to word reading errors was not sufficiently strong to reach significance; both decoding and syllable addition/omission errors were the most frequent error types across participants, potentially affecting power when using raw numbers of errors in analyses (e.g., rarity of Stress Only errors).

For our third research question, we asked to what extent participants’ error productions would consist of lexicalized errors (that is, errors resulting in real English words) and whether this type of error would be related to participants’ reading ability.
About a quarter of children’s error productions were lexicalized, relatively evenly distributed across reading quartiles such that reading ability was not significantly related to the incidence of lexicalization. This is mostly in keeping with previous research showing that reading ability alone is not a sufficient predictor of lexicalization errors: children make fewer lexicalized errors than adult basic education students matched on reading ability (Greenberg, Ehri, & Perin, 2002), and English language learners, described as resembling typically developing young readers, make half as many lexicalized errors as Adult Basic Education students matched on reading ability (Davidson & Strucker, 2002). Counter to our expectations, PA was not related to this type of error; instead, MA and vocabulary exhibited low but significant correlations with the proportion of participants’ lexicalization errors. Both relationships were positive, indicating that the better readers’ MA and vocabulary, the more likely they were to make lexicalization errors. Misreading words as visually similar items may reflect readers’ tendency to examine words as whole units rather than break them apart; that is, this type of error may indicate a stronger reliance on lexical, rather than sublexical, processing (e.g., Nation, 2014; Stuart, Masterson, Dixon, & Quinlan, 1999). For example, previous research has shown that children make more lexicalization errors on nonword items with many orthographic neighbours (e.g., Treiman, Goswami, & Bruck, 1990). This could explain the connection to both vocabulary and MA; if readers are not paying close attention or decoding carefully, they may misconstrue morphological and vocabulary-based cues, producing a near miss (e.g., “attitude” for altitude; “expectation” for exception; “possession” for procession). In order to test this hypothesis, a study using a multisyllabic word list balanced for orthographic neighbourhood size as well as item frequency is required.
In our final research question, we examined error productions involving a shift in lexical stress to investigate whether children were likely to regularize based on statistical properties of English trisyllabic words. Our hypothesis was upheld: just over half of participants’ stress-shifting error productions moved the word's primary stress to the first syllable. Moreover, there was a significant increase in the proportion of regularization errors across reading quartiles. Only children's SSPA was significantly related to their tendency to regularize trisyllabic words’ lexical stress. Together, these results suggest that errors involving a shift in primary stress to a trisyllabic word’s first syllable may indicate stronger reading skills as well as a better sense of the implicit patterns of spoken English, including prosodic awareness. Making an error that indicates a desirable level of proficiency recalls Savage and colleagues’ finding that scaffolding errors indicate age-appropriate reading skills (Savage, Stuart, & Hill, 2001).

**Error Analysis of Multisyllabic Word Reading**

As we noted at the beginning of this chapter, children’s word reading accuracy is typically scored dichotomously; although this method is efficient, it misses the opportunity to provide more information as to the skills and strategies children may be applying to the task by describing and categorizing errors. There are a number of ways to be wrong, especially when words are longer and more complex; similarly, an analysis of word reading errors also contains multitudes: as with qualitative data analysis in general, there are a number of ways one could potentially describe error productions, depending on the items, the purpose at hand, and the patterns that arise within each child’s reading errors. Another benefit of error analysis is that it is less prone to errors than are subjective reports such as
verbal strategy reports (“What strategy did you use to read that word?”) (McGoern, Medford, & Moxon, 2013).

With respect to multisyllabic word reading, our error analyses added nuance and provided a different perspective to the question of how middle school students approach multisyllabic word reading. In keeping with Greenberg and colleagues’ comparison of adults and children, older students (Grade 5s) were more likely to make lexicalized errors than younger ones (Grade 4s), perhaps because they guessed or accessed an item from memory based on partial letter cues (Greenberg, Ehri, & Perin, 2002). However, whether this is due to their relative skills, inclinations, or teaching methods is unknown.

**Coding Scheme for MSWR**

One of the contributions this study makes to the reading literature is the coding scheme, devised specifically for use with multisyllabic words. The primary intention of the scheme was to capture both decoding and linguistic stress errors concurrently, in an exploratory fashion: our aim was to amass a relatively broad base of information about the types of errors that children made on multisyllabic words. Previous analyses of word reading errors began by defining errors by the degree of deviation from the target (Davidson & Strucker, 2002; Greenberg, Ehri, & Perin, 2002; Savage, Stuart, & Hill, 2001). The coding scheme successfully divided decoding errors (Code A) from stress errors (Code B) from those errors in which both decoding and stress errors co-occurred (Code C) and allowed a more qualitative, descriptive analysis of participants’ reading accuracy upon encountering multisyllabic words.

Equal Stress (Code D) and Incorrect Number of Syllables (Code E) were more heterogeneous in terms of the error productions captured. Equal Stress errors tended to fall
broadly into two subcategories of errors: (a) the laborious, phoneme-by-phoneme decodings of readers as yet unable to apply chunking to read longer words and/or to blend long strings of phonemes together (e.g., “k... ek... ssp... or... p... i... p... tin... on” for exception) and (b) the more “robotically” read syllable-by-syllable decodings we expected upon creation of that category (e.g., “deh... preh... shun” for depression). This is an example of how, rather than holding rigidly to the error categories set out a priori, taking a more traditionally qualitative approach involving “emergent” coding may have been resulted in a more representative list of error categories. Another option would have been to accept that the scheme’s codes were not mutually exclusive, allowing for the assignment of multiple codes. Refinement of the coding scheme should consider the options that would most benefit rich description of children’s error production.

Another improvement to the coding scheme would be to begin by categorizing children’s error productions into real words and nonwords (Adams & Huggins, 1985; Greenberg, Ehri, & Perin, 2002). As our analyses demonstrated, children made lexicalization errors regardless of reading quartile; lexicalizations constituted approximately a quarter of the error productions made by these typically developing readers. Beginning here would improve our coding scheme: half of the categories (Decoding Error, Decoding and Stress Error, and Incorrect Number of Syllables) could accurately be applied to either real words or nonwords. In effect, future error coding could be organized into a “tree-root structure,” beginning with a division into real and nonword responses before continuing to further classifications. Finally, the inclusion of an “unclear” category may lead to higher levels of inter-rater reliability (McGeown, Medford, & Moxon, 2013).
Error coding that addresses morphological aspects of MSWR. Based on the strong relationship between MA and MSWR, the application of morphological error codes would certainly add value to our analyses. This would entail a second pass at participants’ error productions and would require items to be carefully selected to include and balance morphological features. Error types could include degree of morphological preservation in non-lexicalized error productions (e.g., root, prefix, suffix accuracy); morphological errors could provide information about how (a) derivational suffixes might influence lexical stress regularization errors, and (b) root and suffix frequency could affect error production. Although our results were in keeping with previous findings that children continue to rely on phonological decoding and that their ability to use it is related to the pattern of errors made (Greenberg, Ehri, & Perin, 2002; Shefelbine & Calhoun, 1991), the strength of MA’s relationship with MSWR suggests that there is a great deal to be gained by including morphological errors in future error analyses.

