The Linguistic Profiles of Spelling Errors in Fourth, Fifth, and Seventh Grade Students

Yi-Chieh Wu
THE FLORIDA STATE UNIVERSITY
COLLEGE OF EDUCATION

THE LINGUISTIC PROFILES OF SPELLING ERRORS IN
FOURTH, FIFTH, AND SEVENTH GRADE STUDENTS

By
YI-CHIEH WU

A Dissertation submitted to the
School of Teacher Education
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Degree Awarded:
Summer Semester, 2013
Yi-Chieh Wu defended this dissertation on June 20, 2013.
The members of the supervisory committee were:

Barbara Foorman
Professor Directing Thesis

Christopher Schatschneider
University Representative

Young-Suk Kim
Committee Member

Beth M. Phillips
Committee Member

The Graduate School has verified and approved the above-named committee members, and certifies that the dissertation has been approved in accordance with university requirements.
To my family and friends who always support me.
ACKNOWLEDGMENTS

This journey of pursuing my dream could never have been possible without the supports of many individuals. I would like to thank my committee members whose guidance has nurtured me to become a confident researcher. I sincerely thank Dr. Barbara Foorman who has been compassionate since our first meeting and has guided me every step throughout the whole journey. I thank Dr. Chris Schatschneider who has transformed my fear of statistics into certainty with his witty and concise suggestions. I thank Dr. Young-Suk Kim who has demonstrated a great example of a dedicate scholar and inspired my research with precise questions. I thank Dr. Beth Phillips who reconstructed my understanding of methodology which provided a solid foundation of my research.

I would like to thank me wonderful colleagues and lab members who have been encouraging me and believing in me. I thank Denise Bishop who kindly offered to administer the spelling test. I thank Annette Nielsen Ladle, Adrea Truckenmiller, Alison Mitchell, and Lakeisha Johnson for being there for me all the time. Your caring kindness warmed my heart during those most difficult times. Without your helps and friendships, my dissertation could never be completed. A particular thanks to Dr. Hodge and Suzanne Wilkinson for arranging the administration at your schools. I would also like to thank the teachers, and students that participated in this study. Without their cooperation, the study would not be possible.

Finally, I would like to thank my family and friends. To my family in Taiwan, I could never achieve this goal without your tolerance and unconditional supports. To my brothers and sisters in our bible study, I thank you for your generous encouragement and the continuous weekly prayers. Most importantly, I thank God for giving me the strength, courage, and wisdom, and bringing this work to completion.
# TABLE OF CONTENTS

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

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

Abstract .......................................................................................................................................... ix

CHAPTER ONE – INTRODUCTION ............................................................................................1

CHAPTER TWO – METHOD ......................................................................................................28

CHAPTER THREE – RESULTS ..................................................................................................40

CHAPTER FOUR – DISCUSSION ..............................................................................................64

APPENDICES ...............................................................................................................................79

A. PERCENTILE RANKS OF FCAT 2.0 READING DSS ..............................................................79

B. THE RESEARCHER-DESIGNED SPELLING TEST ..............................................................82

C. HUMAN SUBJECTS COMMITTEE APPROVAL MEMORANDUM ..................................83

D. SAMPLE OF PARENTAL CONSENT FORM ......................................................................84

References ......................................................................................................................................86

Biographical Sketch .......................................................................................................................95
**LIST OF TABLES**

Table 1 Distribution of Subjects in Different Groups ..............................................................................29

Table 2 The Error Coding System .............................................................................................................31

Table 3 Means and Standard Deviations for All Scores of the Students in the Three Grades .....42

Table 4 Means (Standard Deviations in parentheses) for All Scores of the Average and Poor Readers in the Three Grades ...........................................................................................................48

Table 5 Means (Standard Deviations in parentheses) for All Scores of the Older Poor Readers and Younger Normally Progressing Readers ........................................................................................54

Table 6 Means (Standard Deviations in parentheses) for All Scores of the Unexpectedly Poor Spellers and Poor in Both Students in the Three Grades ........................................................................60

Table 7 The Results of Independent Samples t-tests between Different Subgroups .................63

Table 8 The percentile ranks constructed from FCAT 2.0 Reading DSS (Florida Department of Education, 2012) ..................................................................................................................80
LIST OF FIGURES

Figure 1 Histogram of the Spelling Accuracy Scores for the Fourth Grade Students .............42

Figure 2 Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation, and OL for Orthographic Legality) in the Three Grades ........................................45

Figure 3 Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation, OL for Orthographic Legality, and TES for Total Error Score) in the Three Grades .............................................................................................................................................46

Figure 4 Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation, and OL for Orthographic Legality) for the Poor and Average Readers in the Fourth Grade ..................................................................................................................................................50

Figure 5 Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation, and OL for Orthographic Legality) for the Poor and Average Readers in the Fifth Grade .................................................................................................................................................................50

Figure 6 Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation, and OL for Orthographic Legality) for the Poor and Average Readers in the Seventh Grade ..................................................................................................................................................51

Figure 7 Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation, OL for Orthographic Legality, and TES for Total Error Score) for the Poor and Average Readers in the Fourth Grade ........................................................................................................................................52

Figure 8 Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation, OL for Orthographic Legality, and TES for Total Error Score) for the Poor and Average Readers in the Fifth Grade .................................................................................................................................................................52

Figure 9 Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation, OL for Orthographic Legality, and TES for Total Error Score) for the Poor and Average Readers in the Seventh Grade ..................................................................................................................................................53

Figure 10 Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation, and OL for Orthographic Legality) for the Two Reading-Ability-Matched Groups ..................................................................................................................................................55

Figure 11 Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation, OL for Orthographic Legality, and TES for Total Error Score) for the Two Reading-Ability-Matched Groups ..................................................................................................................................................55

Figure 12 Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation and OL for Orthographic Legality) for the Unexpectedly Poor Spellers and Students with Poor Reading and Spelling Abilities .........................................................................................................................57
Figure 13 Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation, OL for Orthographic Legality, and TES for Total Error Score) for the Unexpectedly Poor Spellers and Students with Poor Reading and Spelling Abilities ..................58

Figure 14 Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation and OL for Orthographic Legality) for the Unexpectedly Poor Spellers and Average Spellers matched on the Reading Ability of the Unexpectedly Poor Spellers ...............59

Figure 15 Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation, OL for Orthographic Legality, and TES for Total Error Score) for the Unexpectedly Poor Spellers and Average Spellers matched on the Reading Ability of the Unexpectedly Poor Spellers .................................................................61
ABSTRACT

The purpose of this study was to investigate the use of linguistic knowledge in spelling by analyzing spelling errors made by 220 students in the fourth, fifth, and seventh grades. A 25-word researcher-designed spelling test with considerations of word frequency, word familiarity, and word type (based on morphological complexity) was administered. An error coding system was established based on the Triple Word Form theory. Each misspelling was coded based on its linguistic features and scored cumulatively in 3 categories: Phonological Representation, Orthographic Legality, and Morphological Legality. The error coding system revealed the linguistic profiles of misspellings and allowed the comparisons among subgroups matched on grades, reading, and spelling ability levels.

The results of profile analyses supported the Overlapping Waves Model, which advocates that spellers use their phonological, orthographic, and morphological knowledge in spelling simultaneously regardless of age, reading, or spelling levels. On the other hand, the study did not find evidence supporting the stage-specific theory, which defines each stage by observations of the consistent use of one strategy in spelling. The linguistic profiles revealed the competition between Phonological Representation and Orthographic Legality, which provided little evidence supporting the specific phonological deficit hypothesis. On the contrary, the researcher found that the key to becoming an average speller is to be able to effectively apply sufficient phonological knowledge in spelling. For students with poor reading ability, they do not just suffer from limited phonological knowledge but also from lack of other linguistic knowledge. For any two students with average reading ability, it is the one who can apply sufficient phonological knowledge that benefit in spelling and perform at the level that matches his or her reading ability. Educational implications are discussed.
CHAPTER ONE
INTRODUCTION

The Issue

In 2008, the National Assessment of Educational Progress released “The Nation’s Report Card: Writing 2007”. The report showed that both American eighth- and twelfth-graders have slightly improved their writing proficiency since 2002. However, the study also cautiously pointed out that only 33 percent of the eighth-graders and 24 percent of the twelfth-graders were considered “at or above proficient,” indicating a decrease in achievement between the two grades. In 2012, the results of the Florida Comprehensive Assessment Test (FCAT Writing) showed that only 27 percent of fourth-graders in Florida earned an acceptable score—a four or better on the six-point grading scale—which was a steep decline from 81 percent in 2011. In addition, eighth- and tenth-graders in Florida also showed similar drops on the FCAT Writing that requires students to write an essay on an assigned topic (Florida Department of Education, 2012).

It is evident that American students are not performing at the highest level in their writing and that instructional reform is desperately needed. According to the Simple View of Writing (Berninger, 2000; Berninger & Graham, 1998), transcription skills and executive functions are the foundational components of text generation. In this model, the more automatic the low-level transcription skills, the more working-memory resources are available for high-level composition skills (McCutchen, 1996). In another words, good spelling skills enable effortless translation from an idea into written language. In order to become a good writer, a student needs to have at least a minimum knowledge of conventional spelling.

However, spelling has received less attention than reading in research and instruction.
Researchers have focused much more on how children learn to read than how they learn to spell. Joshi and Carreker (2009) obtained analytic results from the ISI Web of Science and reported that 10,235 articles were published from 2003 to 2008 on reading, but only 896 on spelling. The strategic neglect of spelling instruction primarily started in the 1980s. A trend away from systematic spelling instruction began with the theory that spelling skills would ultimately be “caught” as children immersed themselves in reading and writing (Smith, 1983). Although language researchers have pointed out the regularities in spoken and printed English (e.g., Chomsky & Halle, 1986; Venezky, 1970), many teachers still believe that spelling can only be learned through memorization (e.g., Schlagal, 2001). In spite of the considerable evidence showing the failure of this approach, spelling instruction has not been significantly improved. Further understanding of spelling development can help educators improve spelling instruction so that students will perform better in their writing endeavors.

Even though there is far less spelling research than reading research, there has been growing interest in the relations between reading and spelling in predicting literacy skills at a later age. As stated by Ehri (2000), learning to read and learning to spell are two sides of the same coin. It is well known that learning to read and spell involves mastery of the association between printed and spoken forms of language (e.g., Adams, 1990; Ehri, 1998; Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998; Stanovich & Siegel, 1994). Ehri (2000) found high correlations (i.e., from .68 to .86) between word reading and spelling across six studies with students of various ages (first grade through post-secondary students). Longitudinal studies of literacy in elementary grades also indicated a bidirectional relation between word reading and spelling (Berninger, Abbott, Abbott, Graham, & Richards, 2002). Recently, Abbott, Berninger, and Fayol (2010) conducted a longitudinal structural equation modeling study and found that
both word reading and spelling were the most predictive and developmentally stable skills among multiple levels and domains of language skills.

On the other hand, researchers have reported several highly predictive early language abilities on later word reading and spelling performance, including alphabetic knowledge (Burgess & Lonigan, 1998), print awareness (Justice & Ezell, 2001), and rapid naming ability (Lovett, Steinbach, & Frijters, 2000). Converging evidence has also shown substantial correlations between word reading and spelling among the shared linguistic features of written language, including phonemic, orthographic, morphological, lexical, syntactic, and discourse features (Berninger, 2000; Shanahan & Lomax, 1986). To confirm the mutual learning process of reading and spelling, Conrad (2008) conducted an experimental study to examine the transfer effects on novel target word reading and spelling. The results showed that the repeated practice of reading and spelling can efficiently promote the solidification of orthographic representation. The literature has given us a theoretical foothold from which we can learn more about the basis of spelling in our understanding of word reading.

Demonstrating added complexity to the relationship between word reading and spelling, it has been shown that a good speller is typically a good reader, but a good reader might not be a good speller (Berninger, Abbott, Abbott, Graham, & Richards, 2002; Bosman & Van Orden, 1997). What makes spelling more often difficult than word reading is the opaque orthography and morpheme-dominant nature of the English language. Lacking such word-specific linguistic knowledge will lead to failure in spelling (Treiman & Bourassa, 2000). Recently, findings from meta-analyses (Graham & Perrin, 2006; Graham & Perrin, 2007) have suggested that spelling is a process that needs to and can be taught, just like reading.

Reed (2012) also pointed out that integrating the learning of reading and spelling can
promote the development of both skills. Practicing with eyes and hands simultaneously not only doubles the exposure dosage of the letter patterns but also provides the spellers opportunities to reflect and confirm that the spelling they produce looks and sounds accurate. Despite these benefits, not all educators realize the importance of spelling instruction and, for those that do, there is often uncertainty about how to teach spelling. Foorman and Petscher (2010) conducted a series of multilevel modeling analyses and found that 17 percent to 33 percent of the variance in spelling performance can be explained by differences at the classroom level in 6 out of 10 grades, indicating that not all students received the same quality or quantity of spelling instruction. More spelling research is needed to increase the understanding of spelling development and to improve spelling instruction which will lead to better spelling performance by students.

**Spelling Development Models**

Among several approaches to understanding spelling development, stage theory might be the most explored model. As early as 1971, Chomsky discovered that young children use their knowledge of letter-sound correspondences to write words. Then, Read (1971) was one of the forerunners of the stage approach with his findings on the systematic linguistic features of invented spellings made by young children. He found some spelling error patterns (i.e., *dot for don’t*) to be quite consistent from 20 children aged three to five. From then on, researchers such as Ehri (1987; 1989), Frith (1980), Gentry (1982), and Henderson and colleagues (Beers, Beers, & Grant, 1977; Henderson, Beers, & International Reading Association, 1980) proposed distinct stage models of spelling development. By analyzing invented spellings, these researchers identified specific stages of early spelling development. As a result of this body of work, educators have learned to assess and analyze the early literacy development of preschoolers and
primary grade students.

Even though stage theorists have no consensus about the exact number and names of stages, they have generally recognized three early stages of spelling development (Sharp, Sinatra, & Reynolds, 2008): (1) a pre-phonetic stage, where children use random symbols or letters to represent sounds; (2) a phonetic stage, where children attempt to map phonemes (sounds) and graphemes (letters); and (3) an orthographic stage, where children analyze words into larger orthographic units. Stage theorists typically have implemented error analysis to identify the stages of spelling acquisition based on the consistent use of a strategy. In other words, they have suggested that spelling development occurs across stages with different spelling strategies at different times. Once a specific strategy is fully mastered, the fluctuation between stages seldom happens (e.g., Ellis, 1994; Gentry, 1982). Beers and Beers (1992) summarized key findings from the stage theories: (1) children’s spelling errors are not random; (2) while children become better at spelling, the stages of orthographic awareness are identifiable; and (3) children proceed through the stages at different rates.

Although studies of young children’s invented spellings have provided valuable information about early spelling development, stage theories have provided less guidance in explaining misspellings typically produced by older spellers. Unlike beginning spellers, older spellers produce words with complex linguistic features and are expected to follow spelling conventions. Hence, typical misspellings provide a more comprehensive picture that encompasses all the related linguistic knowledge rather than just phonological knowledge. Without a sufficient investigation of typical misspellings, stage theorists claimed that some “high-level” strategies, such as orthographic and morphological knowledge, are unavailable to beginning spellers. Some non-English stage theorists have done studies that included older
children (Keuning & Verhoeven, 2008; Lervag & Hulme, 2010; Loizidou-Ieridou, Masterson, & Hanley, 2010; Rahbari & Senenchal, 2010; Sprenger-Charolles, Siegel, Béchennec, & Serniclaes, 2003). However, these studies have primarily been done in more transparent orthographies (i.e., Dutch, Norwegian, or Persian). A transparent orthography is characterized as having consistency in its correspondence of phoneme (sound) and grapheme (letter). Children in these studies mainly demonstrate phonological knowledge in their spellings due to the demands of their own languages. Thus, the results were not particularly informative in explaining spelling development in a more opaque written language, such as English, which requires more than merely phonological knowledge. In fact, the results of these non-English studies showed that children became skilled in different spelling patterns through development (Keuning et al., 2008), and older children tended to use more lexical strategies to spell instead of relying only on sublexical (phonological) strategies (Loizidou-Ieridou et al., 2010).

As stated above, stage theories are limited in their ability to explain the complexity of spelling development. Consequently, many researchers have started to explain spelling development from a more inclusive viewpoint where children use multiple strategies and different types of knowledge to spell from the beginning of their experiences with print (Treiman & Cassar, 1997; Hoijtink & Notenboom, 2004; Rittle-Johnson & Siegler, 1999; Varnhagen, McCallum, & Burstow, 1997). Varnhagen, McCallum and Burstow (1997) examined whether the strategy used in spelling was identifiably different across stages and consistent within stages. They analyzed the spelling production of first through six graders and focused on spellings of (a) silent –e long vowel words, and (b) regularly inflected past tense verbs. The results showed little support for stage theory and revealed the complexity of children’s spelling knowledge throughout spelling development. Similarly, Rittle-Johnson and Siegler (1999) generalized the
concept of the Overlapping Waves Model from cognitive development (Siegler, 1995) and found that the trend to use multiple spelling strategies in first grade was similar to that in second grade after one year. Also applying the Overlapping Waves Model, Hoijtink and Notenboom (2004) traced spelling development from first to sixth grades in a large data set (i.e., the responses of more than 3500 pupils to 245 items). They found no consistent occurrence of a stage-specific strategy within any developmental time cluster. In addition, the results showed that not only a variety of strategies was found, but also that the gradual changes in different rates across strategies confirmed the competition within and fluctuation between stages.