In terms of item selection, a specific aspect of morphology that may be important for multisyllabic words is one that provides meaningful information about pronunciation: derivational suffixes. This type of suffix generally belongs to one of two classes: neutral and nonneutral, both of which provide information about a word’s lexical stress. Neutral suffixes (e.g., -ness, -er, -ment) are relatively easy to learn (Tyler & Nagy, 1989) and do not cause a shift in lexical stress when added to the stem (e.g., enterTAIN – enterTAINment); nonneutral suffixes (e.g., -ity, -ion, -ive), however, necessitate a shift in stress to the syllable immediately preceding it (in the case of trisyllabic words, to the second syllable; e.g., ROtate – roTAtion). Both children and adults have been found to make productive use of the morphological information provided by derivational suffixes, and their ability to do so is
related to their reading ability (e.g., Carlisle, 2000; Clin, Wade-Woolley, & Heggie, 2009; Shankweiler et al., 1995; Wade-Woolley & Heggie, 2015). An analysis of lexical stress errors as they related to derivational suffixes was not possible in the current study; task construction was not focused on the inclusion of items with both neutral and nonneutral suffixes. The fact that participants made lexical stress errors (with and without accompanying decoding errors) suggests a potential direction for future work that focuses on items constructed to determine whether derivational suffixes and lexical stress errors have a meaningful relationship.

**Limitations and Future Directions**

As we have discussed, error analysis provides information about the decoding strategies that children may be applying to word reading. However, the technique is limited in that it provides indirect evidence and of said strategies and furthermore, the information is incomplete: only strategies that children may have been using when they get the wrong outcome, not the strategies that they were using when they read the word correctly. In addition, error analysis does not provide any information about the instructional approach(es), an influential and variable aspect of the children’s word reading context.

As we noted in Chapter 2, ratings of familiarity with the items would have benefited our analyses. It is likely that participants made more errors on items with which they were unfamiliar; just how their errors would break down on known and unknown words, as well as by type (e.g., decoding errors, lexicalizations), is an empirical question. As we also noted in Chapter 2, a larger sample size, over a broader range of grades and reading abilities, would have given us the statistical power and sample diversity for a clearer picture of multisyllabic word reading. We made the decision to present the target words in isolation
in the current study; this limited participants’ ability to use context as a decoding strategy. For our purposes, this decision was appropriate; however, an examination of MSWR in sentence context would better reflect how students in Grades 4 and 5 encounter unfamiliar multisyllabic words. Also, given the importance of morphological awareness for MSWR, a list of items that was compiled specifically to balance morphological elements (e.g., complexity, derivational suffixes, semantic relatedness) would greatly add to our understanding of how MA is applied to multisyllabic words.

Although we were able to describe the types of errors made by students in Grades 4 and 5 in many ways and with some detail, the rough transcriptions limited the types of analyses we were able to perform. Closer phonetic transcription of error productions would provide the basis for more detailed and accurate information about syllable- and phoneme-level errors, particularly with respect to vowels; this is a component of multisyllabic words that is pertinent in light of lexical stress and vowel reduction in English. In addition, we rated items’ lexical stress subjectively (i.e., by ear); although strengthened by inter-rater reliability information, we recognize that subjective ratings of stress placement are not as accurate an assessment as acoustic measures. Future analyses of participants’ error productions should include additional descriptive categories for these types of errors.

Conclusion

In sum, an analysis of the errors children make when decoding provides a more nuanced and detailed picture of multisyllabic word reading ability. The current study aimed to describe some of these errors — focused on decoding errors, lexical stress errors, lexicalization errors, and stress regularization errors — and to make connections between error types and child-level skills. Our results indicate that the types of errors readers make
often depends on their reading ability: good readers make finer grain size errors (decoding errors, stress placement errors) while poorer readers made more basic reading errors such as equal stress or incorrect number of syllable errors. Children’s MA, PA, vocabulary, and naming speed were related to the majority of error types, while SSPA was related solely to errors involving stress regularization.

The application of error analysis to a reading outcome most frequently scored dichotomously constitutes a significant contribution to the literature on multisyllabic word reading, including the development of the MSWR coding scheme. There remains work to be done to refine the error coding process and accurately categorize MSWR errors, but this study constitutes an important step and also highlights child-level skills and aspects of multisyllabic words to focus on in subsequent studies. Error analysis was initially a foundational source of information for theory building in reading development but it now has the potential to add to our knowledge about the nuance underlying individual differences in word reading.
CHAPTER FOUR: General Discussion

I began Chapter 1 by situating my examination of multisyllabic word reading (MSWR) and its importance within a broader context of literacy difficulties and their impact on our lives as readers, communicators, and members of a society reliant on written communication. As I noted then, the resounding importance of oral language for learning to read, as well as the foundation for reading provided by automatized decoding skills, both support our success as readers, writers, and learners. In the following two chapters, I have, through the two studies described therein, tried to make some progress on the question of whether and how reading outcomes might be predicted or explained by focusing specifically on multisyllabic word reading (MSWR). The two manuscripts took different perspectives: one was an overall examination of MSWR and the other analyzed the errors children made as they read these longer words; in each case, the relationship between various reading-related skills and MSWR was also examined.

At the heart of most reading difficulties is a struggle with word reading (Shaywitz & Shaywitz, 2008; Stanovich & Stanovich, 2003); for students who continue to struggle with word reading into secondary school, their challenge typically lies with multisyllabic words (Archer, Gleason, & Vachon, 2003; Carlisle & Katz, 2006; Moats, 1998). Previous studies seem to suggest that the factors involved in reading monosyllabic words aloud are also important for MSWR (e.g., Kearns, 2015; Kearns et al., 2016; Yap & Balota, 2009), but that additional factors also become critical only for these longer words. In this final chapter, I discuss the implications of my results in terms of three broad areas: theory, instruction, and future research.
Multisyllabic Word Reading: Implications for Theory

At the outset, I noted that just by virtue of being longer, multisyllabic words are more difficult to read (New, Ferrand, Pallier, & Brysbaert, 2006). For example, a reader who decodes a 10-letter word (e.g., strawberry, confidence) without the ability to use chunking will struggle as the task exceeds her working memory capacity, even before aspects such as vowel pronunciation ambiguity are considered. But length effects are only the beginning: there are a number of factors that contribute to multisyllabic words’ complexity, reflected in the number of significant relationships observed between child-level skills and MSWR ability across my two studies.