To conclude, Treiman and Bourassa (2000) pointed out that stage theories gave a general overall picture of spelling development but failed to explain the complex relations among phonological, orthographic, and morphological representations and the impact of such relations on spelling development. To better explain spelling development, we need to understand the key skills and linguistic knowledge that a good speller requires. Ideally, a skilled speller should be capable of manipulating sound units, recognizing the use of sound, identifying grammatical and semantic units within words, and appreciating the conventions of written language. These skills are the shared foundation of reading and spelling, which include phonological, orthographic, and morphological knowledge. Recently, Berninger, Abbott, Nagy and Carlisle (2010) conducted a growth curve analysis to investigate the development and contribution of phonological, orthographic, and morphological awareness on learning to read and spell. The results showed that all three kinds of linguistic knowledge were necessary and ought to be coordinated to improve literacy. In the next section, we will learn more about the relationship between the layering of linguistic features of English and the spelling-related cognitive components.
Components of Spelling Knowledge

English is well known as an opaque alphabetic orthography, which means that the correspondences between sounds and letters are less consistent than other more transparent orthographies, such as Finnish or Serbo-Croatian. However, English is not arbitrary (Joshi, Treiman, Carreker, & Moats, 2008-2009). Ziegler, Stone, and Jacobs (1997) reported that almost 70 percent of monosyllabic English words were consistent in terms of their letter to sound mapping. Although there have been different answers to the question of how regular English spelling is, it is true that the orthographic units of English not only correspond to sounds (phonemes and syllables) but also to meanings (morphemes). Thus, English orthography is most accurately described as a morphophonemic system (Chomsky & Halle, 1968). Overall, approximately ninety-six percent of English words follow phonological, orthographic, and morphological rules. It is evident that the acquisition of these rules is crucial and therefore needs to be explicitly taught.

Phonological knowledge is one of the most well-known shared linguistic foundations between reading and spelling. Its influence on the development of reading and spelling has been well documented (e.g., Adam, 1990; Catts, 1991; Ehri & Stahl, 2001; Manis, Doi, & Bhadha, 2000; Snow, Burns, & Griffin, 1998). Phonological knowledge is the ability to identify the sounds of a language and to manipulate the sequence of those sounds within a word. Nation and Hulme (1997) found that phoneme segmentation was a strong predictor of reading and spelling in children aged five to nine. Cunningham, Perry, and Stanovich (2001) reported that phonological awareness explains up to 40% of performance in reading and spelling, and poor phonological awareness is highly correlated with poor spelling ability (Shaywitz & Shaywitz, 2005). Furthermore, a series of intervention studies also showed that training in phonological
awareness improved the spelling and reading of children in low-income, inner-city schools (Graham, 1999). By analyzing spelling errors, we can uncover the use of phonological strategies and reveal whether a speller is having difficulties in identifying, segmenting, or representing each phoneme within a spoken word.

In spite of its predictive validity, phonological knowledge cannot explain all of the variance in children’s reading and spelling. Some researchers have been investigating additional linguistic components that are related to children’s literacy development, such as orthographic knowledge. Orthographic knowledge is comprised of the skills that enable the accurate translation of spoken words into written representations. These skills include (Wasowicz, 2007): 1) knowledge of which specific letter-sound patterns are acceptable (e.g., the /k/ sound can be represented by the letters c, k, ck, cc, lk, ch, and que); 2) knowing which letter patterns are acceptable (e.g., the /k/ sound is almost always spelled with the letter ‘k’ at the end of a word after a long vowel sound); and 3) understanding sound, syllable, and word position constraints on spelling patterns (e.g., the /k/ sound at the beginning of a word is never spelled with the letters ck, cc, lk). Fowler, Shankweiler, and Liberman (1979) investigated the progressive benefits of orthographic regularity and word context on spelling across different ages (i.e., second, third, fourth graders and college students). They found that spellers gained orthographic knowledge regardless of reading experience and that they applied this knowledge to spelling novel words. Spellers who do not have sufficient orthographic knowledge might make systematic mistakes, such as overgeneralizing rules or neglecting within-word environments, when spelling unfamiliar words. Again, error analysis is an informative tool to reveal whether a speller is sensitive enough to the orthographic rules within words.

As mentioned earlier, English orthographic units not only correspond to sounds but also
to meanings. Besides phonological and orthographic knowledge, a skilled speller also needs to understand English morphology—the letter-meaning representation of orthography. When spelling inflected or derived forms of words, morphological knowledge is needed (Carlisle, 1995). Morphological knowledge includes the letter-meaning relationships of morphemes (i.e., suffixes, prefixes, base words, and word roots), the semantic relationships between base words and related words, and the modification rules when adding affixes (Bourassa & Treiman, 2001). Inflected words contain suffixes that provide possession, gender, or number information if the word is a noun; tense, voice, or mood information if the word is a verb; and comparison information if the word is an adjective (Moats, 2000). Derivation is the process of creating a new word on the basis of an existing word (e.g., changing a noun to a verb; glory to glorify). The morphophonemic nature of English is apparent when the addition of a derivational suffix causes a shift in the pronunciation of the base word (e.g., sign, signal).

Unlike the suggestion of stage theorists, research has shown that young spellers not only apply their sound-letter correspondence knowledge but also their morphological knowledge while spelling. Treiman and Cassar (1996) found that young children were significantly more likely to spell the first segment of a final consonant cluster accurately when the cluster itself was an inflection or derivation (e.g., the /n/ sound in tuned is more likely to be spelled accurately than in brand). Devonshire and Fluck (2010) also found that five- to eleven-year olds who tended to make a morphological connection with their spelling scored higher than those who only used phonological or memory-based retrieval strategies.

Spelling is a complicated linguistic task that requires knowledge of sound-letter correspondences, meaningful parts of words, and word origins. Recently, an overview provided converging evidence that supported the Triple Word Form Theory (Berninger, 2004; Berninger,
Abbott, Billingsley, & Nagy, 2001; Berninger & Richards, 2002) based on family genetics, brain imaging, and treatment research at the University of Washington Multidisciplinary Learning Disabilities Center (UWLDC) over the past decade (Richards et al., 2006). Word forms refer to the multiple formats in which the brain may code the same word. Learning to read and spell requires storage and processing of these word forms in working memory (Berninger, Abbott, Thomson, Wagner, Swanson, & Raskind, 2006). Working with dyslexic subjects, the team identified unique and common brain regions involving phonological, orthographic, and morphological coding in working memory and reported that the processing of these three word forms contributed uniquely to the difficulties of dyslexia. In addition, they summarized studies that combined brain imaging before and after instructional treatment and confirmed that treatments targeting the three word forms promoted specific brain activation changes in dyslexic children (Richards et al, 2006). The Triple Word Form Theory provides a theoretical framework not only to determine the relations between the awareness of the three word forms and their parts, but also to explain how these word forms and their parts predict the development of word reading and spelling. Moreover, the cross-word-form mapping effects found in dyslexic children confirms the interactive nature among phonological, orthographic, and morphological knowledge while learning to read and spell (Berninger, Raskind, Richards, Abbott, and Stock, 2008; Richards et al, 2006). In the next section, we will talk about how error analysis untangles the interrelations among these components and how research design leads to different aspects of our understanding of spelling related linguistic skills.

**Spelling Error Analysis**

As stated previously, spelling is a precise measure of literacy-related skills because it requires attention to conventional forms which can be acquired through processing phonological,
orthographic, and morphological coding in working memory. To understand the use of linguistic knowledge, a direct examination of the spelling product is an informative method that provides a good general index of phonological, orthographic, and morphological knowledge and is particularly sensitive to the lack of this knowledge (Ehri, 2000; Treiman, 1998). Spelling error analysis provides a window into a writer’s understanding about his or her own language (Joshi, Treiman, Carreker, & Moats, 2008; Stone, Silliman, Ehren, & Apel, 2004). Treiman and Bourassa (2000) pointed out that traditional classroom spelling tests are typically scored according to a standard of correct or incorrect spelling. This scoring approach does not disclose how closely an error approaches the phonological, orthographic, or morphological structure of the target word. An analysis of spelling errors can reveal the utilization or deficiency of specific linguistic knowledge and deepen our understanding of spelling development. It is important to establish a systematic error coding system for educators to understand the function of each spelling subskill and also to improve spelling instruction accordingly.

When conducting an error analysis, a sufficiently representative spelling sample is needed to extract the most information from the misspellings (Apel, Masterson, & Brimo, 2010). Traditionally, several types of spelling samples have been collected for error analysis, such as free writing in a composition or response to a target word list. However, spelling samples collected from compositions might be biased due to the fact that while writing a student might select words that he or she feels more confident to spell. The analytic results might not completely reflect an understanding the writer has about written language. In addition, without considering word frequency and difficulty, the results might not be generalized to other conditions. For different research purposes, researchers may create experimental spelling measures in order to investigate specific aspects of spelling ability. These measures have been
developed based on interest in particular parts of words. Examples of aspects of interest are: syllable-initial consonant clusters (Treiman, 1991); final consonant clusters (Treiman & Cassar, 1996); spelling patterns (e.g., flaps in Treiman, Cassar, & Zukowski, 1994); or forms of words (e.g., derived words in Carlisle, 1985). Researchers have collected spelling samples using these experimental measures to answer particular research questions rather than to provide information about the typical development of spelling performance.

To assess general spelling performance, researchers may choose norm-referenced, criterion-referenced, or qualitative tests. These spelling tests mostly contain 40 to 50 words that are chosen based on the frequency or familiarity of use in content area textbooks utilized by school systems (Kaufman & Kaufman, 1985; Larsen, Hammill, & Moats, 1999; Woodcock, McGrew, & Mathers, 2001). However, researchers have found high variability in the linguistic structures among several commonly used spelling assessments and their alternate forms (when available) (Calhoon, Greenberg, & Hunter, 2010; Calhoon & Masterson, 2011). Calhoon and her colleagues (2010, 2011) pointed out that word lists in standardized spelling tests should be constructed carefully based on phonological, orthographic, and morphological features. Besides considering the issue of word choice, Moats (1994) recommended that a comprehensive spelling test should include as many items as needed to adequately sample the different types of spelling knowledge an individual possesses. Nevertheless, Masterson and Apel (2000) also suggested that 50 to 100 words may be required to adequately assess English spelling. However, it might be challenging for teacher to fit a long spelling test into their tight classroom schedule. Combining a short spelling test and an efficient coding tool makes it convenient for teachers to decipher the information behind the misspellings without spending a great deal of instructional time on testing.
As was previously stated, stage theorists have been using error analysis to investigate the strategies used in invented spellings (e.g., Bear & Templeton, 1998; Beers & Henderson, 1977; Gentry, 1982). For example, Nelson (1980) scored spelling errors as phonetically correct or incorrect. This method recognized misspellings that sounded the same as the target words as phonetically correct for children using invented spelling strategies. Lombardino, Bedford, Fortier and Carter (1997) analyzed the invented spelling samples of 100 kindergarteners who spelled 12 target words. They scored correctness based on whether each phoneme was represented by one grapheme. The spelling errors were then used to develop 10 spelling patterns and the 21 most frequent spellings of graphemes for the intervention. However, the definitions of phonetically or phonologically correct spellings were not consistent among these researchers. Hence, it was difficult to investigate whether students were using multiple types of linguistic knowledge or only phonological knowledge in their invented spellings. In addition, these results did not provide sufficient information about the typical development of spelling.

Meanwhile, other researchers categorized typically developing misspellings into several general types and assessed plausibility (e.g., Bebout, 1985). For example, Bebout (1985) collected 677 spelling mistakes from a fill-in-the-blank task completed by 2 groups of advanced learners of English: English-speaking children in grades 4 and 6, and Spanish-speaking adults studying English. She first assigned each mistake an acceptability code based on the answer to the following question: “If the misspelling was pronounced according to the most likely patterns of English, would the target word be the result?” (p.575). Next, each misspelling was assigned one or more of eight major error codes: (1) consonant-doubling error, (2) other consonant errors, (3) errors involving schwa /ə/, (4) errors involving silent ‘e’, (5) other vowel errors, (6) letter mis-ordering, (7) unclassifiable items, and (8) homophones (correct spelling). Although this
study provided some instructional suggestions for the most frequent incorrect spelling patterns, the major shortcoming of such coding categories is that the results do not reveal the cause of the misspellings.

Somewhat later, some researchers began to code spelling errors based on visual similarity and phonological accuracy (Lennox & Siegel, 1996). Using a scoring scheme similar to one designed by Bruck and Waters (1988), Lennox and Siegel (1996) introduced two systems for scoring phonological accuracy and visual similarity: 1) a constrained system, which includes position cues; and 2) an unconstrained system, in which position cues were irrelevant. Three scores were assigned to each misspelled word based on phonological and visual similarity. Visual accuracy is the percentage of the combination of correct bigrams and letters in a misspelling. According to the example provided by Lennox and Siegel (1996), the misspelling “heven” for heaven was scored as 3 correct bigrams and 5 correct letters. The visual accuracy score of the misspelling was the ratio of the misspelling score of 8 and the target spelling score of 11, or .73. They further analyzed the misspellings utilizing unconstrained and constrained systems for scoring phonological accuracy. A misspelling could be coded as phonologically correct in an unconstrained system but might be coded as incorrect in a constrained system due to the consideration of position (e.g., “kep” for keep). Considering the complexity of the coding scheme, it is difficult to utilize Lennox and Siegel’s (1996) method to analyze misspellings in a classroom setting.

It is important to point out that researchers have defined the representations of the use of phonological and orthographic knowledge in misspellings in various ways. It is commonly agreed that the integration of phonological and orthographic knowledge is necessary for good spelling. Researchers have typically identified spelling errors with a phonological basis as
phonetic errors. However, lack of precision might decrease the sensitivity of the coding system, and potentially cause more confusion. As mentioned earlier, Lombardino et al. (1997) took using a grapheme to represent a phoneme as the proof of using phonological knowledge. Other researchers included the plausibility (acceptability) or visual similarity standard while recognizing a phonetic spelling error (Bebout, 1985; Lennox & Siegel, 1996; Nelson, 1980). Moreover, several researchers also considered the position within words in their coding systems (Bruck & Waters, 1988; Lennox & Siegel, 1996). For instance, Moats (1983) included phonetic accuracy and serial order in her analysis, where serial order errors included letter order errors (i.e., “srtuck” for struck), duplications, and insertions. In this approach, it is difficult to determine the use of orthographic knowledge in a phonetic spelling error. Ideally, each category of the coding system should have a more precise definition based on the need to collect further information. Thus, having several distinct categories to identify the use of specific linguistic knowledge is preferable.

Alternatively, some researchers designed coding systems to examine the use of specific morphological knowledge in spelling. Carlisle (1996) investigated the use of morphologically complex words in the story writing of second and third graders. She created a coding system that first categorized different types of inflections and derivational forms and then coded errors into errors of omission (the lack of a morphological marker in an obligatory context) or errors of commission (use of incorrect morphemes or forms of words, misspelled morphemes). Green, McCutchen, Schwiebert, Quinlan, Eva-Wood, and Juelis (2003) used Carlisle’s (1996) morphological coding system to investigate the development of morphological skill in third- and fourth-grade students’ narrative writing. Both Carlisle (1996) and Green et al. (2003) found some developmental trends in the use of morphological knowledge. The results indicated that
inflectional morphology is commonly mastered by third grade and derivational morphology continues to develop in fourth grade. In addition, the use of morphological knowledge in spontaneous writing has unique predictive power on reading and spelling performance across second to fourth grades. Some other researchers coded the responses of dictation tasks that were purposely designed. For example, Hauerwas and Walker (2003) gave three groups of students from second to eighth grade a series of dictated spelling tasks consisting of past and progressive tense verbs in context and in isolation. They coded the responses based on phonological representation (the suffix was phonetically correct), morphological representation (the suffix was spelled correctly), and orthographic representation (correct application of the rules to add the suffix) of the spellings of inflected verbs.

Recently, some researchers started to include all three components of spelling – phonological, orthographic, and morphological knowledge -- into their error analyses (Apel & Masterson, 2001; Arndt & Foorman, 2010; Masterson, Apel, & Wasowicz, 2006). Masterson, Apel, and Wasowicz (2006) established the *Spelling Evaluation for Language and Literacy-2* (*SPELL-2*), which included a placement test at the beginning of the assessment to determine which one of four levels of spelling words to administer. For those patterns that are in error more than 40% of the time, they suggested additional analyses to determine which “language knowledge blocks” may be deficient and thus contribute most to the errors. Masterson et al.’s (2006) determination of phonological and orthographic errors is similar to other researchers’ except that they classified misspellings that were incorrect yet plausible or legal as representing an inadequate Mental Graphemic Representation (MGR). In terms of morphological errors, they classified misspellings for juncture modifications and affixes as having difficulties using phonological (for omissions), or morphological knowledge (illegal misspellings), or inadequate
MGRs (legal misspellings). Several researchers have applied Spell-2 for spelling error analysis in order to either compare the performance of different subgroups or to investigate spelling development at different levels (Hart, Scherz, Apel, & Hodson, 2007; Kelman & Apel, 2004; Apel & Apel, 2011).

During this same time period, Arndt and Foorman (2010) gave a 25-word dictated spelling test that followed similar word selection criteria used by Foorman and Ciancio (2005) to 60 second graders and coded the misspellings based on Masterson and Apel (2000) and Treiman (1993). The selected spelling patterns and words fell within 14 categories (e.g., r-controlled). Arndt and Foorman’s (2010) coding categories for linguistic features were: (a) phonological errors: a phoneme is not represented by a grapheme(s) in the spelling of a word (e.g., do/dog); (b) orthographic errors: letter(s) is used to represent a phoneme in a word that is not possible in English orthography; a spelling rule is not applied when required (e.g., bick/back); (c) orthographic image errors: legal/plausible but incorrect representation of the phoneme in that particular word (e.g., boyl/boil); (d) morphological errors: the prefix or suffix is omitted, misspelled, or when the suffix is added, the needed modification to the base is not spelled accurately (e.g., planing/planning); (e) transposition errors: the correct representation of phonemes is selected but two adjacent phonemes occur in the wrong letter sequence (e.g., nad/and).