Multisyllabic Words and Related Skills

Just as learning to read is based on a foundation of oral language (Mattingly, 1984; Nagy & Anderson, 1999), so too is MSWR reliant on a foundation of successful monosyllabic word reading. Building on the knowledge and skills acquired for reading words, children must strategically apply and broaden their knowledge base to accommodate the increase in complexity that comes with multisyllabic words. Namely, longer words have an internal structure that involves different grapheme-phoneme connections from those found in monosyllabic words; these connections require other decoding methods, including morphological analysis (Verhoeven & Perfetti, 2011; Kirby & Bowers, in press). One of the primary aims I set out in each of my studies was to determine whether certain key skills were related to MSWR, both in terms of straightforward accuracy (Study 1) and error patterns (Study 2).

Phonological awareness (PA). While its foundational importance for reading development is no longer questioned in the literature, there has been conflicting evidence
in terms of whether PA’s influence wanes in favour of other skills (e.g., morphological awareness, naming speed), both as word reading skill improves and in response to longer, more complex words. For the Grade 4 and 5 students in my studies, PA was still an important component of their reading process, maintaining an indirect relationship with MSWR via morphological awareness (MA). PA was also related to some specific types of word reading errors: poorer PA was related to a greater incidence of decoding errors (whether or not they were accompanied by a stress shift), equal stress errors, and syllable-level errors. Both correct readings and the errors related to PA share a foundation of decoding skill: errors explicitly involving decoding (at the phoneme level) were often close approximations of the target (e.g., “de-poe-sit” for deposit; “pearl-a-ment” for parliament) while equal stress errors indicated weak decoding or a lack of fluency with decoding skills (e.g., “aw... nn... ser” for observer; “com... plit... ment” for compliment). A relationship between PA and errors involving the wrong number of syllables also suggests a failure to decode, namely one at the syllable level where students failed to accurately identify the syllable structure of the word (e.g., “a-men-din-ment” for amendment; “row-shun” for rotation). PA was not related to stress placement errors (regardless of stress regularization; e.g., “par-LE-ment” for parliament) or to lexicalization errors, suggesting that these errors did not rely on phonological skill.

**Suprasegmental phonological awareness (SSPA).** Much of the work on SSPA and word reading has done so on the basis of the hypothesis that prosody should be especially important for multisyllabic word reading (e.g., Corriveau, Pasquini, & Goswami, 2007; David, Wade-Woolley, Kirby, & Smithrim, 2007; Duncan & Seymour, 2003; Mundy, 2011; Wade-Woolley & Heggie, 2015), even going so far as to speculate that SSPA may "be an
ability that is crucial for the transition from reading monosyllabic words to multisyllabic words” (Harrison & Wood, 2016, p. 88). In my studies, I found that just as PA was indirectly related to MSWR via MA, so too was SSPA. SSPA was related to correct readings of multisyllabic words but it was only related to one of the error types I examined: children’s tendency to regularize the trisyllabic items’ stress to the first syllable. The common thread here may be that an awareness of lexical stress is an aspect of one’s oral language competence that strengthens words’ representation in the mental lexicon (Perfetti, 1985; Perfetti & Hart, 2002), supports accurate pronunciation when the word is known, and is evident in regularization errors when the word is unknown or less familiar. Children’s oral language competence, therefore, may include an implicit sensitivity to the statistical properties of English words, internalized from hearing and using the language.

Although a specific relationship between SSPA and multisyllabic words makes theoretical sense, it is important to remember that it is an over-simplification to posit or conclude that prosody somehow emerges from nothing to become a factor for MSWR. Monosyllabic content words in English are also stress-bearing (e.g., I and dog versus a and the), even in isolation. Prosody at the lexical (word) level is not activated only for multisyllabic words. The context of multisyllabic words is more frequently employed when explaining or defining lexical stress, as it is a more obvious feature of multisyllabic words: the relative strength and weakness of neighbouring syllables, along with the resulting vowel reduction, provides a clearer and more easily understood illustration.

**Development of SSPA over the elementary grades.** My review of the research on SSPA and word reading underscored just how unclear children’s awareness of prosodic information and its application to word reading remain, especially the degree to which
SSPA develops and changes over the elementary school years. Much of the research on SSPA and word reading has focused on the early years (preschool to Grade 1; e.g., Beattie & Manis, 2014; Goodman, Libenson, & Wade-Woolley, 2010; Holliman, Critten, Lawrence, Harrison, Wood, & Hughes, 2014; Holliman, Williams, Mundy, Wood, Hart, & Waldron, 2014; Holliman, Palma, Critten, Wood, Cunnane, & Pillinger, 2016; Holliman, Wood, & Sheehy, 2008, 2010; Kim & Petscher, 2015; Wood & Terrell, 1998; Wood, 2006). While some work has been done with adults (e.g., Chan & Wade-Woolley, 2016; Mundy & Carroll, 2016; Wade-Woolley & Heggie, 2015), there have been far fewer studies on readers in middle and late elementary grades (e.g., Wade-Woolley, 2016; Whalley & Hansen, 2006). Our understanding of SSPA’s development has been hampered by the lack of longitudinal studies; of those that have been conducted (e.g., Calet, Gutierrez-Palma, Simpson, Gonzalez-Trujillo, & Defior, 2015; Holliman, Wood, & Sheehy, 2010), the majority span only one or two years.

One exception, both in terms of design and age range, is Goswami and colleagues’ 4-year longitudinal study of children with dyslexia. Dyslexics’ performance was compared to that of reading- and age-level controls, showing that dyslexic readers were at a significant deficit in SSPA at age 9, measured using a prosodic awareness task that relies on a reiterative speech technique (DEEdee; Whalley & Hansen, 2006); the dyslexics’ SSPA deficit was even more significantly marked than their deficits in PA, which the authors interpreted as evidence that “insensitivity to prosodic structure may be a causal factor in developmental dyslexia in English” (Goswami, Mead, Fosker, Huss, Barnes, & Leong, 2013, p. 9). Half of the participants were tested again four years later, this time using a direct stress perception task (a same/different task; e.g., Difficulty, diFFiculty), dyslexics
performed more poorly than age-matched controls, suggesting that while SSPA does develop in children with dyslexia, it does not do so to age-appropriate levels (Goswami et al., 2013). This longitudinal study, however, was not truly focused on the development of SSPA over the elementary years; it does, together with evidence from the current and other studies suggesting the interactive influence of other variables (e.g., PA, MA, vocabulary, naming speed), support the need for further longitudinal research on SSPA’s development as well as more comprehensive studies to elucidate the various factors influencing MSWR.

**Morphological awareness (MA).** MA demonstrated the strongest relationship to MSWR, despite the fact that the list of items on the MSWR task was not made up exclusively of morphologically complex items. This relationship held even after controlling for a number of variables (i.e., age, nonverbal reasoning, naming speed, vocabulary, PA, and SSPA). The stronger children’s MA, the more likely they were to read multisyllabic words correctly and the less likely they were to make errors of all types (save stress placement errors, which were not related). MA was also related to lexicalized errors, but in reverse: the better participants’ MA, the more lexicalized errors they were likely to make. Here, it may be the case that by Grades 4 and 5, students are more aware of and concerned with real words as the outcome of decoding. Younger readers may be more accustomed to sounding out words only to arrive at an unknown word or nonword. To put it another way, participants with stronger MA have a greater degree of morpheme and word recognition, which they can apply to the reading process, making them more likely to gravitate toward lexicalized error productions.