To conclude, a set of representative spelling responses and an efficient coding system are the two keys needed to conduct a valid error analysis. As mentioned earlier, caution is needed when creating a dictation word list for data collection. In order to assess typical spelling performance, several aspects should be considered, including word frequency, linguistic complexity (e.g., Carlisle, 1985), and word familiarity (Ehri, 2000). With regard to spelling error
coding, researchers commonly agree with expanding the definition of spelling accuracy by adding the concept of legality into the error coding system (Arndt & Foorman, 2010; Bebout, 1985; Hauerwas, & Walker, 2003; Masterson, Apel, & Wasowicz, 2006). Legality refers to the plausibility or acceptability of a misspelled error that can be pronounced or can mean the same as the target word or word parts. A legal error occurs when a speller does not have sufficient orthographic knowledge, yet still tries to spell the target word using his or her phonological or morphological knowledge. A valid error coding system should be sensitive enough to recognize the utilization of specific linguistic knowledge in a misspelling. Given the consensus on the concept of legality, the lack of agreement on the definition of categories across coding systems not only leads to different research results but also invites more confusion. In addition, it is also important to note that seldom do the coding systems take the severity of errors into consideration. Previous coding systems might assign an error as legal or illegal orthographically yet fail to reflect the severity of the lack of phonological or morphological knowledge in a misspelling. A spelling error coding system with precise categories that can reveal the severity of the deficiency of spelling-related linguistic knowledge is needed. In the next section, we will further explore results of some contemporary spelling error analysis systems.

**The Cause of Poor Spelling**

Researchers have investigated and compared spelling errors among groups to enhance the understanding of individual differences in spelling acquisition. In general, poor readers perform worse than their peers on any type of spelling test. It is well known that reading difficulties are primarily caused by a deficiency in processing phonological information (e.g., Catts & Kamhi, 1998). Thus, it is reasonable to expect that poor readers rely more on their non-phonological knowledge in learning to spell and make more mistakes that are not phonetically accurate.
Researchers have studied the misspellings of poor readers to confirm this specific phonological deficit hypothesis (e.g., Bruck, 1988; Friend & Olson, 2008). On the other hand, Frith (1980) reported a unique group of poor spellers who performed fairly well in word reading, namely unexpectedly poor spellers. Aside from the fact that English spelling is more difficult than English reading, researchers have been performing tests to discern other reasons that could cause unexpectedly poor spelling (Bruck & Waters, 1988; Frith, 1980; Holmes & Castles, 2001). One hypothesis is that unexpectedly poor spellers suffer from a mild phonological deficit. However, Frith (1980) found a high percentage of phonologically acceptable spelling errors, indicating that unexpectedly poor spellers did not have phonological difficulties but struggled with an insufficiency of other spelling related knowledge. The discussion has been concerned with whether poor spellers are developmentally behind average spellers in particular component skills or whether they demonstrate a different skill pattern in their spelling.

Researchers have applied a match design to examine the differences in developmental paths by comparing groups of different ages matched on some aspect of reading or spelling performance level (Backman, Mamen, & Ferguson, 1984). A match design provides a powerful way to study atypical performance (Jackson & Butterfield, 1989) because it controls for the effects of skill level and exposure to printed material. The general purpose of a match design is to compare groups who took a greater or lesser number of years to achieve the same performance level and to investigate whether one of the groups had been using a different mix of skills. Aside from matching groups by chronological age, researchers typically paired older and younger students by reading, spelling, or both levels. A chronological-age-match design provides a comparison of children at a similar maturational level (Snowling, 1991). On the other hand, reading-level-match studies examine the specific phonological deficit hypothesis in spelling by
following a similar paradigm that has been well documented in reading, wherein older poor readers have shown significantly lower phonological decoding and/or phonological awareness when compared to younger, normally developing readers at the same absolute level of word reading or reading comprehension. As for the spelling-level-match design, it provides information about whether poor spellers exhibit a developmental lag or are deviant in spelling acquisition. Again, error analysis is commonly applied to determine the misspelling patterns among groups.

In order to examine the specific phonological deficit hypothesis, researchers pair older poor readers with younger normally developing students to understand if poor readers demonstrate severe phonological processing difficulties in their misspellings. Bruck (1988) found that 10-year-old students with dyslexia made a higher proportion of non-phonetic spelling errors than 7-year-old normal students did. As mentioned previously, Lennox and Siegel (1996) investigated the hypothesis using a similar scoring system created by Bruck and Waters (1988) where three scores were assigned to each misspelled word based on phonological and visual similarity. They found differences between younger and older students (i.e., between the ages of six and sixteen) at the same spelling grade level in the use of basic grapheme-phoneme correspondence rules. Curtin, Manis, and Seidenberg (2001) analyzed the spelling errors of third grade students with dyslexia and compared the results with reading-level- and age-matched groups. They classified spelling errors into two categories: phonological and orthographic errors. The constraints and acceptability were both considered. They also found that students with dyslexia produced spelling errors that showed poor knowledge of phoneme-grapheme correspondences. Recently, Friend and Olson (2008) calculated the grgraphotactic accuracy score in a similar manner as orthographic acceptability in Bourassa and Treiman (2003) by breaking
the attempted spelling into syllables, checking the legitimate existence and location of each 
syllable, and dividing the number of graphotactically plausible syllables over the total number of 
syllables for a given word. An analysis of variance on spelling errors showed that children with a 
spelling disability had significantly lower phonological accuracy than younger, normally 
progressing children. These results indicated that poor readers were more likely to rely on their 
non-phonological skills to spell and make less phonologically accurate misspellings than a 
younger normal comparison group.

However, the examination of the specific phonological deficit hypothesis did not turn out 
to be a one-sided result. By using a skill-level-match design, several researchers compared the 
spelling error types and found no evidence to support that the groups took a deviant 
developmental path in learning to spell (e.g., Schwartz, 1983; Varnhagen & Varnhagen, 1986, 
1992). Moats (1983) scored misspellings with a phonetic/non-phonetic approach and compared 
the results of fourth to eighth graders with dyslexia and normal second graders. No evidence was 
found to support that students with dyslexia produced more non-phonetic errors than the matched 
comparison group. Carpenter (1983) applied discriminant analyses to compare the spelling error 
patterns of older disabled readers and younger able readers. He found the difference between the 
two groups mentioned was less than that between disabled readers and their peers. Treiman 
(1997) reviewed several studies and concluded that the “non-phonetic” errors that children with 
dyslexia made were phonologically motivated and based on conventional English writing rather 
than random errors. As mentioned previously, Bourassa and Treiman (2003) scored spelling 
performance based on the degree of accuracy in phoneme representation, graphotactic 
convention, and conventional spelling. Even though they found subtle differences in specific 
types of errors (e.g., adding finale e in words with a short vowel), they did not find evidence
supporting a phonological deficit in spelling. Similarly, Cassar, Treiman, Moats, Pollo, and Kessler (2005) also found no differences between the misspellings made by 10-year-old students with dyslexia and 6-year-old typical students in either phonological or graphic acceptability. Moreover, Bourassa, Treiman, and Kessler (2006) further investigated whether morphology conditioned the difficulty in spelling and caused the differences between the strategies used by students with dyslexia and typical students. The result once again rejected the specific phonological deficit hypothesis.

Kamhi and Hinton (2000) reviewed several contemporary skill-level-match design studies. They pointed out that the inconsistency among studies was mostly due to the definition of phonetically acceptable spelling errors. They concluded that poor readers generally made more spelling errors than average readers, but their so called “non-phonetic” errors are phonologically motivated and based on conventional English writing. In Bruck and Treiman (1990), they found that students with dyslexia tended to omit the second consonants of word-initial clusters when compared to younger normally developing students (e.g., “tip” for trip). However, both group had difficulties producing legal spellings of syllables with initial clusters. The explanation of the omission was that students with dyslexia might attempt to spell by sound, but fail to capture all of the phonemes in the target words. Thus, it was not equitable to categorize such misspelling as non-phonetic errors. Evidently, further research on the hypothesis of specific phonological deficits in spelling is still needed. Fortunately, it is commonly agreed that a deficiency in the phonological process is one cause of difficulties in the acquisition of word reading and spelling. As noted in Cassar and Treiman (2004), the inconsistent results led us to explore a deficiency in other linguistic skills that might also cause poor spelling so that we can further understand the differences between the subgroups of poor spellers.
As noted earlier, some researchers have paid attention to unexpectedly poor spellers (Frith, 1980; Bruck & Waters, 1988; Bruck & Waters, 1990; Waters, Bruck, & Seidenberg, 1985). The main question about unexpectedly poor spellers is whether their misspellings shared similar patterns with poor spellers with low reading ability or good spellers. Unexpectedly poor spellers would be expected to make more phonetically accurate errors if the difficulty is due to the ambiguity of sound-spelling correspondences in English. On the other hand, if unexpectedly poor spellers equally experience the same phonological process deficiency as poor spellers with poor reading ability, their error pattern would be similar, exhibiting a smaller proportion of phonetically accurate misspellings than good spellers. Frith (1980) suggested that some fairly good readers might be unexpectedly poor spellers and identified a group of 12-year-old good readers with poor spelling skills. She analyzed their misspellings and discovered that unexpectedly poor spellers produced a high percentage of phonologically accurate misspellings and suffered from a lack of orthographic knowledge. In other words, unexpectedly poor spellers can represent phonemes in a word using corresponding graphemes, but fail to select the accurate one based on consideration of the environment within the word. Frith (1980) concluded that the unexpectedly poor spellers performed fairly in reading skills using a partial cue reading strategy and did not have complete mental representations of word spellings.

Kamhi and Hinton (2000) discussed the work of Waters, Bruck, and Seidenberg (1985) and explained the differences between Waters et al. (1985) and Frith (1980). Unlike Frith (1980) who found that unexpectedly poor spellers made significantly more phonetic errors than poor spellers, Waters et al. (1985) found that both groups had the same proportion of phonetic errors, which was significantly lower than that for the good spellers. Kamhi and Hinton (2000) pointed out that the methodological differences between the two studies included the age rage of the
subject, the measure of reading or spelling ability, and the definition of phonetically plausible errors. Bruck and Waters (1988) tried to experimentally control these confounding factors and found some conflicting results in third and six graders. Regardless of age or reading ability, they found that unexpectedly poor spellers and poor spellers had similar difficulty producing positionally-appropriate graphemes and produced proportionally fewer phonetic errors than good spellers. However, they found that older unexpectedly poor spellers had relatively good visual memory of words and knowledge of basic sound-letter correspondences. Recall the findings of Lennox and Siegel (1996): they found a significant difference in the use of basic grapheme-phoneme correspondence rules between poor spellers and the spelling-level-matched group. These results indicated that the use of basic sound-letter correspondence rules (i.e., orthographic knowledge) might be sensitive enough to classify poor spellers and younger average spellers, but certainly not sufficient to differentiate older poor spellers with different reading levels. Thus, we may conclude that a spelling error coding system should include consideration of the positional constraints within words in order to efficiently diagnose spelling difficulties.

The Present Study

Spelling is the foundation of writing and shares common linguistic knowledge with reading (e.g., Berninger, 2000; Ehri, 2000). The purpose of this study was to investigate the use of linguistic knowledge in spelling by analyzing spelling errors made by fourth, fifth, and seventh grade students. The spelling outcomes of subgroups were compared based on both students’ age and reading levels. Many researchers have been working on invented spelling or emerging spelling development (e.g., Read, 1971; Gentry, 1982). Only a few studies have been conducted on later spelling development. It is commonly agreed that learners in intermediate grades are in a transitional phase. They are shifting from “learning to read” to “reading to learn”
(Chall, 1983). The more readers are exposed to literature, the more they might be able to apply the linguistic knowledge that they have learned in their spelling. According to the Common Core State Standards in English Language Arts, students above fourth grade are expected to accurately spell using the conventions of the English written language (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). However, little is known about the interrelations among phonological, orthographic, and morphological representations and the impact of such relations on spelling development (Treiman & Bourassa, 2000). By investigating the spellings of the target age group in this study, it was predicted that further understanding of the components of spelling knowledge and development would be achieved.

As mentioned previously, error analysis of spelling errors provides a window for researchers and educators to understand how closely a misspelling approaches the phonological, orthographic, and morphological structure of the target word (e.g., Treiman, 1998). The morpheme-dominant nature of the English written language directs the development of the error coding system in this study. Words were separated into their morphological elements before any graphophonemic analysis was performed. In the error coding system, the affix and the juncture, if applicable, in a misspelling were scored under the category of Morphological Legality. Next, each representation of the phonological information in the base word was evaluated in the category called Phonological Representation. Finally, the graphemes that were represented were judged to see if they were legally following the English orthographic conventions in a category called Orthographic Legality. The sequential and weighting approach of the coding system reveals the severity of the deficiency in each category. Considering the spelling demands of words students are asked to write above fourth grade, the researcher manipulated the
morphological complexity of the target words used in the spelling test. Finally, the spelling test
and error coding system were created to capture the linguistic features of misspellings of words
with diverse familiarity, frequency, and complexity.

Several theoretical hypotheses of reading- or spelling-level-matched design have been
described previously (e.g., Bruck, 1988; Frith, 1980). In this study, the grade- and skills-level-
match designs were applied to answer the following research questions:

1. How do the fourth, fifth, and seventh grade students perform on the prior year’s Florida
   Comprehensive Assessment Test (FCAT) 2.0 Reading Developmental Scale Scores (DSS,
   Florida Department of Education, 2012) and on the researcher-designed spelling test? What are
   the linguistic profiles of the spelling errors made by these students?
2. How do the poor readers perform on the spelling test? How do the poor readers’ linguistic
   profiles of misspellings differ from their (grade-matched) peers with average reading ability?
3. How do the linguistic profiles of the older poor readers differ from those normally progressing
   younger (reading-level matched) readers?
4. What is the incidence of unexpectedly poor spellers in the fourth, fifth, and seventh grades?
   How do their linguistic profiles differ from those of students matched on reading or spelling
   ability?
5. How do item characteristics (i.e., word frequency and morphological complexity) and subject
   characteristics (i.e., grade, reading ability, and word familiarity) affect spelling outcomes?
CHAPTER TWO

METHOD

Participants

Approximately 220 of the 4th, 5th and 7th grade students at two northern Florida elementary and middle schools with parent consent and child assent were included in the present study. Schools with moderate to high percentages of students qualifying for free and reduced price lunches, and with 2% or less students identified as English Language Learners, were targeted. The schools were chosen to match the expected sample size and state-level reading test score distribution necessary for the current study. There were 72 (38 boys, 34 girls), 73 (34 boys, 39 girls), and 75 (32 boys, 33 girls) students in each grade, respectively. The demographic characteristics of the students were 44% African-American, 32% Caucasian, 12% Hispanic, 2% Asian, and 10% Multiracial.

Several grouping criteria were applied to answer the research questions. FCAT 2.0 Reading Developmental Scale Scores (FCAT 2.0 Reading DSS, Florida Department of Education, 2012) from the previous academic year were used to identify poor (at and below the 20th percentile) and average (at and above the 40th percentile) readers. The percentile ranks were constructed from the FCAT 2.0 Reading DSS by the researcher’s colleague (see Appendix A). The Spelling Accuracy Score from the researcher-designed spelling test was also used to identify poor spellers (at and below the 20th percentile). To answer the third research question, poor readers in 7th grade (n=13) were matched with students in 4th grade (n=22) with normal reading ability (i.e., at and above the 40th percentile) based on the FCAT 2.0 Reading DSS. Unexpectedly poor spellers were students with normal reading ability (at and above the 40th percentile in the FCAT 2.0 Reading DSS) but poor spelling performance (at and below the 20th percentile in the
researcher-developed spelling test). Table 1 displays the distributions of subjects in different groups.

Table 1

<table>
<thead>
<tr>
<th>Grades</th>
<th>Average Readers</th>
<th>Poor Readers</th>
<th>Unexpectedly Poor Spellers</th>
<th>Poor Spellers</th>
<th>Poor in Both</th>
<th>Average Spellers</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 4</td>
<td>59 (81.9%)</td>
<td>6 (8.3%)</td>
<td>12 (16.7%)</td>
<td>18 (25.0%)</td>
<td>3 (4.2%)</td>
<td>10 (13.9%)</td>
<td>72</td>
</tr>
<tr>
<td>Grade 5</td>
<td>50 (68.5%)</td>
<td>13 (17.8%)</td>
<td>8 (11%)</td>
<td>15 (20.5%)</td>
<td>5 (6.8%)</td>
<td>9 (12.3%)</td>
<td>73</td>
</tr>
<tr>
<td>Grade 7</td>
<td>56 (74.7%)</td>
<td>13 (17.3%)</td>
<td>8 (10.7%)</td>
<td>17 (22.7%)</td>
<td>8 (10.7%)</td>
<td>9 (12.0%)</td>
<td>75</td>
</tr>
</tbody>
</table>

Note. Sample size and the percentage for each group. a Average spellers matched on reading ability of unexpectedly poor spellers

Materials and Procedure

Researcher-designed spelling test

A 25-word dictated spelling test was used to assess spelling performance across all groups. Items were selected from the Academic Word List (Coxhead, 2000). Since the 570 words included in this list are grouped into 10 sublists that reflect word frequency, words were selected from each sublist to ensure the variation of word frequency across the 25 items. For word type, morphological complexity was considered while forming the spelling test. Half of the words on the spelling test were less complex and half were more complex. A base word without any affixes (e.g., route) or words requiring no spelling change with the addition of an affix (e.g., restrict) were categorized as less morphologically complex. Conversely, words with more than 2 affixes (e.g., subsequently) or words requiring a spelling change when an affix is added (e.g., occur to occurrence) were considered highly morphologically complex. A list of the target words
and sentences used in the spelling test can be found in Appendix B. The spelling test was group-administered using a paper-pencil format. The researcher read out loud a script containing child assent and test instructions. A trained native English speaker followed the instructions and administered the spelling test. Each item was first read aloud, then repeated in the context of a sentence, and then read aloud again. Students were asked to check if the word was familiar to them or not, and then asked to spell the word they heard on a blank sheet of paper. Students had around 5 seconds to respond to each item. No repetitions of the target words were allowed.