Notably, several participants made the same lexicalized errors on the same items (e.g., “attitude” for altitude, “argument” for arrangement, “defense” for defiance, “illustration” for
illusion, “mystery” for misery, and “organ” for origin). These errors were often near matches to the targets, preserving morphemes (e.g., -ment, -(t)ion, -y) as well as syllabic structure. It is reasonable to conclude, then, that participants’ MA helped them to analyze the items and read them as near matches, even if they did not read the words correctly.

I noted in Chapter 1 that English adheres to the isomorphism principle, which states that morphemes are to be given priority over phonology; this principle is a factor in the “inconsistent” and opaque orthography of English, wherein orthographic consistency comes at the expense of reduced grapheme-phoneme consistency (Verhoeven, Schreuder, & Baayen, 2003). I discussed the isomorphism principle as an underlying reason for word reading difficulty in English, but this orthographic principle may also be reflected in the mediating role that morphological awareness played in PA and SSPA’s relationship to MSWR in Study 1. Phonological information does not become unimportant for English in the face of morphological information, but the interaction between the two has consequences for how words are spelled. A stronger and more proximal relationship between MA and word reading reflects the preservation of morphological information at the expense of the phonological, which enjoys a weaker and more distal relationship with word reading as the items to be read become longer and more complex. Thus although MA is hypothesized to be important for MSWR because of the frequent morphological complexity of these items (e.g., Deacon, Whalen, & Kirby, 2011; Nagy, Berninger, & Abbott, 2006), the more general multisyllabic nature of the items and the adjustment of reading skills brought to bear on the task may also contribute to MA’s importance. This remains an empirical question to be tested in future studies.
Vocabulary. Like the majority of the variables I investigated, vocabulary was inversely related to nearly all types of errors (Stress Only were not related), as well as positively related to correct readings of multisyllabic words. These relationships were second only in strength to those with morphological awareness. Participants’ vocabulary was also related to the incidence of lexicalized errors: just as with MA, the better participants’ vocabulary, the more lexicalized errors they were likely to make. (This pattern may also be due in part to the strong correlation between MA and vocabulary \(r = .769\).) If we assume, as noted above, that by Grades 4 and 5 students likely expect to produce a real word – and necessarily, one that is known to them – as a result of their decoding efforts, then it is not surprising that they would produce more lexicalized errors if they also had a greater vocabulary. A breakdown in decoding resulting in a nonword would not be accepted. Instead, these readers may subconsciously tap into their vocabulary skills to find a near match, a word they have in their oral vocabulary but for which they may not have a solid orthographic representation in their written vocabulary.

Support for this hypothesis comes from the fact that in several cases, many participants produced the very same lexicalized errors, on the same items. As I noted above, examples included altitude read as “attitude,” arrangement as “argument,” defiance as “defense,” illusion as “illustration,” misery as “mystery,” and origin as “organ.” Although all the items on the MSWR task were selected to be at or just above participants’ grade level, when lexicalized errors were produced, more frequent words were substituted. Of these six examples, the most frequently lexicalized items, the mean SFI (Standard Frequency Index\(^8\))

\(^8\) As noted in Appendix B, Standard Frequency Index (SFI) can range between 3.5 (low frequency per million) and 88.3 (high frequency per million) on a logarithmic scale (Zeno et al., 1995). For
of the target items was 49.35 while the mean SFI of the lexicalized error productions was 54.2; however, these substituted words were also ones that children would have been more likely to see in print: targets appeared a mean of 2.34 times in Grade 4 and 4.67 times in Grade 5 written corpora, while lexicalized error productions appeared 8.67 and 11.67 times, respectively (Zeno et al., 1995).

To accurately decode words in an opaque orthography such as English, it is important to know the word — especially when the item is less predictable or familiar. This is just as true for monosyllabic (e.g., irregular words such as yacht) as for multisyllabic words. Studies of single word reading report moderate correlations with vocabulary (e.g., Scarborough, 2001; Sénéchal, Ouellette, & Rodney, 2006); in keeping with this trend, vocabulary was moderately correlated with multisyllabic word reading ($r = .611$) in my studies.

**Naming speed.** Like MA, naming speed retained its significant relationship with MSWR accuracy in the context of the other child-level variables (i.e., age, nonverbal reasoning, vocabulary, MA, PA, SSPA). Naming speed was related to all types of errors (except Stress Only errors), although its relationship was strongest with correct readings of multisyllabic words. Both naming speed and MSWR steadily increase over the elementary school years: naming speed’s contribution to word reading increases (Kirby, Parrila, & Pfeiffer, 2003; Scarborough, 1998; Van den Bos, Zijlstra, & Spelberg, 2002), as does the frequency of multisyllabic words in texts (Hiebert, Martin, & Menon, 2005; Kearns et al., 2016; Zeno, et al., 1995).

---

example, an SFI of 50.0 is approximately 10 per million, such as words like abuse, sphere, grammar, and charter (Zeno et al., 1995, p. 17).
It should not be surprising that naming speed retained its relationship to MSWR, even in the context of other variables, as it has been described as one of the “most universal, longitudinal, and concurrent predictors of reading ability” (Araújo, Reis, Petersson, & Faisca, 2015, p. 869), even a “microcosm” of reading (Wolf & Bowers, 1999). Although it has been included in studies of multisyllabic word reading (e.g., Kearns et al., 2016), it is not often the focus. One of the main outcomes of the current studies is support for a multi-component network of skills and word-level factors that all contribute to MSWR accuracy. Naming speed may be related to word reading as an assessment of the degree to which the reading network (such as the one described in the triangle model) is automated, with fluid connections between phonological, orthographic, and semantic aspects (Al Dahhan, Kirby, & Munoz, 2017).