Error coding system

The spelling responses on the researcher-developed spelling test were first marked as correct or incorrect. Each incorrect response was coded based on its linguistic features and then scored cumulatively in 3 categories: Phonological Representation (PR), Orthographic Legality (OL), and Morphological Legality (ML). The sum of error points in each category reflects the severity of the linguistic knowledge deficit shown in a misspelling. In other words, each misspelling was coded and assigned 3 cumulative scores ranging from 0 to the maximum number of possible error points. As shown in Table 2, the coding system was designed to be implemented sequentially in the order of ML, PR, and OL. Each category weights the linguistic features in a misspelled grapheme or morpheme differently by assigning the features 1, 2, or 3 points.

For each misspelling, the base word (or root if applicable) and the affix (and the juncture if applicable) were first identified. The later part of the misspelling was coded in the category of Morphological Legality (ML). The unit of error analysis in the ML category is the morpheme in order to determine the use of morphological knowledge in a misspelled word. A morpheme is the most elemental unit that forms the meaning of a word. In the current error coding system, only
affixes and junctures in a misspelling were coded. An affix is a morpheme attached before or after a base word or a root to modify its meaning (or sound). A juncture is the modification of a base word when attached by an affix (e.g., when transforming *reduce* to *reduction*, one needs to eliminate *e*). Taking one of the target words in the spelling test, *commission*, for example, there are one affix, *ion*, and one juncture, *ss*, which are the two units that should be coded in a misspelling. Each affix or juncture was coded in a weighted fashion. The error points of all the affixes and junctures in one misspelling were summed as the total error points.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>The Error Coding System</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coding Target</td>
<td>Step 1: Morphological Legality (ML)</td>
</tr>
<tr>
<td>Coding Unit</td>
<td>All of the affix(es) and juncture(s) if applicable</td>
</tr>
<tr>
<td>Coding Unit</td>
<td>Morpheme</td>
</tr>
<tr>
<td>3 point</td>
<td>Omission/Addition</td>
</tr>
<tr>
<td>2 point</td>
<td>Illegal spelling error</td>
</tr>
<tr>
<td>1 point</td>
<td>Legal error: Alternative affix or juncture following morphological rules</td>
</tr>
<tr>
<td>Total score</td>
<td>Cumulated error points of each suffix and juncture if applicable in a misspelling</td>
</tr>
</tbody>
</table>
According to the findings of Carlisle (1996) and Green et al. (2003), inflectional morphology is commonly mastered by third grade and derivational morphology continues to develop in fourth grade. Considering the age of the target subjects in this study, an omission or addition of a morpheme was evaluated as a more severe error in a misspelling; therefore, the omission/addition was weighted with 3 points. In the example word, *commission*, an omission of *ss* or *ion* was scored as 3 points. To determine the legality of a misspelled affix or juncture, one needs to understand the opacity of English morphology and both the grammatical and semantic function that a morpheme might serve. A morpheme might contribute multiple meanings to different words (e.g., *re* means “again” in *recharge* and means “back” in *revoke*). On the other hand, one grammatical or semantic function can be fulfilled by different morphemes (e.g., both *ance* in *appearance* and *ence* in *difference* serve as a noun suffix to form a noun from a verb). Considering the opacity of English morphology, a speller might inaccurately choose an alternative morpheme to serve the original grammatical or semantic function in a word. Thus, a misapplied alternative morpheme was considered as morphologically legal and coded as 1 error point. For a misspelled morpheme, any other attempts were considered illegal and were coded as 2 error points. The example word, *commission*, might be commonly misspelled as *commition* or *commision*. Even tough -*tion* and -*sion* could both possibly be used to form a noun, just like -*ssion* in *commission*, English language only allows the sound /ʃən/ to be spelled as ssion in words containing *mission*, such as *omission* and *permission*. As the result, using either *t* or *s* to replace *ss* in *commission* would be considered as illegal and be scored as 2 points.

After the affix and the juncture of a misspelling were coded in Morphological Legality, the remaining part (i.e., base word) was coded in the categories of Phonological Representation (PR) and Orthographic Legality (OL). A base word is a free morpheme that can be expanded
with affixes to produce longer new words. However, some longer words may contain a root, a bound morpheme, which can only appear as part of a larger word and be used to form a family of words with related meanings (e.g., clar in clarify or declaration; Moats, 2010). Under the superordinate category which contained PR and OL, a base word or a root that carried the principal meaning was analyzed after removing all affixes and junctures. The coding unit was changed from morpheme to grapheme. In the example word, commission, there are four sounds in the main morpheme, commi, needed to be represented using four graphemes. It is important to note that it is possible to use more than one letter in a grapheme, such as mm in commission is representing the sound /m/. By applying a sequential coding approach, each grapheme within the main morpheme was analyzed sequentially so that the penalty was not given twice for the same reason across the 2 categories.

In the category of Phonological Representation (PR), a response was only penalized when a sound in the base or root word was under- or over-represented by one or a set of letters. Error points were summed after examining each grapheme in a misspelling. The error coding system in this study applied a more stringent approach to define each category. Based on the work of Lombardino, Bedford, Fortier, and Carter (1997) in invented spelling, an excessive use of graphemes was considered to be more serious than the omission of a phoneme. Thus, the misrepresentation of a phoneme was coded as 1 or 2 points accordingly. Each phoneme that was not represented with a grapheme was considered as an under-represented phoneme and was coded as 1 point (e.g., cake is spelled as “ca_” or “ko_”). An over-represented phoneme was coded as 2 points when a speller used an unnecessary additional grapheme to represent a non-existent phoneme (e.g., cake is spelled as “scack”). It is also important to note that the use of more than one letter as a grapheme to represent a phoneme was not penalized because an over-
represented phoneme might comprise two, three, or even four written letters. The weighted coding approach aimed to provide information regarding the severity of the deficiency in the use of phonological knowledge.

Following the sequential coding principle, none of the over-represented graphemes were penalized an additional time in the category of Orthographic Legality (OL). In the OL category, only graphemes used to represent existing phonemes were coded and scored cumulatively. By precisely separating the PR and OL categories, the error coding system in this study provided a clearer picture of the use of phonological and orthographic knowledge in misspellings.

Regarding the severity of the lack of orthographic knowledge, misspelling a grapheme by using an alternative English spelling pattern was considered to be a less egregious mistake than misspelling with random graphemes. To determine the legality of a misspelled grapheme, one needs to follow English orthographic rules strictly. Due to the opacity of English orthography, it is important to acknowledge both the position of a grapheme and the environment within a word. A misspelled grapheme that may sound equivalent was not taken as legal if it was at the wrong position in a word. In the OL category, a legally misspelled grapheme was coded as 1 point, while an illegally misspelled grapheme was coded as 2 points.

For further delineation of the OL category, the legality of a misspelled grapheme was determined by the position and the environment within a word. For example, when *cake* was spelled as “caik,” *ai* in the misspelled response was coded as orthographically legal (i.e., 1 point) because *ai* can have the same pronunciation as a long *a* in the middle of a word without a silent *e* at the end. On the other hand, “cace” as a misspelled response to *cake* was coded as illegal with 2 points because the second *c* can only be pronounced as /s/ before *e* and was not phonologically equivalent to /k/. Another illegal misspelling example of *cake* was “ckake”. Even though *ck*
sounds /k/, which is equivalent, it is not allowable to put ck at the beginning of a word in English orthography. Thus, in the examples of “cace” and “ckake”, the words cannot be pronounced equivalently or they violate orthographic rules after considering the location and the environment within the word.

**Inter-rater reliability**

To ensure high inter-rater reliability, a second rater was trained by the researcher. The rater practiced scoring responses of 5 randomly selected students first and discussed disagreements in the error coding with the researcher until 100% agreement was reached. To obtain inter-rater reliability, 10% (i.e., the spelling responses of 20 students) of the data were randomly selected and scored a second time. Ottenbacher (1986) suggested a point-by-point approach to establish inter-rater reliability. In this method, each item would be rated as an agreement or disagreement between raters. To calculate percentage of agreement, the number of agreements between the two raters was divided by the total number of responses. In the current study, 90% agreement between the researcher and the second rater was considered adequate. The results of point-by-point interrater reliability were: 99.03% for Morphological Legality; 95.12% for Phonological Representation; and 93.35% for Orthographic Legality.

**Data Analysis**

In the current study, descriptive analysis and Profile Analysis (Tabachnick & Fidell, 1996) were utilized to answer the research questions. To answer the first research question, descriptive statistics were reported and followed by tests of normality and homogeneity of variance to ensure that test assumptions were met. Since each cumulative score from the three-category spelling error coding system was measuring the same cognitive ability (i.e., spelling), Profile Analysis (Tabachnick & Fidell, 1996) was used to examine the differences across group
profiles to answer the second, third, and fourth research questions. For the fifth question, Hierarchical Linear Modeling (Bryk & Raudenbush, 1992) was used to explain the crossed random effects of item- and subject-level characteristics on spelling outcome.

Profile analysis has been widely used by researchers in different fields to provide information regarding unique profiles of target observed data (e.g., Mooney, Leventhal, Hatsukami, 2006; Eyduran, Yazgan, Ozdemir, 2008). As Tabachnick and Fidell (1996) pointed out, profile analysis is an effective way to "compare profiles of two or more groups measured on several different scales, all at one time" (p. 391). Profile analysis is the multivariate equivalent of repeated measures or mixed-design ANOVA when there are multiple dependent variables on the same scale (Tabachnick & Fidell, 1996). In this study, the within-subject factor was the result of the coding categories with 3 levels, and the between-subject factor was the grouping variable with 2 or 3 levels accordingly. The profile plots will be presented and compared to test the hypotheses.

The three major hypotheses of profile analysis are the equal levels, parallelism, and flatness hypotheses (Morrison, 1976). If any of the hypotheses are rejected, follow-up tests will be conducted accordingly. The equal levels hypothesis states that regardless of whether individual scores differ across the 3 categories, the group averages are the same. In the current study, it was hypothesized that older or average readers scored fewer total spelling error points than younger or poor readers. To answer the research questions, the number of subgroups for each comparison might include two groups or six groups accordingly (e.g., poor readers and average readers in 3 grades for the second research question). The group averages over the 3 categories and the total accurate scores were calculated and one-way analysis of variance (ANOVA) was used to test the hypotheses. Rejection of the equal levels hypothesis implies that
the profile of at least one group, although possibly parallel to the others, has a significantly different average score for the 3 categories than at least one of the others. In this case, contrasts among group means were implemented to determine which group or groups were significantly different.

The parallelism hypothesis is generally the primary focus in profile analysis. The parallelism hypothesis states that the patterns across the 3 coding categories are the same for each group, although some groups may be elevated or depressed in comparison with others. This technique allows discrimination of overall effects (rejection of the equal levels hypothesis) from interactions between within- and between-subject factors (rejection of the parallelism hypothesis). Non-parallelism indicates that the shape of the profile for at least one group differs from the others. Tests of simple effects within a factorial ANOVA framework were used to describe the category effect within each level of group effect. Scheffé and Bonferroni’s critical values were calculated and the smaller one was chosen to control for Type I error. It was expected that the interactions between group and category factors would be significant.

In a traditional factorial ANOVA, the significance of interaction effects determines the type of follow up tests that are used. When the interaction is not significant, a main effect test and a post hoc test are conducted. On the other hand, a significant interaction leads to tests of simple effects while ignoring the main effects. In profile analysis, the flatness hypothesis states that no significant change exists within groups across the three categories. Thus, testing the flatness hypothesis is the same as testing the within-subjects main effect no matter whether the test of parallelism is significant. Rejection of the flatness hypothesis implies that the profile of at least one group changes between any two of the three categories. It was expected that older students or average readers would have higher scores in the Orthographic Legality than in the
Phonological Representation category.

The final question was to investigate the effects of subjects and items on coding category scores. In the current study, the scores of the three coding categories were affected by the sampled subjects and items. In order to generalize the results beyond the particular sample to a larger population of interest, the random effects of subjects and items both need to be taken into account in data analysis. Since every subject responded to the same 25 items, the random effects were crossed, creating a mixed-effects model. Mixed-effects modeling (Quene & van den Bergh, 2008) allows for a full analysis with multiple random factors. The method is also known as the multilevel model, hierarchical linear model (HLM), or variance component model. HLM can capture dependency among individuals that is ignored in an ordinary regression by adding a random effect for each group so that the unique group-level contribution can be counted.

The purpose of the fifth question was to investigate the item-level and subject-level effects on the spelling performance. A mixed-effects model was established using the Spelling Accuracy Score of each item from each student as an outcome variable. The random effects of subject and item factors were analyzed to answer the research question. In this study, subject and item factors were completely crossed, which means that there was exactly one observation for each subject on each item. In order to estimate the predictive power of the subject and item characteristics, several fixed effect factors were introduced to the models. At the item level, fixed factors included Word Frequency (WFR) and Word Type (WT). Each item had a WFR value determined by the sublist from the Academic Word List (Coxhead, 2000). The WFR value ranges from 1 to 10, which is continuous. A WT value reflects the degree of morphological complexity of each item, which was be scored dichotomously. At the subject level, Grade (G), Reading Ability (RA), and Word Familiarity (WFA) were included as predictors. FCAT 2.0
Reading Developmental Scale Scores (Florida Department of Education, 2012) were used to represent the Reading Ability of each subject. WFA was collected based on the self-reported responses of subjects which was dichotomous (“This word is familiar to me” vs. “This word is not familiar to me”). The unconditional model was as follows:

Level 1

\[ Y_{ijk} = \pi_{0jk} + e_{0jk}; \quad e_{0jk} \sim N (0, \sigma^2) \]

Level 2

\[ \pi_{0jk} = \theta_0 + b_{00j} + c_{00k} + d_{0jk}; \quad b_{00j} \sim N (0, \tau_{b00}^2); \quad c_{00k} \sim N (0, \tau_{c00}^2); \quad d_{0jk} \sim N (0, \tau_{d00}^2) \] (1)

At level 1, \( Y_{ijk} \) is the response \( i \) of subject \( k \) on item \( j \), and \( \pi_{0jk} \) is the cell mean across all subject and items. \( e_{0jk} \) is the deviation of each response from the cell mean and is assumed to be normal distribution. At level 2, \( \theta_0 \) is the grand mean of all subjects across all items. \( b_{00j} \) indicates the item random effect, and \( c_{00k} \) indicates the subject random effect. Finally, \( d_{0jk} \) is the random interaction of item and subject. All three of the random parts at level 2 are also assumed to be normally distributed. In the current study, the researcher hypothesized that the results from the conditional model using one of the three spelling error coding categories (i.e., Phonological Representation) as the outcome would be as follows:

Level 1

\[ \text{PRERROR}_{ijk} = \pi_{0jk} + e_{0jk}; \quad e_{0jk} \sim N (0, \sigma^2) \]

Level 2

\[ \pi_{0jk} = \theta_0 + \beta_0X_k + \gamma_0W_j + b_{00j} + c_{00k} + d_{0jk}; \quad b_{00j} \sim N (0, \tau_{b00}^2); \quad c_{00k} \sim N (0, \tau_{c00}^2); \quad d_{0jk} \sim N (0, \tau_{d00}^2); \] (2)

where \( X_k \) indicates the subject fixed factors, and \( W_j \) indicates the item fixed factors.
CHAPTER THREE

RESULTS

The purpose of this study was to investigate the use of linguistic knowledge in spelling by analyzing spelling errors made by fourth, fifth, and seventh grade students. Spelling performance was assessed by a 25-word dictated spelling test across all grades. Each misspelling was coded based on its linguistic features and scored cumulatively in 3 categories: Phonological Representation (PR), Orthographic Legality (OL), and Morphological Legality (ML). The error coding system revealed the linguistic profiles of misspellings produced by intermediate and higher grade students. In order to determine the significant differences that existed among subgroups based on multiple grouping criteria, appropriate statistical techniques were utilized. Moreover, the current study intended to examine the effects of item/subject characteristics on spelling performance.

Five research questions were proposed in this study. In order to untangle the complicated relations among multiple factors, some grouping variables and matching designs were applied. To answer the first question, an overall understanding of the spelling performance of all students at different grade levels was achieved by utilizing descriptive statistics. Several profile analyses were run to answer the second, third, and fourth questions. The purpose of the second question was to compare the results of poor readers and their peers with those of average reading ability. The results of the comparisons were expected to provide some insight into general spelling development at different ages. To answer the third research question, a match design was utilized to pair poor readers in seventh grade with normally progressing readers in fourth grade. By comparing these two reading-ability-matched groups, the study examined the specific phonological deficit hypothesis (e.g., Bruck, 1988; Friend & Olson, 2008) in spelling. In terms of
unexpectedly poor spellers as mentioned in the fourth question, this study proposed to clarify the use of orthographic knowledge in the spelling of readers with discrepant spelling ability. Finally, the fifth question focused on the effects of manipulating item characteristics in the researcher-designed spelling tests and subject-related variables.

1. How do fourth, fifth, and seventh grade students perform on the prior year’s Florida Comprehensive Assessment Test (FCAT) 2.0 Reading Developmental Scale Scores (DSS, Florida Department of Education, 2012) and on the researcher-designed spelling test? What are the linguistic profiles of the spelling errors made by these students?

The purpose of the first question was to have a general understanding of the distribution of spelling performance among the three grades. The results of the prior year’s FCAT 2.0 Reading DSS showed that students in higher grades performed better than younger students in general, as can be seen in Table 3. None of the results from the Shapiro-Wilk normality test indicated a need for concern about the skewness in the data ($p = .123, p = .447, p = .698$, respectively). According to the percentile ranks constructed from the norm of the FCAT 2.0 Reading DSS in 2012 (see Appendix A), the participants in this study performed slightly better than the means of the state norms. After confirming the homogeneity of variances using Levene’s Test of Equality of Variance ($p = .19$), a one way ANOVA showed that the increase of FCAT 2.0 Reading DSS across the three grades was statistically significant, $F (2, 194) = 22.44, p < .0001$. According to the post-hoc Scheffé test, the seventh graders outperformed the fifth graders, $M$ difference $= 13.70, p < .0001$, and the fifth graders outperformed the fourth graders, $M$ difference $= 8.82, p = .04$. 