**The Spreading Activation of an Interconnected Network**

One of the implications for theory raised by an examination of MSWR is whether and how our mental representation of words requires adjustments when we are dealing with longer words. Although I initially presented a stage theory of reading (Chall, 1983, 1996) as a framing device for MSWR, my studies’ results suggest that a more dynamic system, capable of juggling multiple components concurrently, may be more appropriate. Is multisyllabic word reading an example of complexity that must be met with complexity? Word reading, formed as it is on a foundation of language competence and cognition (Mattingly, 1984; Nagy & Anderson, 1999), must necessarily posit a reliance on multiple sources of information that weave their way in and out of prominence as reading skill develops. Stage theories (e.g., Chall, 1983, 1996) and the Simple View of Reading (Hoover & Gough, 1990) are parsimonious illustrations of reading that can be quickly and clearly
explained to novices of the field. Theories such as the triangle model that take a systems approach wherein interconnections and integration of information (Seidenberg & McClelland, 1989; Seidenberg, 2005) may require a more sophisticated understanding of reading and cognition. It is not my intention to imply, however, that reading skill — even limited to reading monosyllabic words in isolation — lacks complexity; if decades of research have taught us nothing else, it is that learning to read “is rocket science” (Moats, 1999). It is well understood that to build skilled reading, children must integrate increasingly strategic components of language comprehension (e.g., language structures, background knowledge) with increasingly automatic word recognition skills (e.g., phonological awareness, decoding, sight word recognition) (Scarborough, 2001). More parsimonious theories of reading development may underestimate the influence of additional factors, but an approach in which the causal factors are interwoven, contributing to reading skill from an early time (e.g. Scarborough’s “twisted braid” (2001) or the multiple systems approach (Parrila & McQuarrie, 2015)), embraces and describes the complexity more accurately. For example, children have extensive morphological knowledge as part of their oral language competence in and prior to the primary grades (e.g., Brown, 1972; Clark, 1982) and there is evidence that they are able to apply it from the beginning, before it was hypothesized in the majority of stage theories (Berko, 1958; Bowers, Kirby, & Deacon, 2010; Lyster, Lervåg & Hulme, 2016).

**Overlapping Waves Model.** This model states that there is variability both between and within children: there are a variety of strategies available and they will choose among them based on the nature of the problem at hand — in this case, the word to be decoded (Siegler, 1991). Most often applied to domains that support algorithmic strategies such as
mathematics, the Overlapping Waves framework has also been applied to reading and spelling (e.g., Farrington-Flint, Vanuxem-Cotterill, & Stiller, 2009; Lindberg et al., 2011). In this context, the “strategies” that readers “choose” are much more likely to be selected implicitly or unconsciously than the explicit selection of a strategy to solve a math problem; with this difference in mind, the application of the Overlapping Waves Model holds. Children may also choose to employ more than one strategy to complete a task or solve a problem; the model describes their thinking as characterized by three key characteristics: wide variability, adaptive choice, and gradual change in strategies over time (Siegler, 1995). In fact, one of the papers reviewed in Chapter 3 (Greenberg, Ehri, & Perin, 2002) was examined using the lens of the overlapping waves model: Greenberg and colleagues compared strategy use between two groups, typically developing children and adult literacy learners. The two groups contrasted in terms of relative reliance on phonological and orthographic strategies, depending on their skill level with each type, and this difference was described as an example of how skill base influences readers’ strategy choice using the Overlapping Waves model (McGeown, Medford, & Moxon, 2013).

While I did not measure or focus directly on students’ reading strategies, I began with the assumption that an analysis of word reading errors could provide a window into the types of strategies that children were applying when decoding multisyllabic words. Results suggested that there are a number of skills involved in MSWR, an outcome that is in keeping with a toolbox of strategies to be applied to the problem solving process of MSWR. It is also in keeping with phase models of reading development, which update the stage model of reading and note specifically that phases can occur, change and may even interact – at any point during literacy development (Ehri, 2005, 2014). A wide variety of available tools and
strategies for analysis, from which readers select – unconsciously and adaptively – to suit the word to be read, as well as a gradual change in the available tools and strategies over time are consistent with an overlapping waves framework.

The variety in error types made by individual participants also supports this hypothesis. Upon examination of individual items from the MSWR task, a variety of error types were made on most items. For example, of the 93 times it was read, origin was decoded correctly 35 times, involved a decoding error 11 times, a stress error twice, equal stress twice, and involved the wrong number of syllables 43 times; error productions on this item were lexicalized 37 times, 29 as organ. This variety was also seen in Farrington-Flint & Wood’s (2007) examination of young children’s self-reported strategy choices, which suggested that participants displayed abundant variability of strategies (e.g., grapheme-phoneme correspondence rules, decoding by analogy) for decoding nonwords, and that they made their selections purposefully, depending on the context (i.e., the word). Further, there were individual differences in reading strategy (i.e., lexical analogies, alphabetic decoding strategies (GPC rules), or a combination of the two), despite the fact that all the readers in Farrington-Flint & Wood’s sample had received the same form of reading instruction.

Thus, when we consider the implications for theory, there are many factors that contribute to readers’ MSWR accuracy, as well as to the types of errors they tend to make. The results of my studies suggest that a framework such as Scarborough’s “twisted braid” or Siegler’s Overlapping Waves model may best describe how children approach MSWR. The orthographic avalanche of longer, more complex words that begins in the middle
elementary grades appears to be best met by a broad and flexible system of word reading skills and strategies.

**Multisyllabic Word Reading: Implications for Instruction**

Strictly speaking, my studies’ designs do not allow me to make definitive conclusions regarding instructional implications; nevertheless, it remains valuable to discuss the possibilities. The results of my error analysis – and, I suspect, teachers’ experience with students’ word reading development – highlighted the variety of errors evident in children’s decoding attempts on multisyllabic words. If these errors do in fact reflect differing and flexible decoding strategies, then a similarly nuanced and responsive instructional method could be devised; whether this adaptive reading instruction is categorized as a type of Differentiated Instruction (Tomlinson, 1999) or as a way to locate each student’s Zone of Proximal Development (Vygotsky, 1987; Schnozt & Kürschner, 2007) in word reading is a theoretical framing decision, but either of these reflects a fit in terms of instructional design.

The strong relationship between MA and MSWR suggests that increased instruction in morphological knowledge\(^9\) (MK) may similarly be related to an improvement in MSWR. Support for explicit instruction in MK rests both on the fact that more than 60% of the new multisyllabic words encountered in the middle elementary grades have relatively transparent morphological structures, able to be broken down and assigned meanings to aid decoding (Nagy & Anderson, 1984; Nagy & Scott, 1989) and on evidence suggesting that weaker readers not only tend to have specific difficulties decoding multisyllabic words but

\(^9\) I use the term morphological knowledge (MK) here to make the distinction between one's predominantly implicit awareness or understanding of morphology (MA) and the knowledge resulting from explicit instruction in morphology (MK). However, it is noted that many researchers tend to use the two terms interchangeably.
also tend to lag behind their peers in morphological knowledge (Nagy, Carlisle, & Goodwin, 2014). Explicit instruction in MK (e.g., Bowers, Kirby, & Deacon, 2010) is generative, providing another set of tools for children to use when decoding multisyllabic words (e.g., suffixes such as -ment (“action or process”) and prefixes such as dis- (“not, opposite of”) occur frequently across multisyllabic words and can provide cues to meaning). There has been evidence that morphological instruction can begin early, at or just prior to the dramatic increase in multisyllabic words in elementary texts: data show significant relations between MA and word reading as early as Grade 2 or 3 (Carlisle, 2003; Lyster, Lervåg, & Hulme, 2016). Finally, of the two most strongly related child-level factors I investigated, MK is more amenable to instruction than naming speed, for which evidence is minimal and mixed at best (Kirby, Georgiou, Martinussen, & Parrila, 2010).

An additional benefit of morphological instruction is that it enhances vocabulary. Meta-analyses of MA/MK and vocabulary have shown that morphological instruction positively impacts vocabulary acquisition with moderate effect sizes (e.g., Bowers, Kirby, & Deacon, 2010; Goodwin & Ahn, 2010). Explicit instruction in morphological knowledge has also been effective at increasing struggling readers’ vocabulary (Harris, Schumaker, & Deshler, 2011; Wysocki & Jenkins, 1987).

Based on my results and on the literature, it appears that instructional approaches aimed at improving multisyllabic word reading should be primarily based in morphology. The question of whether an additional component, lexical stress, may contribute to a program of morphological instruction remains an empirical question. Stress is an aspect of multisyllabic words that is implicit in morphological instruction (e.g., shifts associated with the addition of certain types of derivational suffixes); it has yet to be included as an
integrated element, however, and thus it remains to be seen whether doing so would further improve reading (and spelling) outcomes.

Some researchers have initiated discussions around SSPA-based interventions (e.g., Harrison & Wood, 2016), but to date none have combined SSPA with MK instruction nor have any specifically addressed MSWR. While Harrison and Wood’s discussion included addressing the needs of dyslexic learners and their difficulty detecting phonological properties of language, another group of readers was not addressed: those who are learning English as a second language (ESL). Depending on their first language (e.g., Dutch, French), English prosodic structure will be more (Dutch) or less (French) familiar, and challenges specific to the languages’ similarity/differences will arise accordingly. Without explicit orthographic stress markers (e.g., diacritics) as in Greek and Spanish to act as guidelines, the rules of English pronunciation (and thus word reading) may be frustratingly unclear. ESL readers may have the most to gain from a multisyllabic word reading intervention that involves a SSPA component; again, this is an empirical question, to be addressed in future research.

**Multisyllabic Word Reading: Implications for Future Research**

There are relatively few studies of multisyllabic word reading and as such the experimental field is wide and full of possibility. As I noted in Chapters 2 and 3, a longitudinal design would provide the best opportunity to delve more deeply into the contribution of and the interplay between the myriad factors that contribute to multisyllabic word reading. Provided appropriate items could be collected for the MSWR task to support such an age range, beginning in Grade 3 (to coincide with the sharp increase in multisyllabic words in school texts) and ending in Grade 6 or 8 would assure a rich data
set. Improved and extended measures for several constructs (e.g., SSPA, PA), as well as the addition of new measures (e.g., familiarity ratings, orthographic processing), would add further depth to a longitudinal study’s design. Closer transcription of errors (i.e., using the International Phonetic Alphabet) and a more open-ended coding process that further refines the scheme put into use in Chapter 3 will benefit future work on MSWR errors.

Another important consideration for future work is the inclusion of a word reading measure composed solely of monosyllabic words, to control for participants’ monosyllabic word reading. I have argued that multisyllabic word reading is a complex process that builds on one’s ability to read monosyllabic words; the inclusion of both types of word reading measures would allow for the isolation of the child-level skills that are relevant specifically for reading both short and long words. Given the empirical support for the many skills important for word reading in general, in addition to those that may be specifically relevant for MSWR, such a measure would make a valuable contribution to future research on MSWR to tease these skill sets apart.

As I noted in Chapter 3, error analysis is interesting in part because there are many different ways to misread a word, which in turn reveal how readers are processing the information in the word. Within the boundaries set for the current studies, I scored only participants’ final answers and did not code for the number of decoding attempts, false starts, or repetitions prior to a correct or incorrect final reading. Thus, I am unable to determine how many of the items were read by sight (i.e., directly retrieved from long-term memory; Ehri, 2005). Considering the relationships between MSWR accuracy and both naming speed and vocabulary, more detailed coding could provide insight into these relationships.
Few studies of SSPA and word reading have directly examined vocabulary's role; instead, vocabulary is most often included as a control variable. An exception to this is Holliman and colleagues’ (2014b) examination of SSPA’s relationship to word reading; using path analysis, the two were found to be related indirectly, partially mediated in two paths involving vocabulary: vocabulary to MA, and vocabulary to PA. This study was conducted with early readers (Grade 1); given the relationships between multisyllabic word reading and vocabulary in my studies, it would be interesting to delve further into vocabulary’s role in MSWR.

**Overall Conclusion**

In this multi-manuscript dissertation, I have aimed to make some progress with respect to two broad questions: what skills are important for multisyllabic word reading accuracy, and what can the types of errors children make tell us about multisyllabic word reading? The specific variables of interest (phonological awareness, suprasegmental phonological awareness, morphological awareness, vocabulary, and naming speed) included in my studies all played a role in MSWR accuracy and error production. The results of my studies suggest that morphological awareness and naming speed are particularly important for MSWR, but also that there are other child-level variables, as part of a complex system of integrated skills, which contribute to the word reading skills and strategies children apply adaptively to the task of MSWR. My analysis of the various errors children made when decoding multisyllabic words added nuance and detail to the complexities of this system.

Research focused specifically on multisyllabic word reading is still relatively scarce, especially when contrasted with the breadth and depth of evidence supporting reading
development in general. This study serves as a basis for further examinations of the
cognitive and linguistic variables that contribute to multisyllabic word reading, whether
examined using dichotomous scoring or analysis of error productions.
References


https://doi.org/10.2307/1593592


https://doi.org/10.3758/BF03200465

https://doi.org/10.3758/BF03193014

https://doi.org/10.2307/747673


https://doi.org/10.1037/0033-2909.130.6.858


https://doi.org/10.1017/S0142716401003083


https://doi.org/10.4159/harvard.9780674732469


https://doi.org/10.1006/drev.1993.1012


https://doi.org/10.1007/s11145-005-5766-2

https://doi.org/10.1598/RRQ.40.4.3

https://doi.org/10.1007/978-1-4757-3720-2_2


https://doi.org/10.1016/S0010-0277(03)00164-1

https://doi.org/10.1093/mind/os-XI.42.220


https://doi.org/10.1111/1467-9817.12083


https://doi.org/10.1016/0028-3932(76)90075-0

Deshler, D. D., & Hock, M. F. (2007). Adolescent literacy: Where we are, where we need to go. In M. Pressley, A. K. Billman, K. H. Perry, K. E. Reffitt, & J. Moorhead Reynolds (Eds.) *Shaping literacy achievement: Research we have, research we need* (pp. 8-128). New York: The Guilford Press.