Table 3
Means and Standard Deviations for All Scores of the Students in the Three Grades

<table>
<thead>
<tr>
<th></th>
<th>Grade 4 n = 72</th>
<th>Grade 5 n = 73</th>
<th>Grade 7 n = 75</th>
<th>All n = 220</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCAT 2.0 Reading DDS</td>
<td>207.78 18.02</td>
<td>216.37 18.14</td>
<td>230.97 21.76</td>
<td>218.54 21.59</td>
</tr>
<tr>
<td>Spelling Accuracy Score</td>
<td>6.35 4.50</td>
<td>9.48 4.90</td>
<td>13.31 4.21</td>
<td>9.76 5.35</td>
</tr>
<tr>
<td>Error Coding System</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phonological Representation (PR)</td>
<td>14.08 15.01</td>
<td>11.03 15.65</td>
<td>5.59 7.65</td>
<td>10.17 13.63</td>
</tr>
<tr>
<td>Orthographic Legality (OL)</td>
<td>29.61 13.09</td>
<td>23.45 11.33</td>
<td>14.95 7.38</td>
<td>22.57 12.36</td>
</tr>
<tr>
<td>Total Error Score</td>
<td>67.83 37.46</td>
<td>51.39 37.00</td>
<td>29.64 16.30</td>
<td>49.36 35.24</td>
</tr>
</tbody>
</table>

Note. M = Mean; SD = Standard Deviation.

Figure 1. Histogram of the Spelling Accuracy Scores for the Fourth Grade Students

As shown in Table 3, the average Spelling Accuracy Scores on the 25-item spelling test also increased by grade level. It is noticeable that the researcher-designed spelling test was difficult for most of the students in this study even with typical reading ability in the different
grades. According to the results of the Shapiro-Wilk normality test, the data of fourth grade students showed a statistic of .95 and a significance of .007, indicating that the distribution violated the assumption. Figure 1 illustrates the histogram of the Spelling Accuracy Score in fourth grade students with the range of 0 to 21 and demonstrates a floor effect, indicating the spelling test was too difficult for fourth graders. The result of Levene’s Test of Equality of Variance shows that the data match the assumption ($p = .476$). A one way ANOVA also showed that the increase across three grades was statistically significant, $F (2, 194) = 37.293$, $p < .0001$. According to the post-hoc Scheffé test, the seventh graders outperformed the fifth graders, $M$ difference = 3.86, $p < .0001$, and the fifth graders outperformed the fourth graders, $M$ difference = 2.94, $p < .0001$.

As mentioned previously, the Spelling Accuracy Scores increased by grade levels. Since each misspelling was scored in the error coding system, it was reasonable to expect a gradual decrease by grade level in the Total Error Score (i.e., the sum of the cumulative scores from the three coding categories). The results confirm this expectation, as can been seen in Table 3. Unlike the FCAT 2.0 Reading DSS and Spelling Accuracy Score, the scores from the error coding system were not expected to fit either the normality test or the homogeneity of variance. In fact, substantial variance in the scores was preferred. As expected, the standard deviation values of the Total Error Scores were rather large, which indicates a wide range of scores. The result of Levene’s Test of Equality of Variance was significant ($p < .0001$), indicating unequal variances across the three grades. The results of the Robust Tests of Equality of Means were both significant, Welch’s $F (2, 108.94) = 31.20$, $p < .0001$ and Brown-Forsythe’s $F (2, 149.36) = 21.42$, $p < .0001$. We can conclude that at least two of the three grades differed significantly on their Total Error Scores. According to the post-hoc Games-Howell test, the seventh grade
students produced significantly fewer Total Error scores than the fifth and fourth graders did, both with \( p < .0001 \).

A Profile analysis was performed on the three dependent variables--Morphological Legality, Phonological Representation, and Orthographic Legality--to compare the profiles of the three grades. The means and standard deviations for each dependent variable for each grade are shown in Table 3. The repeated measures module under GLM in SPSS was used for the profile analyses. While applying a multivariate analysis of variance, the technique takes the correlation among the different levels into account. Hence, the assumptions of sphericity and homogeneity of treatment difference variances are not required. The first test in a profile analysis is for equal level. The tests of between-subject effects show whether there is a significant difference between groups after averaging across all dependent variables, which suggests rejection of the equal level hypothesis. The results show significant differences among the three grades, \( F(2, 217) = 26.91, p < .0001 \). A univariate ANOVA was performed using the average of the three categories as the dependent variable to see where the differences occurred. The result of Levene’s Test of Equality of Variance was significant \( (p < .0001) \), indicating unequal variances across the three grades. The results of the Robust Tests of Equality of Means were both significant, Welch’s \( F(2, 121.70) = 37.26, p < .0001 \) and Brown-Forsythe’s \( F(2, 168.22) = 26.57, p < .0001 \). According to the post-hoc Games-Howell test, there was a significant mean difference between the fourth and fifth grades, \( M \) difference = 5.48, \( p = .02 \); and between the fifth and seventh grades, \( M \) difference = 7.25, \( p < .0001 \).

For the flatness test, the results of the multivariate tests showed that the scores from the error coding system for the three categories deviated significantly from the flatness among the three grades. The results of the Pillai’s Trace Test were \( F(2, 216) = 108.25 \) with \( p < .0001 \). This
indicates that the performances from the three grades were significantly different across the three categories in the error coding system. Figure 2 graphed the linguistic profiles of the three grades. Visual inspection of the plotted profiles indicates that students in the three grades made more error points in Orthographic Legality and Morphological Legality than in Phonological Representation. The third test in a profile analysis is for parallelism. This is found through the interaction term in the multivariate test box, also using Pillai’s Trace values. The significant result, $F(4, 434) = 3.54, p = .007$, indicates that the students from the three grades performed differently over the three categories in different ways. As can been seen in Figure 2, the profile of the fourth grade demonstrated the steepest slopes across Morphological Legality, Phonological Representation and Orthographic Legality among the three grades.

![Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation, and OL for Orthographic Legality) in the Three Grades](image)

*Figure 2. Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation, and OL for Orthographic Legality) in the Three Grades*

As the analysis revealed, it was not easy to interpret the significant interaction of the profile as shown in Figure 2. The researcher was interested in whether the ratios of scores from
the three categories to Total Error Score might provide more information for understanding the linguistic characteristics in the misspellings. Three new dependent variables were converted from the raw error score of each category divided by the Total Error Score (viz., Ratio of Morphological Legality to Total Error Score, Ratio of Phonological Representation to Total Error Score, and Ratio of Orthographic Legality to Total Error Score). In Figure 3, visual inspection showed that the profiles from the three grades all had the highest ratios in Orthographic Legality and the lowest ratios in Phonological Representation, regardless of grades.

Figure 3. Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation, OL for Orthographic Legality, and TES for Total Error Score) in the Three Grades.

2. How do the poor readers perform on the spelling test? How do the poor readers’ linguistic profiles of misspellings differ from their (grade-matched) peers with average reading ability?
To answer the second research question, the prior year’s FCAT 2.0 Reading DSS (Florida Department of Education, 2012) were used to group students in the three grades. Students with the FCAT 2.0 Reading DSS at or below the 20th percentile were considered poor readers while those with the FCAT 2.0 Reading DSS at or above the 40th percentile were considered average readers. According to the percentile ranks constructed from the norms of the FCAT 2.0 Reading DSS in 2012 (see Appendix A), scores of 182 and 195 represented the 20th and 40th percentiles in third grade; scores of 196 and 208 represented the 20th and 40th percentiles in fourth grade; and, scores of 206 and 219 represented the 20th and 40th percentiles in sixth grade. Table 4 demonstrates the means and the standard deviations of all the measures for average and poor readers in these three grades.

Because the FCAT 2.0 Reading DSS were used to identify groups, it was expected to find that the average readers outperformed the poor readers in the FCAT 2.0 Reading DSS. Homogeneity of variance was assessed by Levene’s Test for Equality of Variances with \( p = .10 \) for fourth graders and \( p = .09 \) for seventh graders. The results of the \( t \) tests for these two grades showed that average readers outperformed the poor readers with the differences of 37.07 (\( t(63) = 6.09, p < .0001 \)) in fourth grade and 42.87 (\( t(67) = 9.65, p < .0001 \)) in seventh grade. Because homogeneity of variance in fifth grade was questionable (\( p = .015 \)), a \( t \) test that did not assume equal variance was conducted (\( t(67) = 9.65, p < .0001 \)). The significant result also indicated a difference between average and poor readers in the FCAT 2.0 Reading DSS. In addition, the FCAT 2.0 Reading DSS of poor readers increased significantly by grade level (\( F(2, 29) = 14.01, p < .0001 \)).
Table 4
Means (Standard Deviations in parentheses) for All Scores of the Average and Poor Readers in the Three Grades

<table>
<thead>
<tr>
<th></th>
<th>Grade 4</th>
<th>Grade 5</th>
<th>Grade 7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average Readers</td>
<td>Poor Readers</td>
<td>Average Readers</td>
</tr>
<tr>
<td></td>
<td>n = 59</td>
<td>n = 6</td>
<td>n = 50</td>
</tr>
<tr>
<td>FCAT 2.0 Reading</td>
<td>213.24 (14.73)</td>
<td>176.18 (5.12)</td>
<td>226.22 (11.77)</td>
</tr>
<tr>
<td>DSS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spelling Accuracy</td>
<td>6.95 (4.60)</td>
<td>3.50 (2.67)</td>
<td>10.04 (4.46)</td>
</tr>
<tr>
<td>Score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coding System</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morphological Legality</td>
<td>20.75 (15.46)</td>
<td>36.50 (19.58)</td>
<td>14.36 (13.84)</td>
</tr>
<tr>
<td>Phonological Legality</td>
<td>12.92 (6.50)</td>
<td>6.50 (6.98)</td>
<td>7.16 (12.25)</td>
</tr>
<tr>
<td>Orthographic Legality</td>
<td>29.27 (12.96)</td>
<td>26.17 (9.28)</td>
<td>20.24 (13.05)</td>
</tr>
<tr>
<td>Total Error Score</td>
<td>62.93 (34.46)</td>
<td>69.17 (33.74)</td>
<td>41.76 (29.62)</td>
</tr>
</tbody>
</table>

**Note.** Mean (with standard deviations in parentheses)

As for the Spelling Accuracy Score, homogeneity of variance was assessed by Levene’s Test for Equality of Variances with $p = .17$ for fourth graders and $p = .07$ for fifth graders. The results of the $t$ tests for these two grades showed that there were no significant differences between average and poor readers in Spelling Accuracy Scores, $p = .08$ in fourth grade and $p = .14$ in fifth grade. However, in seventh grade, even with a questionable result from Levene’s Test for Equality of Variances ($p = .01$), a significant difference between average and poor readers in Spelling Accuracy Scores was found, $M$ difference = 5.46, $t (48.768) = 7.759$, $p < .0001$. On the other hand, the result of a one way ANOVA showed that poor readers performed significantly better by grade, Welch’s $F (2, 11.81) = 10.34$, $p =.003$. Poor readers in the seventh grade outperformed those in the fourth grade according to the post-hoc Games-Howell test with $p <.0001$.

All of the Total Error Scores collected from the error coding system in the three grades met the assumption of homogeneity of variance, $p = .67$ in fourth grade, $p = .03$ in fifth grade,
and \( p = .94 \) in seventh grade. In addition, there were significant differences between average and poor readers in fifth and seventh grades (\( t (15.09) = 4.19, p < .0001; t (17.08) = 4.50, p < .0001 \), respectively). According to the result of a one-way ANOVA, poor readers performed significantly better by grade level, Welch’s \( F (2, 11.14) = 5.05, p = .02 \). The poor readers in seventh grade performed better than those in fifth grade according to the post-hoc Games-Howell test with \( p = .02 \). It should be noted that poor readers in the fifth grade made more Total Error Scores than those in the fourth grade.

Three profile analyses were performed on three dependent variables: Morphological Legality, Phonological Representation, and Orthographic Legality in three grades. The grouping variable was reading ability with two levels based on the percentiles of the FCAT 2.0 Reading DSS as mentioned above. One potential problem was that the sample sizes of each group were not equal and, therefore, the results of Pillai’s Trace Test were used. The first test in a profile analysis is for equal level. The results show significant differences in fifth and seventh grade levels, \( F (1, 61) = 17.55, p < .0001 \) and \( F (1, 67) = 22.37, p < .0001 \), respectively. However, no significant difference was found in the fourth grade, \( F (1, 63) = .18, p = .67 \).

For the flatness test, the results of Pillai’s Trace Test in three grades were \( F (2, 62) = 29.36 \) for fourth grade, \( F (2, 60) = 15.10 \) for fifth grade, and \( F (2, 66) = 18.33 \) for seventh grade and all with \( p < .0001 \). This indicated that the performance of the average and poor readers in the three grades was significantly different across the three categories in the error coding system. Figure 4, 5, and 6 graphed the profiles of both groups in the three grades. Visual inspection of the plotted profiles indicated that both groups in the three grades made more error points in Morphological Legality and Orthographic Legality than in Phonological Representation. The third test in a profile analysis is for parallelism. Only the fourth grade had a significant result, \( F \)
(2, 62) = 6.77, \( p = .002 \), as can be seen in Figure 4. It should be noted that the poor readers in fourth grade made more error points in Morphological Legality, but unexpectedly made fewer in Phonological Representation and Orthographic Legality than the average readers did. On the other hand, the fifth and seventh grades had parallel profiles between average and poor readers \( (F(2, 60) = 1.87, p = .17; F(2, 66) = .147, p = .86) \).

**Figure 4.** Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation, and OL for Orthographic Legality) for the Poor and Average Readers in the Fourth Grade

**Figure 5.** Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation, and OL for Orthographic Legality) for the Poor and Average Readers in the Fifth Grade
As can be seen in Figure 5 and 6, the profiles look similar between the two groups in fifth and seventh grade. The researcher was interested in whether the ratio of scores in each category to Total Error Score might provide more information for understanding the linguistic characteristics in the misspellings. Three new dependent variables were converted from the raw error score of each category divided by the Total Error Score (viz., Ratio of Morphological Legality to Total Error Score, Ratio of Phonological Representation to Total Error Score, and Ratio of Orthographic Legality). The plotted profiles, as can be seen in Figure 7, 8, and 9, deviated from flatness \( F (2, 62) = 30.75, p < .0001 \); \( F (2, 60) = 41.48, p < .0001 \); \( F (2, 66) = 39.87, p < .0001 \). Although the tests of interaction show significance only in the fifth grade \( F (2, 60) = 8.84, p < .0001 \), it is observable in the graphs of the fourth and seventh grades. The results from the fifth and seventh grades indicated that poor readers tend to make more error points in Phonological Representation then average readers proportionally, and poor readers tend to make fewer error points in Orthographic Legality then average readers, again, proportionally. However, the results from the fourth grade showed that poor readers had the same unexpected trends when using the ratios as dependent variables.
Figure 7. Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation, OL for Orthographic Legality, and TES for Total Error Score) for the Poor and Average Readers in the Fourth Grade

Figure 8. Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation, OL for Orthographic Legality, and TES for Total Error Score) for the Poor and Average Readers in the Fifth Grade
3. **How do the linguistic profiles of the older poor readers differ from those of younger normally progressing (reading-level matched) readers?**

To answer the third research question, the prior year’s FCAT 2.0 Reading DSS (Florida Department of Education, 2012) were used to match seventh graders with poor reading ability (i.e., at and below the 20th percentile) with fourth graders with average reading ability (i.e., at and above the 40th percentile). The FCAT 2.0 Reading DSS of the poor readers in seventh grade ranged from 167 to 206. Thus, normally progressing readers in fourth grade were then selected accordingly, with a range of FCAT 2.0 Reading DSS from 195 to 206. Table 5 demonstrates the means and the standard deviations of all the measures for older and younger readers.

Three independent sample t tests were conducted to compare the means of the FCAT 2.0 Reading DDS, the Spelling Accuracy Scores, and the Total Error Scores between older poor readers and younger normally progressing readers. As for the FCAT 2.0 Reading DDS, a t test showed that there was no significant difference between the two groups ($t(33) = .98$, $p = .33$) with no significant result from Levene’s Test for Equality of Variances ($p = .034$). This result
was expected since the two groups were matched on the FCAT 2.0 Reading DSS. As for the Spelling Accuracy Score, a t test showed that there was a significant difference between the two groups in Spelling Accuracy Scores, with $M$ difference $= 3.54$, $t (30.71) = 3.77$, $p = .001$ with significance in Levene’s Test for Equality of Variances ($p = .002$). The Total Error Scores collected from the error coding system also failed to meet the assumption of homogeneity of variance, $p = .001$. However, the result of a t test showed again that there was a significant difference between the two groups in Total Error Score, with $M$ difference $= 35.65$, $t (29.82) = 3.79$, $p = .001$.

Table 5  
*Means (Standard Deviations in parentheses) for All Scores of the Older Poor Readers and Younger Normally Progressing Readers*

<table>
<thead>
<tr>
<th>Coding System</th>
<th>Normally Progressing Readers in Fourth Grade</th>
<th>Poor Readers in Seventh Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n = 22</td>
<td>n = 13</td>
</tr>
<tr>
<td>FCAT 2.0 Reading DSS</td>
<td>199.95(3.88)</td>
<td>197.54(10.49)</td>
</tr>
<tr>
<td>Spelling Accuracy Score</td>
<td>5.45(3.86)</td>
<td>9.00(1.63)</td>
</tr>
<tr>
<td>Coding System</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morphological Legality</td>
<td>28.23(20.00)</td>
<td>15.15(5.87)</td>
</tr>
<tr>
<td>Phonological Representation</td>
<td>19.14(14.40)</td>
<td>11.08(14.02)</td>
</tr>
<tr>
<td>Orthographic Legality</td>
<td>34.59(13.461)</td>
<td>20.08(7.29)</td>
</tr>
<tr>
<td>Total Error Score</td>
<td>81.95(39.30)</td>
<td>46.31(15.41)</td>
</tr>
</tbody>
</table>

*Note.* Mean (with standard deviations in parentheses)

A profile analysis was performed on the three dependent variables: Morphological Legality, Phonological Representation, and Orthographic Legality. The tests of between-subject effects showed a significant difference in levels, $F (1, 33) = 9.711$, $p = .004$. For the flatness test, the results of the multivariate tests showed that the scores of the three categories from the error coding system deviated significantly from flatness, $F (2, 32) = 9.08$, $p = .001$. In Figure 10, both groups performed differently across the three categories. Visual inspection of the plotted profiles indicates that students in the two groups made more error points in Orthographic Legality and Morphological Legality than in Phonological Representation. The third test in a profile analysis
is for parallelism. No significant interaction effect was found $F (2, 32) = .719, p = .495$. Using the ratios as dependent variables, Figure 11 showed the two groups shared identical trends in their linguistic profiles.

*Figure 10.* Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation, and OL for Orthographic Legality) for the Two Reading-Ability-Matched Groups

*Figure 11.* Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation, OL for Orthographic Legality, and TES for Total Error Score) for the Two Reading-Ability-Matched Groups
4. **What is the incidence of unexpectedly poor spellers in the fourth, fifth, and seventh grades?**

How do their linguistic profiles differ from those of students who were matched on their reading or spelling ability?

To answer the fourth research question, the prior year’s FCAT 2.0 Reading DSS (Florida Department of Education, 2012) and the Spelling Accuracy Score were used to identify unexpectedly poor spellers and students who were poor at both spelling and reading in the three grades. Unexpectedly poor spellers were average readers with a Spelling Accuracy Score at or below the 20th percentile. Students with FCAT 2.0 Reading DSS and the Spelling Accuracy Score at or below the 20th percentile were identified as poor at both abilities. The percentile ranks of the Spelling Accuracy Score were constructed based on the data in each grade: a score of 2 represented the 20th percentile in fourth grade; a score of 4 represented the 20th percentile in fifth grade; and a score of 9 represented the 20th percentile in sixth grade. Table 6 demonstrates the group sizes, means and the standard deviations of all the measures for subgroups in the three grades.

Considering the small sample sizes in the three grades, the researcher decided to compare groups regardless of grades. Since the FCAT 2.0 Reading DSS and the Spelling Accuracy Score were used to identify different groups, it was expected to find differences among group means accordingly. The expectation was that the unexpectedly poor spellers would outperform the students with poor reading and spelling ability in the FCAT 2.0 Reading DSS. The expectation was confirmed, $t(42) = 6.16, p < .0001$. However, the results from the Spelling Accuracy Score ($t(25.48) = 1.26, p = .22$) and Total Error Score ($t(42) = .446, p = .67$) were both not significant, indicating the two groups performed similarly on both variables.
The researcher first investigated the difference between the linguistic profiles from unexpectedly poor readers and their peers who are poor at both reading and spelling. In other words, the two groups were matched on their poor spelling ability. A profile analysis was performed on the three dependent variables: Morphological Legality, Phonological Representation, and Orthographic Legality. The results of the test for equal level showed that there was no significant difference in levels, \( F (1, 42) = .20, p = .658 \). The profiles again showed a significant deviation from flatness, \( F (2, 41) = 15.52, p < .0001 \). Finally, there was no evidence of interaction, \( F (2, 41) = .09, p = .92 \) (see Figure 12). In addition, a profile analysis using the ratios as dependent variables (see Figure 13) demonstrated the almost identical trends between the two spelling-level-matched groups.

*Figure 12. Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation and OL for Orthographic Legality) for the Unexpectedly Poor Spellers and Students with Poor Reading and Spelling Abilities*
Figure 13. Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation, OL for Orthographic Legality, and TES for Total Error Score) for the Unexpectedly Poor Spellers and Students with Poor Reading and Spelling Abilities

The unexpectedly poor spellers seemed to perform similarly to students with poor reading and spelling abilities. The researcher then investigated whether the unexpectedly poor spellers performed differently from the average spellers at the same grade matched on the reading ability of the unexpectedly poor spellers ($n = 28$). In other words, the two groups were matched on their average reading ability. The results of t-tests showed that the two groups performed differently in the Spelling Accuracy Score and Total Error Score, all with $p < .0001$. As shown in Figure 14, the results of a profile analysis showed a significant difference in levels, $F (1, 54) = 28.67, p < .0001$. The profiles had a significant deviation from flatness, $F (2, 53) = 23.16, p < .0001$. Visual inspection of the plotted profiles showed that the two groups both made more error points in Orthographic Legality and Morphological Legality than in Phonological Representation. There was no significant interaction, $F (2, 53) = 2.59, p = .08$. The results of a profile analysis using the ratios from the three categories as dependent variables are shown in Figure 15. The
plotted profiles deviated from flatness ($F(2, 53) = 61.44, p < .0001$). Furthermore, there was also a significant interaction, $F(2, 53) = 10.00, p < .0001$.

Figure 14. Linguistic Profiles (ML for Morphological Legality, PR for Phonological Representation and OL for Orthographic Legality) for the Unexpectedly Poor Spellers and Average Spellers matched on the Reading Ability of the Unexpectedly Poor Spellers
Table 6
Means (Standard Deviations in parentheses) for All Scores of the Unexpectedly Poor Spellers and Poor in Both Students in the Three Grades

<table>
<thead>
<tr>
<th></th>
<th>Fourth Grade</th>
<th></th>
<th>Fifth Grade</th>
<th></th>
<th>Seventh Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unexpectedly Poor Spellers n = 12</td>
<td>Poor in Both n = 3</td>
<td>Unexpectedly Poor Spellers n = 8</td>
<td>Poor in Both n = 5</td>
<td>Unexpectedly Poor Spellers n = 8</td>
</tr>
<tr>
<td>FCAT 2.0</td>
<td>205.25</td>
<td>177.00</td>
<td>219.75</td>
<td>191.00</td>
<td>236.63</td>
</tr>
<tr>
<td></td>
<td>(8.70)</td>
<td>(4.58)</td>
<td>(8.00)</td>
<td>(7.55)</td>
<td>(14.66)</td>
</tr>
<tr>
<td>Reading DSS</td>
<td>1.08</td>
<td>1.33</td>
<td>3.50</td>
<td>1.80</td>
<td>7.25</td>
</tr>
<tr>
<td></td>
<td>(.793)</td>
<td>(1.16)</td>
<td>(.76)</td>
<td>(1.48)</td>
<td>(1.04)</td>
</tr>
<tr>
<td>Spelling Accuracy</td>
<td>38.33</td>
<td>47.33</td>
<td>36.00</td>
<td>46.60</td>
<td>16.63</td>
</tr>
<tr>
<td></td>
<td>(18.89)</td>
<td>(15.18)</td>
<td>(21.94)</td>
<td>(14.45)</td>
<td>(5.80)</td>
</tr>
<tr>
<td>Coding System</td>
<td>Phonomologal</td>
<td>Legality</td>
<td>24.83</td>
<td>9.00</td>
<td>20.88</td>
</tr>
<tr>
<td></td>
<td>39.33</td>
<td>25.67</td>
<td>29.25</td>
<td>41.00</td>
<td>22.25</td>
</tr>
<tr>
<td></td>
<td>(13.33)</td>
<td>(8.66)</td>
<td>(24.96)</td>
<td>(14.18)</td>
<td>(5.11)</td>
</tr>
<tr>
<td></td>
<td>(11.80)</td>
<td>(22.14)</td>
<td>(11.40)</td>
<td>(14.16)</td>
<td>(8.28)</td>
</tr>
<tr>
<td>Orthographic Legality</td>
<td>102.50</td>
<td>82.00</td>
<td>86.13</td>
<td>129.40</td>
<td>50.75</td>
</tr>
<tr>
<td></td>
<td>(31.87)</td>
<td>(44.24)</td>
<td>(44.28)</td>
<td>(28.95)</td>
<td>(13.02)</td>
</tr>
</tbody>
</table>

*Note: Mean (with standard deviations in parentheses)*
5. How do item characteristics (i.e., word frequency and morphological complexity) and subject characteristics (i.e., grade, reading ability, and word familiarity) affect spelling outcomes?

In the researcher-designed spelling test, students were asked to check whether they felt familiar with the target word before spelling the word. While doing data entry, the researcher noticed that some students checked the familiarity box without spelling the target word. The result of the Pearson correlation showed moderate evidence of an association between the overall Spelling Accuracy Score and Word Familiarity, $r (219) = .44, p < .001$. However, the result from the poor readers was not significant, $r (32) = .33, p = .07$. A series of Pearson correlations were conducted between the Spelling Accuracy Score and Word Familiarity of each target word in order to investigate the relations more specifically. Most of the $r$ values were below .20 and were not significant. Some target words even had negative correlations between the Spelling Accuracy Score and Word Familiarity, such as preceding ($r (219) = -.04, p = .54$). Therefore, Word Familiarity was not informative enough to be used as a subject-related variable.
When constructing the spelling test, 25 target words were evenly selected from 10 sublists of the Academic Word List (Coxhead, 2000) to ensure diversity of word frequency in the test. Among the 25 target words, Word Frequency was assigned to each target word based on which sublist it was selected from, ranging from 1 representing the most frequent word to 10 representing the least frequent word. Because words with low frequency tend to be more difficult to spell, a negative correlation was expected between the Spelling Accuracy Score and Word Frequency. However, the result of the Pearson correlation showed that there was no significant relation between the Spelling Accuracy Score and Word Frequency, \( r(24) = -.38, p = .06 \). Among the three grades, only the data from the seventh grade showed a significant negative correlation between the two variables, \( r(24) = -.44, p = .03 \).

As for Word Type, the 25 target words were deliberately chosen and manipulated based on their morphological complexity. A base word without any affixes or words requiring no spelling change with the addition of an affix was categorized as having low morphological complexity. Conversely, words with more than 2 affixes or words requiring a spelling change when an affix is added were considered to have high morphological complexity. Among the 25 target words, 10 items had low morphological complexity and the rest (15) had high morphological complexity. The researcher divided the 25 words into two parts and calculated the Spelling Accuracy Score separately. The average score from items (\( n = 10 \)) with low morphological complexity was named Spelling Accuracy Score-Low and the average score from the rest of the items (\( n = 15 \)) was named Spelling Accuracy Score-High.

From the results of the previous research questions, it was evident that students from different subgroups made significantly different mean number of error points in Morphological Legality. The researcher further investigated whether Word Type, which was defined based on
the morphological complexity of each item, would have an impact on the Spelling Accuracy Scores of different subgroups. A Mixed-effects model (Quene & van den Bergh, 2008) was measured including the Spelling Accuracy Score of each item as the outcome variable. However, the reduction of the deviance by adding Word Type as a predictor at the item level was only 3.7 with one degree of freedom difference. The result from the Hierarchical Linear Modeling (Bryk & Raudenbush, 1992) showed that Word Type did not seem to be able to explain a significant amount of variance in the spelling performance.

Table 7
The Results of Independent Samples t-tests between Different Subgroups

<table>
<thead>
<tr>
<th>Comparisons between groups</th>
<th>Spelling Accuracy Score-High</th>
<th>95% Confidence Interval</th>
<th>Spelling Accuracy Score-Low</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Diff</td>
<td>Std. Error Diff</td>
<td>Lower Bound</td>
<td>Upper Bound</td>
</tr>
<tr>
<td>Average and Poor Readers</td>
<td>.17***</td>
<td>.03</td>
<td>.23</td>
<td>.11</td>
</tr>
<tr>
<td>Average and Poor Spellers</td>
<td>.29***</td>
<td>.02</td>
<td>.34</td>
<td>.25</td>
</tr>
<tr>
<td>Older Poor Readers and Younger Normally Progressing Readers</td>
<td>.12*</td>
<td>.05</td>
<td>.21</td>
<td>.02</td>
</tr>
<tr>
<td>Unexpectedly Poor Spellers and Average Spellers matched on Reading Ability</td>
<td>.27***</td>
<td>.04</td>
<td>.36</td>
<td>.18</td>
</tr>
<tr>
<td>Unexpectedly Poor Spellers and Students with Poor Reading and Spelling Abilities</td>
<td>.02</td>
<td>.04</td>
<td>.10</td>
<td>.06</td>
</tr>
</tbody>
</table>

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

The researcher was interested in whether Word Type affected spelling performance differently among subgroups. A series of t-tests were performed to examine the differences in
Spelling Accuracy Score-Low and Spelling Accuracy Score-High among groups. Table 7 showed that all the comparisons among groups were statistically significant including those of the unexpected poor spellers and students with poor reading and spelling abilities, $t (42) = .49, p = .62$ for Spelling Accuracy Score-Low and $t (42) = .54, p = .59$ for Spelling Accuracy Score-High. The result showed that students with average reading or spelling abilities outperformed those with poor reading or spelling abilities on both target words with high and low morphological complexity. In addition, older poor readers also had higher accuracy than younger normally progressing readers.
CHAPTER FOUR

DISCUSSION

The purpose of this study was to explore the misspellings made by students in intermediate and higher grades on a 25-word group-administered, spelling test. A sample of 220 students with parent consent and child assent was recruited from two schools in northern Florida and consisted of a representative range of demographics characteristics. Linguistic profiles were synthesized from the three categories in the researcher-designed spelling error coding system. Inspired by the Triple Word Form (e.g., Berninger et al., 2001, 2008), each misspelling was sequentially coded based on its linguistic features across the three categories: Morphological Legality, Phonological Representation, and Orthographic Legality. The cumulative scores from each category created a linguistic profile for each misspelling. Compiling the information from each misspelling, the researcher was able to compare the differences in performance of students at different grades, reading, and spelling ability levels. Profile analysis was the major statistical technique to examine the differences across group profiles by testing the equal levels, flatness, and parallelism hypotheses (Morrison, 1976).

1. How do fourth, fifth and seventh grade students perform on the prior year’s Florida Comprehensive Assessment Test (FCAT) 2.0 Reading Developmental Scale Scores (DSS, Florida Department of Education, 2012) and on the researcher-designed spelling test? What are the linguistic profiles of the spelling errors made by these students?

The purpose of the first question was to understand the general distribution of reading and spelling performance among the three grades. In addition, the information provided a foundation for the researcher to further investigate the differences among subgroups. As expected, students in higher grades did perform significantly better than those in lower grades in
both reading and spelling. The normal distribution and the expected increase by grade levels in the prior year’s FCAT 2.0 Reading DSS indicated that the data were representative. On the other hand, it seemed that the 25-item spelling test was very difficult for the sample. Students demonstrated low performance accuracy rates (viz., Spelling Accuracy Score) at each grade (i.e., 25% for the fourth grade, 38% for the fifth grade, and 53% for the seventh grade). Consequently, the floor effect in the fourth grade was notable. In this study, the spelling test served as a tool to collect misspellings from the three grades so that more information could be obtained through the error coding system. Hence, it was necessary to maintain the difficulty level across the three grades. The difficulty issue might be a reason for caution when using the Spelling Accuracy Score as a grouping variable in the fourth research question, which will be discussed in greater depth later. However, it raised an even more severe concern with respect to spelling instruction.

The floor effect in the fourth grade indicated that the spelling test developed for this study was very challenging. The target words in the researcher-designed spelling test were selected from the Academic Word List (Coxhead, 2000), which did not include the first 2,000 words of English as given in the General Service List (West, 1953). However, the Common Core State Standards in English Language Arts clearly state that students above fourth grade are expected to accurately spell using the conventions of the English written language (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). In other words, fourth-grade students are expected to be fair spellers even when encountering novel words. The results of this study suggest that current levels of spelling instruction may not be adequate to achieve this educational goal. Spelling is a process that needs to and can be taught, just like reading (Graham & Perrin, 2006; Graham & Perrin, 2007). According to the findings of Foorman and Petscher (2010), some teachers are teaching spelling
in a more efficient way than others. It is possible to improve spelling instruction in order to help students become more competent spellers.

The Total Error Score was the sum of the cumulative scores from the three coding categories. As expected, the Total Error Scores decreased by grade level, indicating that students in higher grades made fewer error points in their misspellings. The wide range of score distribution suggested that students performed differently from each other and provided a foundation for the researcher to examine the differences among the linguistic profiles of the subgroups. By using Morphological Legality (ML), Phonological Representation (PR), and Orthographic Legality (OL) as the dependent variables, a profile analysis was performed to compare the differences among the linguistic profiles from the three grades. The results of the profile analysis showed that the linguistic profiles were different across the grades and the categories. In addition, a significant grade by category interaction was also detected.

For the older students above fourth grade, the linguistic profiles showed a significant increase in the error points from the three categories by grade level. This result confirmed the hypothesis that older students are more skilled in coordinating their linguistic to achieve better spelling accuracy. The similar trends shared among the three grades indicated no evidence supporting the stage-specific theory. Stage theorists identified the development across stages by the observation of the consistent use of one specific strategy (e.g., Ellis, 1994; Gentry, 1982). From the result of this study, we can conclude that students in fourth, fifth, and seventh grades were getting more and more skilled in spelling while they applied all three of the components of spelling knowledge simultaneously.

The discrepancies among the three categories in the linguistic profile of the fourth grade students were larger than those in the profiles of the fifth and seventh grade students, which
caused the significant grade by category interaction. The ratios of the three coding categories to the Total Error Score provided another perspective to interpret the differences among the linguistic profiles from the three grades. The plotted profiles showed almost identical results, indicating that students from the three grades made proportionally similar error points across the three categories. The lowest ratios were located in Phonological Representation and the highest ratios were located in Orthographic Legality, regardless of grades. Given the sequential approach used in the error coding system, each misspelt letter or letter set was penalized only for Phonological Representation or Orthographic Legality in a weighted manner. Theoretically, the more sounds that are represented with an accurate number of letters in a misspelling, the more likely the representing letters will be scored as legal orthographically. These results further confirmed that the patterns of usage among the three categories might be following the same developmental path across the three grades in this study. The findings also support the Overlapping Waves Model (e.g., Hoijtink & Notenboom, 2004), which recognizes the complexity of the use of spelling knowledge throughout development. Similarly, Berninger, Abbott, Nagy and Carlisle (2010) found that all three kinds of linguistic knowledge were necessary and ought to be coordinated to facilitate accurate spelling.