https://doi.org/10.3102/0013189X10370204

https://doi.org/10.1037/0022-0663.95.3.465

https://doi.org/10.1111/1467-9817.00190


https://doi.org/10.3102/00346543071003393

https://doi.org/10.1598/RRQ.36.3.2

https://doi.org/10.1037/0022-0663.99.2.326

https://doi.org/10.1348/026151008X383148

https://doi.org/10.1121/1.412063

https://doi.org/10.1207/S15326985EP3501_5


https://doi.org/10.1017/S0021963099005284
https://doi.org/10.1177/0963721412469398


https://doi.org/10.1002/(SICI)1099-0909(200004/06)6:2<133::AID-DYS160>3.0.CO;2-A


https://doi.org/10.1007/s11145-009-9186-6

https://doi.org/10.1016/j.cortex.2012.05.005

https://doi.org/10.1016/j.jml.2013.03.001

https://doi.org/10.1073/pnas.122368599

https://doi.org/10.1017/S0142716403000134


https://doi.org/10.1080/03637750903310360

https://doi.org/10.1177/0956797613480187


Heggie, L., & Wade-Woolley, L. (Submitted.) Prosodic awareness and punctuation ability in adult readers.


https://doi.org/10.1044/0161-1461(2005/029)

https://doi.org/10.1007/s11145-016-9687-z


https://doi.org/10.1037/0096-1523.16.1.92


https://doi.org/10.1016/0749-596X(90)90044-Z


https://doi.org/10.1037/a0037518


https://doi.org/10.1016/j.lindif.2012.12.017


https://doi.org/10.1007/BF01027276


https://doi.org/10.1023/A:1008136012492


https://doi.org/10.1080/02702710801982050


McKenna, M. C., & Picard, M. C. (2006). Revisiting the role of miscue analysis in effective teaching. *The Reading Teacher*, 60(4), 378-380. [https://doi.org/10.1598/RT.60.4.8](https://doi.org/10.1598/RT.60.4.8)


https://doi.org/10.1016/j.jecp.2008.01.004


https://doi.org/10.1023/B:READ.0000044598.81628.e6

https://doi.org/10.1111/j.1467-9280.1995.tb00324.x


https://doi.org/10.1006/jecp.1998.2481

https://doi.org/10.1017/S0954579408000631


factors/teacher factors: Issues in literacy research and instruction (pp. 169—177).

Chicago: National Reading Conference.


144


[https://doi.org/10.1006/jecp.1996.0014](https://doi.org/10.1006/jecp.1996.0014)
https://doi.org/10.3758/BF03197098

https://doi.org/10.1007/978-3-642-69113-3

https://doi.org/10.1016/0749-596X(89)90002-8

https://doi.org/10.1016/0010-0277(90)90052-L

https://doi.org/10.1016/j.jecp.2015.11.016

https://doi.org/10.1207/S1532799XSSR0601_02


https://doi.org/10.1037/0033-2909.101.2.192


https://doi.org/10.1037/0012-1649.33.3.468


https://doi.org/10.1111/j.1467-9817.2006.00309.x


https://doi.org/10.1037/0022-0663.91.3.415


## Appendix A: Correlations Between Measures

<table>
<thead>
<tr>
<th></th>
<th>Vocab</th>
<th>Matrix</th>
<th>RAN</th>
<th>Blending</th>
<th>Elision</th>
<th>ASA</th>
<th>SC</th>
<th>NWSC</th>
<th>MA Comp</th>
<th>Word ID</th>
<th>MSWR</th>
</tr>
</thead>
<tbody>
<tr>
<td>WASI Vocab</td>
<td></td>
<td>.245*</td>
<td>-.238*</td>
<td>.373***</td>
<td>.421***</td>
<td>.254*</td>
<td>.732***</td>
<td>.689***</td>
<td>.769***</td>
<td>.612***</td>
<td>.611***</td>
</tr>
<tr>
<td>WASI Matrix</td>
<td></td>
<td></td>
<td>.060</td>
<td>.250*</td>
<td>.345**</td>
<td>.069</td>
<td>.455***</td>
<td>.383***</td>
<td>.454***</td>
<td>.314**</td>
<td>.311**</td>
</tr>
<tr>
<td>RAN Numbers</td>
<td></td>
<td></td>
<td></td>
<td>.001</td>
<td>-.152</td>
<td>-.036</td>
<td>-.284**</td>
<td>-.300**</td>
<td>-.316**</td>
<td>-.420***</td>
<td>-.429***</td>
</tr>
<tr>
<td>Blending Words</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.328**</td>
<td>.136</td>
<td>.365***</td>
<td>.432***</td>
<td>.431***</td>
<td>.368***</td>
<td>.381***</td>
</tr>
<tr>
<td>Elision</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.199</td>
<td>.422***</td>
<td>.490***</td>
<td>.494***</td>
<td>.472***</td>
<td>.466***</td>
</tr>
<tr>
<td>Aural Stress Assignment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.155</td>
<td>.354**</td>
<td>.276**</td>
<td>.358***</td>
<td>.318**</td>
</tr>
<tr>
<td>Suffix Completion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.706***</td>
<td>.924***</td>
<td>.717***</td>
<td>.736***</td>
</tr>
<tr>
<td>Nonword Suffix Choice</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.924***</td>
<td>.769***</td>
<td>.786***</td>
<td></td>
</tr>
<tr>
<td>MA Composite</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.805***</td>
<td>.824***</td>
</tr>
<tr>
<td>Word ID</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.938***</td>
</tr>
</tbody>
</table>

*Note. * p < .05, ** p < .01, *** p < .001

Note. All variables use raw scores except Elision (transformed score)
### Appendix B: MSWR Items and Attributes