2. How do the poor readers perform on the spelling test? How do the poor readers’ linguistic profiles of misspellings differ from their (grade-matched) peers with average reading ability?

To answer the second research question, the prior year’s FCAT 2.0 Reading DSS (Florida Department of Education, 2012) were used to identify average and poor readers in the three grades. The results of the descriptive analyses confirmed that the poor readers performed worse than their peers in spelling. However, the results of the t-test showed that the Spelling Accuracy Scores were significantly different between average and poor readers only in the seventh grade.
The results might be due to the high level of difficulty of the spelling test, which may have diminished the power of discrimination. The Total Error Score results revealed that the scores that were added-up from the error coding system differentiated performance in both the fifth and seventh grades. These results demonstrated the limitation of the traditional scoring method based on accuracy only. It was expected that the linguistic profiles might reveal more information about the use of Morphological, Phonological, and Orthographic knowledge in misspellings.

Another noteworthy finding was that the Spelling Accuracy Score and Total Error Score yielded different results for poor readers across the three grades. While the Spelling Accuracy Score differentiated the poor readers in the seventh grade from those in the fourth grade, the Total Error Score revealed that poor readers in the fifth grade made even more error points than those in the fourth grade. This conflicting result might be due to the normality issue in the data from the fourth grade. In fact, the researcher found an extremely low correlation between the FCAT 2.0 Reading DSS and the Spelling Accuracy Score for the poor readers in the fourth grade (r = .17). As mentioned earlier, the spelling test was very difficult for the fourth grade students and provided limited variability in spelling levels. In other words, students in the fourth grade generally performed poorly in the spelling test regardless of their reading ability. Some students with average reading ability performed even worse than their peers with low reading ability.

For the fifth and seventh grades, the results of the profile analyses demonstrated significant main effects both between and within subjects (as was shown in Figure 5 and 6). Although the poor readers produced significantly more error points across the three categories, they did share the same trend with the average readers. The profile analysis using the ratios as dependent variables showed some distinct competitions between Phonological Representation and Orthographic Legality. The competitions supported the findings of the Overlapping Waves
Model (Hoijtink & Notenboom, 2004) that a variety of strategies is used by spellers and each strategy demonstrates gradual change at different rates in spelling development. While both average and poor readers suffered the most from a deficit in how to use orthographic knowledge in spelling, average readers utilized their sufficient phonological knowledge significantly better than other linguistic knowledge. This distinct discrepancy found in the average readers in the fifth and seventh grades supported the hypothesis that good readers not only benefit from phonological knowledge in reading but also in spelling.

The benefits of sufficient phonological knowledge in both reading and spelling did not positively confirm the specific phonological deficit hypothesis (e.g., Bruck, 1988; Friend & Olson, 2008). On the contrary, it suggested that students in the fifth or seventh grade need to be a good reader first to be able to benefit from sufficient phonological knowledge in their spelling. On the other hand, there was no evidence to support that poor readers would rely more on orthographic and morphological knowledge and make a higher proportion of error points in Phonological Representation because of their specific phonological deficit. In other words, poor readers not only suffered from limited phonological knowledge but also limited orthographic and morphological knowledge in spelling.

The profile analysis of the data from the fourth grade demonstrated a different pattern from those of the fifth and seventh grades. As mentioned previously, the spelling test was too difficult for the fourth grade students in this study. The unexpected low correlation between reading and spelling ability indicated that reading ability might not be an effective factor to predict spelling performance. Therefore, it would be meaningless to answer whether students with different reading abilities might perform differently in their spelling. We might conclude that the effects of reading ability on the linguistic profiles of misspellings occurred significantly
in the higher grades, such as fifth or seventh grades in this study. For the students in the fourth grade, reading ability was not a sufficient factor to differentiate the use of linguistic knowledge when encountering novel words. In other words, the students in the fourth grade did not seem to benefit from their reading ability when spelling difficult words. This result reminds us of the complex relations between reading and spelling.

3. How do the linguistic profiles of the older poor readers differ from those younger normally progressing (reading-level matched) readers?

To answer the third research question, the prior year’s FCAT 2.0 Reading DSS were used to match poor readers in the seventh grade with some average readers in the fourth grade who shared the same reading ability with the older poor readers, viz., younger normally progressing readers. Although they shared the same reading level, the older poor readers outperformed the younger normally progressing readers in both the Spelling Accuracy Score and Total Error Score. This result suggests that the spelling of older poor readers indeed benefited from additional reading exposure. However, as mentioned previously, it is necessary to note that most of the fourth graders had difficulty spelling the target words even with average reading ability. Thus, there might be other factors causing the differences between older poor readers and younger normally progressing readers.

Result of the profile analysis showed that, controlling for reading level, younger normally progressing readers made more error points across the three categories but produced similar trends compared to the profile of older poor readers. Additionally, the two groups shared almost identical profiles using the ratios as dependent variables. The results did not provide evidence to support the specific phonological deficit hypothesis because the older poor readers did not make relatively more errors in the Phonological Representation category. The older poor readers proportionally utilized their linguistic knowledge while spelling in a very similar way to the
younger normally progressing readers. Even though the older poor readers outperformed the younger normally progressing readers in spelling, it was other factors that contributed to this difference in outcome, such as item difficulty, reading exposure or word familiarity.

4. What is the incidence of unexpectedly poor spellers in the fourth, fifth, and seventh grades? How do their linguistic profiles differ from those of students who were matched on their reading or spelling ability?

To answer the fourth research question, the prior year’s FCAT 2.0 Reading DSS and Spelling Accuracy Score were used to identify unexpectedly poor spellers and students who were poor at both spelling and reading in the three grades. The unexpectedly poor spellers were first matched with the poor readers on their poor spelling ability and then matched with the average spellers on their average reading ability. The results of comparing the means revealed that the unexpectedly poor spellers and those students with poor reading and spelling ability performed similarly on the Spelling Accuracy Score and the Total Error Score. The plotted profiles from the two profile analyses also showed no evidence of difference between the two groups.

On the other hand, the results from the comparison between the unexpectedly poor spellers and the average spellers who were matched with their average reading ability showed some significant differences between the groups. As expected, their Spelling Accuracy Score and Total Error Score were significantly different. The results of the profile analyses demonstrated a similar pattern as was seen in the previous comparisons. Using the raw scores from the three categories as the dependent variables, the plotted profiles from the two groups showed similar trends across the categories but a significant difference at the levels. Using the ratios as the dependent variables, the average spellers made proportionally fewer error points in Phonological Representation, indicating that the average spellers benefited from their sufficient phonological
knowledge in spelling. Even sharing the same reading level, the unexpectedly poor spellers obviously did not efficiently utilize phonological knowledge in spelling. In addition, they made proportionally more error points in Orthographic and Morphological Legality. This result indicated that the unexpectedly poor spellers did not just suffer from insufficient phonological knowledge, but also from limited orthographic and morphological knowledge in spelling.

5. How do item characteristics (i.e., word frequency and word type) and subject characteristics (i.e., grade, reading ability, and word familiarity) affect spelling outcomes?

To answer the fifth research question, the researcher started with some preliminary analyses. The results confirmed that some of the variables did not provide enough information for further investigation. At the subject level, the differences in the spelling outcomes caused by grade and reading abilities were well documented in the previous research question. Word Familiarity was expected to explain the relations between self-perception and spelling outcomes. However, during the administration of the spelling test, the researcher noticed that some students in this study checked Word Familiarity even before they heard the target word, or they left empty responses to the target words that they had claimed they knew. The resulting low correlations confirmed the observation. Asking students to evaluate their familiarity with the target words right before spelling them might not be a feasible method to obtain valid information on self-perception of spelling. At the item level, Word Frequency was expected to have a high correlation with the Spelling Accuracy Score. As mentioned previously, the spelling test was very difficult for many students in this study. Considering the normality issue in the fourth grade, it was predictable that Word Frequency might not be informative for further analysis. It was confirmed that Word Frequency only had significant correlation with the Spelling Accuracy Score in the seventh grade.
On the other hand, Word Type seemed to promise some information that could shed light on spelling performance. Although the result of the Mixed-effects model (Quene & van den Bergh, 2008) showed that Word Type did not explain much of the variance in the Spelling Accuracy Score, the researcher was interested in whether Word Type might affect different subgroups differently. As expected, most of the comparisons showed that there were significant differences among groups, regardless of whether the comparisons of the average accuracy scores were of items with low or high morphological complexity. This result confirmed the findings of the profile analyses discussed previously. In addition, no significant differences were found between the unexpectedly poor spellers and their peers who were poor at both reading and spelling. The result reaffirmed the similarity between the profiles of the two groups. Not only had the two groups shared an identical pattern of use of phonological and orthographic knowledge, but also of morphological knowledge in their spelling. It was worthy to note that among the comparisons, none of them showed inconsistent results between Spelling Accuracy Score-High and Spelling Accuracy Score-Low. In other words, the two sets of target words in the researcher-designed spelling test both caused average mean differences among groups regardless of their morphological complexity.

Summary

The purpose of the study was to explore the differences in spelling performance among the students in the fourth, fifth, and seventh grades. The main assumption in this study was supported by the Overlapping Waves Model (e.g., Hoijtink & Notenboom, 2004), which advocates that spellers use all of their related linguistic knowledge in spelling simultaneously regardless of their age, reading, or spelling abilities. The researcher designed a 25-word spelling test and a spelling error coding system in order to efficiently decipher the use of linguistic
knowledge in the misspellings. The error coding system consisted of three categories, namely, Morphological Legality (ML), Phonological Representation (PR), and Orthographic Legality (OL). The three categories were established based on the Triple Word Form theory (Berninger, 2004; Berninger, Abbott, Billingsley, & Nagy, 2001; Berninger & Richards, 2002). It was evident that the error coding system provided more information than the traditional accuracy scoring method (i.e., Spelling Accuracy Score in this study).

This study did not find evidence supporting the stage-specific theory in spelling development. The result of the profile analysis of the three grades showed a consistent development of the use of linguistic knowledge and a stable trend of coordination among the three error coding categories. Although students became more and more skillful at using linguistic knowledge in general as they progressed in grade level, they consistently made the most error points in Orthographic Legality. In the error coding system, the affix and juncture of a misspelling was first coded in Morphological Legality, and the base word (or root) of that misspelling was then coded in Phonological Representation and Orthographic Legality. Any spelling error in a base word was penalized only once in either of the two categories. The sequential coding approach facilitated a valid competition between Phonological Representation and Orthographic Legality, which allowed us to further interpret the differences in the use of linguistic knowledge in misspellings among subgroups.

It was such competition between Phonological Representation and Orthographic Legality that raised uncertainty with respect to the specific phonological deficit hypothesis. The comparisons between the average and poor readers in the fifth and seventh grades revealed that average readers benefited from their phonological knowledge in spelling and the poor readers did not specifically suffer from their limited phonological knowledge in spelling. Controlling for
reading ability, the older poor readers used their linguistic knowledge better than the younger normally progressing readers in spelling. However, their poor reading ability hindered them from applying their phonological knowledge as sufficiently in spelling as their peers.

On the other hand, it was also confirmed that having average reading ability does not guarantee average spelling ability. The unexpectedly poor spellers failed to sufficiently apply their phonological knowledge in spelling. They demonstrated a similar linguistic profile with those students who were poor at both reading and spelling. Another comparison of the unexpectedly poor spellers and their peers who were matched on reading ability proved that average readers need to be able to sufficiently apply their phonological knowledge in spelling in order to be fair spellers.

In short, the results in this study did not support the specific phonological deficit hypothesis. On the contrary, the researcher found that the key to becoming a fair speller is to be able to effectively apply sufficient phonological knowledge in spelling. For students with poor reading ability, they do not just suffer from limited phonological knowledge but also from lack of other linguistic knowledge. For any two students with average reading ability, it is the one who can apply sufficient phonological knowledge that will benefit in spelling and perform at the level that matches his or her reading ability.

**Limitations and Recommendations for Future Research**

It is important to note that there were several limitations in this study. Due to the unfortunate reality of subject recruitment in this research, the sample size of each subgroup was limited. There were only 6, 13, and 13 poor readers, and 12, 8, and 8 unexpectedly poor spellers in each grade, respectively. This might cause a reduction in the power of the analyses. In the current study, the researcher followed the recommended procedure of using the results of the
Pillai’s Trace Test for the profile analyses. To facilitate research using a valid match design, power analyses should be implemented in order to obtain enough samples for each comparison.

In addition, the study did not provide more qualitative information on how each subject responded to the 25 target words. As mentioned previously, Fowler, Shankweiler, and Liberman (1979) found that spellers without sufficient orthographic knowledge tended to make systematic mistakes such as overgeneralizing rules or neglecting within-word environments. While doing data entry, the researcher found some students did make systematic mistakes throughout the entire spelling test (i.e., replacing *tion* with *sion*). The variation within subject was not included in this study, but certainly could be an informative direction for future research.

Finally, the scoring method of Morphological Legality in the error coding system was not very functional when coding the responses from the spelling test. Morphological Legality in the error coding system applied a weighting approach to score each misspelled affix and juncture. There were many omissions in target words like *automatically*, which was scored as 3 points. The students in this study typically dropped *al* at the end. There were other errors made regarding the orthographic rules but not regarding the morphological rules, such as replacing *hypo* in *hypothesis* with *highpo*, which was considered as an illegal error and scored 2 points. However, there were not a lot of misspellings given 1 point in the ML category. Among the 25 target words in the spelling test, only the target word, *coincidence*, provided a potential opportunity to make a legal error in Morphological Legality by replacing *ence* with *ance*. For future research, it would be more enlightening to design a spelling test with more potential opportunities to make legal errors.
Educational Implications

The main purpose of this research was to utilize the error coding system in order to disclose different linguistic profiles made by the students at different grades with different levels of reading and spelling abilities. However, the use of the error coding system should not be limited for research purposes. The error coding system could be a very powerful and informative tool in spelling instruction. Teachers can apply the error coding system to evaluate spelling production in class and can individualize spelling instruction accordingly. In the current study, each misspelling was divided into the affix and base word to be scored separately in the Morphological Legality, Phonological Representation, and Orthographic Legality categories. By keeping the affix and juncture scored under Morphological Legality only, it was not required to determine whether or not the affix or juncture was spelled following orthographic rules. Information on the use of phonological and orthographic knowledge in a misspelling could be obtained under Phonological Representation and Orthographic Legality. Nonetheless, classroom teachers do not need to follow the division procedure of putting affix and juncture under Morphological Legality only and not also under Phonological Representation and Orthographic Legality.

As mentioned previously, students apply all three types of linguistic knowledge to spell accurately. They might not consciously facilitate their specific linguistic knowledge separately when spelling an affix or a base word. Thus, it is reasonable to evaluate how close a misspelling approaches the accurate linguistic structure by using all three categories simultaneously. The finding of this study emphasized the importance of sufficient phonological knowledge on spelling development. Teachers can start doing spelling error analysis by counting if each sound in a target word is represented by a letter(s). When coding a sounding affix, the omission of
using letter(s) to represent a sound (i.e., omitting *al* in *automatically*) should be considered as a severe deficit in applying phonological knowledge in spelling. Next, it should be determined whether each letter(s) used to represent a sound follows orthographic rules or not. When encountering a pronunciation shift due to the morphophonemic nature of English, it is important to keep in mind that some students may have misapplied their orthographic knowledge and failed to maintain accurate spelling. For example, a student might spell the word *sign* as *sine*, which is orthographically legal, but fail to maintain the morphological relation with the word *signal*.

By analyzing the misspellings, teachers can better understand how closely an error approaches the phonological, orthographic, and morphological structure of the target word. Here is another set of examples. Students make different kinds of mistakes when forming plural nouns. For students who spell *boxes* as *boxs*, they fail to represent all the sounds in the target word making the misspelling a phonological error. For students who spelled *babies* as *babys*, they do not have solid orthographic knowledge to produce a correct spelling following the rule for dropping *y*. Students make morphological mistakes when spelling words like *knife*. The plural form of *knife* is *knives* when used as a noun. However, *knife* can also be used as a verb which means to cut with a knife. When students spell *knives* as *knifes*, they are turning *knife* from a noun to a verb in the simple present tense using the subject he, she, or it. Once teachers have learned how to follow the coding steps, they can apply the technique on any types of spelling outcomes, such as classroom spelling tests, writing products, or even in norm-referenced spelling tests. Furthermore, teachers who are familiar with the coding steps will also be aware of the need to include the three major components (i.e., phonological, orthographic, and morphological knowledge) in their reading and spelling teaching.
APPENDIX A

PERCENTILE RANKS OF FCAT 2.0 READING DSS

Table 8 The percentile ranks constructed from FCAT 2.0 Reading DSS (Florida Department of Education, 2012)

<table>
<thead>
<tr>
<th>Percentile Ranks</th>
<th>Grade 3</th>
<th>Grade 4</th>
<th>Grade 6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DSS Low</td>
<td>DSS High</td>
<td>DSS Low</td>
</tr>
<tr>
<td>1</td>
<td>153</td>
<td>168</td>
<td>178</td>
</tr>
<tr>
<td>2</td>
<td>154</td>
<td>169</td>
<td>173</td>
</tr>
<tr>
<td>3</td>
<td>158</td>
<td>174</td>
<td>176</td>
</tr>
<tr>
<td>4</td>
<td>161</td>
<td>177</td>
<td>178</td>
</tr>
<tr>
<td>5</td>
<td>164</td>
<td>179</td>
<td>180</td>
</tr>
<tr>
<td>6</td>
<td>166</td>
<td>181</td>
<td>182</td>
</tr>
<tr>
<td>7</td>
<td>168</td>
<td>183</td>
<td>183</td>
</tr>
<tr>
<td>8</td>
<td>170</td>
<td>184</td>
<td>185</td>
</tr>
<tr>
<td>9</td>
<td>171</td>
<td>186</td>
<td>186</td>
</tr>
<tr>
<td>10</td>
<td>172</td>
<td>187</td>
<td>187</td>
</tr>
<tr>
<td>11</td>
<td>174</td>
<td>188</td>
<td>188</td>
</tr>
<tr>
<td>12</td>
<td>175</td>
<td>189</td>
<td>189</td>
</tr>
<tr>
<td>13</td>
<td>176</td>
<td>190</td>
<td>190</td>
</tr>
<tr>
<td>14</td>
<td>177</td>
<td>191</td>
<td>191</td>
</tr>
<tr>
<td>15</td>
<td>178</td>
<td>192</td>
<td>192</td>
</tr>
<tr>
<td>16</td>
<td>179</td>
<td>193</td>
<td>193</td>
</tr>
<tr>
<td>17</td>
<td>180</td>
<td>194</td>
<td>194</td>
</tr>
<tr>
<td>19</td>
<td>181</td>
<td>195</td>
<td>195</td>
</tr>
<tr>
<td>20</td>
<td>182</td>
<td>196</td>
<td>196</td>
</tr>
<tr>
<td>21</td>
<td>183</td>
<td>197</td>
<td>197</td>
</tr>
<tr>
<td>22</td>
<td>184</td>
<td>198</td>
<td>198</td>
</tr>
<tr>
<td>24</td>
<td>185</td>
<td>199</td>
<td>199</td>
</tr>
<tr>
<td>25</td>
<td>186</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>27</td>
<td>187</td>
<td>201</td>
<td>201</td>
</tr>
<tr>
<td>28</td>
<td>188</td>
<td>202</td>
<td>202</td>
</tr>
<tr>
<td>30</td>
<td>189</td>
<td>203</td>
<td>203</td>
</tr>
<tr>
<td>31</td>
<td>190</td>
<td>204</td>
<td>204</td>
</tr>
<tr>
<td>33</td>
<td>191</td>
<td>205</td>
<td>205</td>
</tr>
<tr>
<td>35</td>
<td>192</td>
<td>206</td>
<td>206</td>
</tr>
<tr>
<td>36</td>
<td>193</td>
<td>207</td>
<td>207</td>
</tr>
<tr>
<td>38</td>
<td>194</td>
<td>208</td>
<td>208</td>
</tr>
<tr>
<td>40</td>
<td>195</td>
<td>209</td>
<td>209</td>
</tr>
<tr>
<td>42</td>
<td>196</td>
<td>210</td>
<td>210</td>
</tr>
<tr>
<td>Percentile Ranks</td>
<td>Grade 3 DSS Low</td>
<td>Grade 3 DSS High</td>
<td>Grade 4 DSS Low</td>
</tr>
<tr>
<td>------------------</td>
<td>----------------</td>
<td>-----------------</td>
<td>----------------</td>
</tr>
<tr>
<td>43</td>
<td>197</td>
<td>197</td>
<td>211</td>
</tr>
<tr>
<td>45</td>
<td>198</td>
<td>198</td>
<td>212</td>
</tr>
<tr>
<td>47</td>
<td>199</td>
<td>199</td>
<td>213</td>
</tr>
<tr>
<td>49</td>
<td>200</td>
<td>200</td>
<td>214</td>
</tr>
<tr>
<td>51</td>
<td>201</td>
<td>201</td>
<td>215</td>
</tr>
<tr>
<td>53</td>
<td>202</td>
<td>202</td>
<td>216</td>
</tr>
<tr>
<td>54</td>
<td>203</td>
<td>203</td>
<td>217</td>
</tr>
<tr>
<td>56</td>
<td>204</td>
<td>204</td>
<td>218</td>
</tr>
<tr>
<td>58</td>
<td>205</td>
<td>205</td>
<td>219</td>
</tr>
<tr>
<td>60</td>
<td>206</td>
<td>206</td>
<td>220</td>
</tr>
<tr>
<td>62</td>
<td>207</td>
<td>207</td>
<td>221</td>
</tr>
<tr>
<td>63</td>
<td>208</td>
<td>208</td>
<td>222</td>
</tr>
<tr>
<td>65</td>
<td>209</td>
<td>209</td>
<td>223</td>
</tr>
<tr>
<td>67</td>
<td>210</td>
<td>210</td>
<td>224</td>
</tr>
<tr>
<td>68</td>
<td>211</td>
<td>211</td>
<td>225</td>
</tr>
<tr>
<td>70</td>
<td>212</td>
<td>212</td>
<td>226</td>
</tr>
<tr>
<td>71</td>
<td>213</td>
<td>213</td>
<td>227</td>
</tr>
<tr>
<td>73</td>
<td>214</td>
<td>214</td>
<td>228</td>
</tr>
<tr>
<td>74</td>
<td>215</td>
<td>215</td>
<td>229</td>
</tr>
<tr>
<td>76</td>
<td>216</td>
<td>216</td>
<td>230</td>
</tr>
<tr>
<td>77</td>
<td>217</td>
<td>217</td>
<td>231</td>
</tr>
<tr>
<td>79</td>
<td>218</td>
<td>218</td>
<td>232</td>
</tr>
<tr>
<td>80</td>
<td>219</td>
<td>219</td>
<td>233</td>
</tr>
<tr>
<td>81</td>
<td>220</td>
<td>220</td>
<td>234</td>
</tr>
<tr>
<td>82</td>
<td>221</td>
<td>221</td>
<td>235</td>
</tr>
<tr>
<td>84</td>
<td>222</td>
<td>222</td>
<td>236</td>
</tr>
<tr>
<td>85</td>
<td>223</td>
<td>223</td>
<td>237</td>
</tr>
<tr>
<td>86</td>
<td>224</td>
<td>224</td>
<td>238</td>
</tr>
<tr>
<td>87</td>
<td>225</td>
<td>225</td>
<td>239</td>
</tr>
<tr>
<td>88</td>
<td>226</td>
<td>226</td>
<td>240</td>
</tr>
<tr>
<td>89</td>
<td>227</td>
<td>228</td>
<td>241</td>
</tr>
<tr>
<td>90</td>
<td>229</td>
<td>229</td>
<td>243</td>
</tr>
<tr>
<td>91</td>
<td>230</td>
<td>230</td>
<td>244</td>
</tr>
<tr>
<td>92</td>
<td>231</td>
<td>232</td>
<td>246</td>
</tr>
<tr>
<td>93</td>
<td>233</td>
<td>233</td>
<td>248</td>
</tr>
<tr>
<td>94</td>
<td>234</td>
<td>235</td>
<td>250</td>
</tr>
</tbody>
</table>
Table 8 – continued

<table>
<thead>
<tr>
<th>Percentile Ranks</th>
<th>Grade 3</th>
<th>Grade 4</th>
<th>Grade 6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DSS Low</td>
<td>DSS High</td>
<td>DSS Low</td>
</tr>
<tr>
<td>95</td>
<td>236</td>
<td>237</td>
<td>253</td>
</tr>
<tr>
<td>96</td>
<td>238</td>
<td>240</td>
<td>258</td>
</tr>
<tr>
<td>97</td>
<td>241</td>
<td>243</td>
<td></td>
</tr>
<tr>
<td>98</td>
<td>244</td>
<td>248</td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>249</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## APPENDIX B

### THE RESEARCHER-DESIGNED SPELLING TEST

<table>
<thead>
<tr>
<th>Target Words</th>
<th>Word Frequency</th>
<th>Word Type</th>
<th>Sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Period</td>
<td>1</td>
<td>Low</td>
<td>The traffic light changed after a long period of time.</td>
</tr>
<tr>
<td>Definition</td>
<td>1</td>
<td>High</td>
<td>The students did not understand the definition of the word.</td>
</tr>
<tr>
<td>Restricted</td>
<td>2</td>
<td>High</td>
<td>Entrance to the park is restricted after dark.</td>
</tr>
<tr>
<td>Commission</td>
<td>2</td>
<td>High</td>
<td>My commission for selling the tickets was not very high.</td>
</tr>
<tr>
<td>Emphasis</td>
<td>3</td>
<td>Low</td>
<td>Place the emphasis on the correct word.</td>
</tr>
<tr>
<td>Justification</td>
<td>3</td>
<td>High</td>
<td>She had no justification for her behavior.</td>
</tr>
<tr>
<td>Hypothesis</td>
<td>4</td>
<td>Low</td>
<td>The hypothesis was proven wrong.</td>
</tr>
<tr>
<td>Subsequently</td>
<td>4</td>
<td>High</td>
<td>The book was very popular and was subsequently translated into several languages.</td>
</tr>
<tr>
<td>Implementation</td>
<td>4</td>
<td>High</td>
<td>The implementation of the plan was harder than the students expected.</td>
</tr>
<tr>
<td>Summarization</td>
<td>4</td>
<td>High</td>
<td>Summarization is a great strategy for improving comprehension.</td>
</tr>
<tr>
<td>Challenge</td>
<td>5</td>
<td>Low</td>
<td>It is important to face a challenge with courage.</td>
</tr>
<tr>
<td>Stability</td>
<td>5</td>
<td>High</td>
<td>More supports were needed to maintain the stability of the structure.</td>
</tr>
<tr>
<td>Externalization</td>
<td>5</td>
<td>High</td>
<td>The act of externalizing something is called externalization.</td>
</tr>
<tr>
<td>Orientation</td>
<td>5</td>
<td>High</td>
<td>It is easy to lose one’s orientation in a maze.</td>
</tr>
<tr>
<td>Revealed</td>
<td>6</td>
<td>Low</td>
<td>The inspection revealed many problems.</td>
</tr>
<tr>
<td>Preceding</td>
<td>6</td>
<td>High</td>
<td>The preceding act was based on a true story.</td>
</tr>
<tr>
<td>Success</td>
<td>7</td>
<td>Low</td>
<td>The album was an instant success.</td>
</tr>
<tr>
<td>Innovation</td>
<td>7</td>
<td>High</td>
<td>Innovation should be encouraged in school.</td>
</tr>
<tr>
<td>Guidelines</td>
<td>8</td>
<td>Low</td>
<td>There are guidelines for changing a tire in the owner’s manual.</td>
</tr>
<tr>
<td>Inspection</td>
<td>8</td>
<td>Low</td>
<td>All passengers were seated during the inspection of luggage on the plane.</td>
</tr>
<tr>
<td>Automatically</td>
<td>8</td>
<td>High</td>
<td>She automatically assumed the best outcome would occur.</td>
</tr>
<tr>
<td>Scenario</td>
<td>9</td>
<td>Low</td>
<td>Consider the best and the worst scenario when solving a problem.</td>
</tr>
<tr>
<td>Coincidence</td>
<td>9</td>
<td>High</td>
<td>It was a coincidence that we were both in Georgia at the same time.</td>
</tr>
<tr>
<td>Conceive</td>
<td>10</td>
<td>Low</td>
<td>Did you conceive that notion independently?</td>
</tr>
<tr>
<td>Compilation</td>
<td>10</td>
<td>High</td>
<td>The compilation of all of our findings helped us write a strong group paper.</td>
</tr>
</tbody>
</table>
APPENDIX C
HUMAN SUBJECTS COMMITTEE APPROVAL
MEMORANDUM

Office of the Vice President For Research
Human Subjects Committee
Tallahassee, Florida 32306-2742
(850) 644-8673 - FAX (850) 644-4392

APPROVAL MEMORANDUM

Date: 08/10/2012
To: YiChao Wu <XXX@fsu.edu>
Address: XXXXXXXXXXXXXXXX
Dept.: EDUCATION
From: Thomas L. Jacobson, Chair
Re: Use of Human Subjects in Research
The linguistic profiles of spelling errors in 4th, 5th, and 7th grade students

The application that you submitted to this office in regard to the use of human subjects in the research proposal referenced above has been reviewed by the Human Subjects Committee at its meeting on 07/11/2012. Your project was approved by the Committee.

The Human Subjects Committee has not evaluated your proposal for scientific merit, except to weigh the risk to the human participants and the aspects of the proposal related to potential risk and benefit. This approval does not replace any departmental or other approvals which may be required.

If you submitted a proposed consent form with your application, the approved stamped consent form is attached to this approval notice. Only the stamped version of the consent form may be used in recruiting research subjects.

If the project has not been completed by 07/10/2013 you must request a renewal of approval for continuation of the project. As a courtesy, a renewal notice will be sent to you prior to your expiration date; however, it is your responsibility as the Principal Investigator to timely request renewal of your approval from the Committee.

You are advised that any change in protocol for this project must be reviewed and approved by the Committee prior to implementation of the proposed change in the protocol. A protocol change amendment form is required to be submitted for approval by the Committee. In addition, federal regulations require that the Principal Investigator promptly report, in writing, any unanticipated problems or adverse events involving risks to research subjects or others.

By copy of this memorandum, the chairman of your department and/or your major professor is reminded that he/she is responsible for being informed concerning research projects involving human subjects in the department, and should review protocols as often as needed to ensure that the project is being conducted in compliance with our institution and with DHHS regulations.

This institution has an Assurance on file with the Office for Human Research Protection. The Assurance Number is IRB00000446.

Cc: Barbara Foorman <bfoorman@fsu.edu>, Advisor
HSC No. 2012.8412
APPENDIX D

SAMPLE OF PARENTAL CONSENT FORM

Dear Parents,

My name is Yi-Chieh (Sophie) Wu and I am a doctoral student in the Reading Education and Language Art Program at Florida State University (FSU). I am working on my dissertation under the guidance of Dr. Barbara Foorman at the Florida Center for Reading Research (FCRR). I am interested in learning about spelling patterns demonstrated by children in elementary school. Helping teachers understand these patterns will lead to better spelling instruction and better writing performance by students. The research board at FSUS has agreed to allow me to conduct this study in 4th, 5th, and 7th grades. I would like to invite your child to participate in this research study. Please take a moment to read this form carefully before agreeing to allow your child to take part in this study. If you agree to allow your child to participate, please sign and return it to your child’s teacher.

The study: The goal of the study is to understand the use of linguistic knowledge in spelling by analyzing spelling mistakes made by 4th, 5th, and 7th grade students on a researcher-developed spelling dictation test. One visit per classroom will be scheduled during school hours. My colleague and I will administer the spelling test in one class period together. The spelling test will be group-administered using a paper-pencil format. If you agree to allow your child to take part, your child will be asked to check if the word is familiar to them or not, and then asked to spell the word on a blank sheet of paper. The test will take about 20 minutes to complete. Your child’s previous year’s FCAT 2.0 Reading Score will also be obtained for the purpose of this study.

Risks and benefits: The risk in this study is that your child might feel a bit frustrated when asked to spell difficult or unfamiliar words in the spelling test. However, your child can stop the test whenever he or she desires. There are no incentives offered for participation in this study to you or your child. It is expected that the results of this study will inform researchers, teachers, curriculum developers, and policy makers about spelling instruction in elementary school.

Confidentiality: The records of this study will be kept confidential, to the extent permitted by law. Once the results of the spelling test and FCAT reading scores are merged, a virtual ID will be created for each student. There will be no connection between your child’s name and the assigned virtual ID. It will not be possible to figure out your child’s answers. The responses of the spelling test will be kept secure for three (3) years after this study ends in a locked cabinet and office.

Voluntary Participation: Your child’s participation in this study is completely voluntary. Your decision whether or not to allow your child to take part will not affect your current or future relationship with FSU or with FSUS. If you decide to allow your child to take part, your child is free to skip any questions, or stop at any time. You are free to withdraw your child at any time without affecting your relationship with FSU or FSUS.

You may reach me at (XXXX) XXX-XXXX, or YYYY@my.fsu.edu. My advisor, Dr. Barbara Foorman, can be reached at (XXXX) XXX-XXXX, or XXXXXX@fsu.edu. Please feel free to ask any questions you have now, or at any point in the future. If you have any questions or concerns about your child’s rights as a research participant, you may contact the FSU Institutional Review Board (IRB) at (850)944-8333 or humansubjects@magnet.fsu.edu. You may also access their website at http://www.fsu.research.edu. You will be given a copy of this consent form for your records.

Please enter your child’s name and sign below if you give consent for your child to participate in this study.

Your child’s name: ___________________________ Your signature ___________________________ Date ___________________________

Sincerely,
Yi-Chieh (Sophie) Wu
Doctoral Candidate
Reading Education and Language Art Program
Florida State University

FSU Human Subjects Committee Approved on 9/10/2012. Void after 7/10/2013. HSC # 2012.8412
REFERENCES


Frith, U. (1978). From print to meaning and from print to sound, or how to read without knowing how to spell. *Visible Language*.


Moats, L. C. (2000). *Speech to print: Language essentials for teachers*


Morrison, D. F. *Multivariate statistical methods*


Stahl, S. A. (1990). "Beginning to read: Thinking and learning about print" by marilyn jager adams. A summary University of Illinois, Summary, P.O. Box 2276, Station A, Champaign, IL 61825-2276.


Wasowicz, J. (2007). What do spelling errors tell us about language knowledge?


BIOGRAPHICAL SKETCH

Yi-Chieh Wu had been a certified Special Educator for a decade in Taiwan before receiving the admission of Florida State University in 2009. She received her bachelor’s and master’s degree in Special Education with a focus in Learning Disorder and Emotional Behavioral Disorder from National Taiwan Normal University in 1998 and 2007, respectively. Yi-Chieh had worked as a chief teacher of resource classroom in Taipei Municipal Beitou Junior High School since 1999 and become a full-time member of Positive Behavior Support Team in 2008. During her teaching career, she had also attended intensive trainings and been certified as an Educational Diagnostician in 2000 and a Consultant Educational Diagnostician in 2006.

Under the guidance of Dr. Barbara Foorman, Yi-Chieh received her doctoral degree in the Reading Education and Language Art Doctoral Program with a focus in spelling. Her experiences as a special education teacher are the foundation of her current and future researcher interests and goals, which are to conduct scientific method oriented research, to translate evidence-based research results into practice, and to improve literacy development through effective professional development.