*Multisyllabic word reading items (n=54) and their attributes.*

<table>
<thead>
<tr>
<th>Item</th>
<th>Stress Pattern</th>
<th>SFI</th>
<th>Grade 4</th>
<th>Grade 5</th>
<th>Grade 6</th>
<th>#Lett</th>
<th>#Phon</th>
<th>#Morph</th>
</tr>
</thead>
<tbody>
<tr>
<td>altitude</td>
<td>200</td>
<td>48.7</td>
<td>3</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>amendment</td>
<td>020</td>
<td>52.5</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>9</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>announcement</td>
<td>020</td>
<td>47.9</td>
<td>5</td>
<td>7</td>
<td>5</td>
<td>12</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>appointment</td>
<td>020</td>
<td>50.3</td>
<td>3</td>
<td>4</td>
<td>6</td>
<td>11</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>approval</td>
<td>020</td>
<td>52.2</td>
<td>2</td>
<td>7</td>
<td>5</td>
<td>8</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>arrangement</td>
<td>020</td>
<td>54.5</td>
<td>4</td>
<td>8</td>
<td>12</td>
<td>11</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>assistance</td>
<td>020</td>
<td>52</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>10</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>attitude</td>
<td>200</td>
<td>55.3</td>
<td>3</td>
<td>9</td>
<td>13</td>
<td>8</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>cabinet</td>
<td>200</td>
<td>52.8</td>
<td>4</td>
<td>8</td>
<td>10</td>
<td>7</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>candidate</td>
<td>200</td>
<td>52.7</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>committee</td>
<td>020</td>
<td>56.1</td>
<td>6</td>
<td>17</td>
<td>23</td>
<td>9</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>compliment</td>
<td>200</td>
<td>44.3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>10</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>consequence</td>
<td>200</td>
<td>49.5</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>10</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>corridor</td>
<td>200</td>
<td>49</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>8</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>defiance</td>
<td>020</td>
<td>45.2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>deposit</td>
<td>020</td>
<td>51.4</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>depression</td>
<td>020</td>
<td>54.3</td>
<td>3</td>
<td>8</td>
<td>12</td>
<td>10</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>dignity</td>
<td>200</td>
<td>50.8</td>
<td>5</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>discussion</td>
<td>020</td>
<td>56.5</td>
<td>6</td>
<td>11</td>
<td>16</td>
<td>10</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>distinction</td>
<td>020</td>
<td>50.9</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>10</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>election</td>
<td>020</td>
<td>54.8</td>
<td>5</td>
<td>10</td>
<td>13</td>
<td>8</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>enterprise</td>
<td>200</td>
<td>51.9</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>10</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>estimate</td>
<td>200</td>
<td>51.8</td>
<td>1</td>
<td>5</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>exception</td>
<td>020</td>
<td>52.3</td>
<td>2</td>
<td>6</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>gallery</td>
<td>200</td>
<td>47.1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>7</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>illusion</td>
<td>020</td>
<td>46.3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>8</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>improvement</td>
<td>020</td>
<td>51.9</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td>11</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>insurance</td>
<td>020</td>
<td>56.2</td>
<td>3</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>janitor</td>
<td>200</td>
<td>42.7</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>7</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>loyalty</td>
<td>200</td>
<td>51.7</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>misery</td>
<td>200</td>
<td>48.1</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>observer</td>
<td>020</td>
<td>49.2</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>8</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>origin</td>
<td>200</td>
<td>53.3</td>
<td>1</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>paradise</td>
<td>200</td>
<td>46.8</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>8</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>parliament</td>
<td>200</td>
<td>50.4</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>10</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>performance</td>
<td>020</td>
<td>55.6</td>
<td>4</td>
<td>7</td>
<td>11</td>
<td>11</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>physician</td>
<td>020</td>
<td>50.8</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>9</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>pioneer</td>
<td>102</td>
<td>49.6</td>
<td>4</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Item</td>
<td>SFI</td>
<td>Stress Pattern</td>
<td>Grade 4</td>
<td>Grade 5</td>
<td>Letters</td>
<td>Phonemes</td>
<td>Morphemes</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-----</td>
<td>----------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>----------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>primary</td>
<td>200</td>
<td>56.7</td>
<td>1</td>
<td>7</td>
<td>10</td>
<td>7</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>principle</td>
<td>200</td>
<td>55.4</td>
<td>1</td>
<td>4</td>
<td>7</td>
<td>9</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>procession</td>
<td>020</td>
<td>47.8</td>
<td>5</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>proportion</td>
<td>020</td>
<td>52.6</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>10</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>punishment</td>
<td>200</td>
<td>52.4</td>
<td>5</td>
<td>8</td>
<td>11</td>
<td>10</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>pyramid</td>
<td>200</td>
<td>50</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>register</td>
<td>200</td>
<td>51.8</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>8</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>relation</td>
<td>020</td>
<td>52.8</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>8</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>resentment</td>
<td>020</td>
<td>46.3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>10</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>rotation</td>
<td>020</td>
<td>49.8</td>
<td>4</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>sentiment</td>
<td>200</td>
<td>45.2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>9</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>specialist</td>
<td>200</td>
<td>48</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>10</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>substitute</td>
<td>200</td>
<td>51</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>10</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>suggestion</td>
<td>020</td>
<td>50.6</td>
<td>6</td>
<td>9</td>
<td>8</td>
<td>10</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>surrender</td>
<td>020</td>
<td>48.1</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td>9</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>tendency</td>
<td>200</td>
<td>52.6</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>8</td>
<td>2</td>
</tr>
</tbody>
</table>

*Note.* Using the MRC Psycholinguistic Database (Coltheart, 1981), Stress Pattern is set out by syllable: 2 (primary stress), 1 (secondary stress), 0 (unstressed), such that the first item, “corridor,” receives 200 because its primary stress is on the first syllable and the next two syllables are unstressed. Items’ Standard Frequency Index (SFI) can range between 3.5 (low frequency per million) and 88.3 (high frequency per million) on a logarithmic scale (Zeno et al., 1995). For example, an SFI of 50.0 is approximately 10 per million, such as words like abuse, sphere, grammar, and charter (Zeno et al., 1995, p. 17). Columns “Grade 4” and “Grade 5” indicate the number of item occurrences in grade level texts (e.g., “corridor” occurs 10 times in Grade 4 texts and 9 times in Grade 5 texts; Zeno et al., 1995). The final three columns denote the number of letters, phonemes, and morphemes in each item (Balota et al., 2007). Note that the list of items is sorted alphabetically, not in order of presentation.
Appendix C: Approval Letter from General Research Ethics Board

January 12, 2015

Ms. Lindsay Heggie
Ph.D. Candidate
Faculty of Education
Queen's University
Duncan McArthur Hall
511 Union Street West
Kingston, ON, K7M 5R7

GREB Ref #: GEDUC-762-14; Romeo # 6014525
Title: "GEDUC-762-14 Multisyllabic Word Reading and Prosodic Awareness in Elementary School Children"

Dear Ms. Heggie:

The General Research Ethics Board (GREB), by means of a delegated board review, has cleared your proposal entitled "GEDUC-762-14 Multisyllabic word reading and prosodic awareness in elementary school children" for ethical compliance with the Tri-Council Guidelines (TCPS) and Queen's ethics policies. In accordance with the Tri-Council Guidelines (article D.1.6) and Senate Terms of Reference (article G), your project has been cleared for one year. At the end of each year, the GREB will ask if your project has been completed and if not, what changes have occurred or will occur in the next year.

You are reminded of your obligation to advise the GREB, with a copy to your unit REB, of any adverse event(s) that occur during this one year period (access this form at https://eservices.queensu.ca/romeo_researcher/ and click Events - GREB Adverse Event Report). An adverse event includes, but is not limited to, a complaint, a change or unexpected event that alters the level of risk for the researcher or participants or situation that requires a substantial change in approach to a participant(s). You are also advised that all adverse events must be reported to the GREB within 48 hours.

You are also reminded that all changes that might affect human participants must be cleared by the GREB. For example you must report changes to the level of risk, applicant characteristics, and implementation of new procedures. To make an amendment, access the application at https://eservices.queensu.ca/romeo_researcher/ and click Events - GREB Amendment to Approved Study Form. These changes will automatically be sent to the Ethics Coordinator, Gail Irving, at the Office of Research Services or irvingg@queensu.ca for further review and clearance by the GREB or GREB Chair.

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

Yours sincerely,

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

c:  Dr. Lesly Wade-Woolley, Faculty Supervisor
    Dr. Chris DeLuca, Chair, Unit REB
    Mrs. Angelina Gencarelli, c/o Graduate Studies and Bureau of Research