DIFFERENTIATING AMONG STUDENTS: THE VALUE ADDED OF A DYNAMIC ASSESSMENT OF MORPHOLOGICAL PROBLEM-SOLVING

by

Kathleen M. Stanfa

B. A. in English, University of California, Los Angeles, 1989

M. A. in English, New York University, 1994

M. Ed. in Special Education, Slippery Rock University, 1997

Submitted to the Graduate Faculty of

School of Education in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy in Special Education

University of Pittsburgh

2010

UNIVERSITY OF PITTSBURGH

SCHOOL OF EDUCATION

This dissertation was presented

by

Kathleen M. Stanfa

It was defended on

December 1, 2010

and approved by

Charles Perfetti, Ph.D., University Professor of Psychology;

Director, Learning Research and Development Center

Kimberly Gomez, Ph.D., Associate Professor, Department of Instruction and Learning

Christopher Lemons, Ph. D., Assistant Professor, Department of Instruction and Learning

Dissertation Advisor: Naomi Zigmond, Ph.D., Distinguished Professor of Special Education,

Department of Instruction and Learning

DIFFERENTIATING AMONG STUDENTS: THE VALUE ADDED OF A DYNAMIC ASSESSMENT OF MORPHOLOGICAL PROBLEM SOLVING

Kathleen M. Stanfa, Ph.D.

University of Pittsburgh, 2010

In this study, a dynamic assessment was used to measure the morphological analysis skills of participating sixth graders. The primary aim of this research was to examine the utility of this dynamic assessment and specifically to identify whether this measure can differentiate among students based on reading proficiency.

Twenty-seven participants, including three identified as needing special education services, underwent two testing sessions. Four assessments were administered: a standardized measure of receptive vocabulary, an oral reading fluency task, and two measures of morphological knowledge, one static in nature and the other a dynamic task which utilizes a standardized graduated prompting approach. Scores on a standardized reading outcome measure were also obtained for each participant.

Several significant results can be drawn from this study. Evidence was generated that the dynamic task had high estimates of internal consistency and seemed to perform similarly to other measures of morphological problem-solving found in the literature. Correlations between the dynamic task and measures of vocabulary knowledge and reading fluency were moderate to large in size and positive, also making it consistent with other studies and suggesting that the dynamic measure captures skills related to these critical literacy constructs.

Other interesting findings occurred in this study which extend our understanding of the role of morphological problem-solving in reading performance. First, the dynamic task emerged

from the analyses as a potentially useful screening measure, capable of differentiating sixth graders at risk for reading difficulty and including procedures and content more sensitive than a static measure of morphological knowledge. The predictive validity of the DATMA was also compared to the oral reading fluency task, a commonly used reading screening measure. Results showed that the dynamic measure added significantly to the prediction of reading outcomes. Exploratory analyses examined the use of the dynamic assessment as a supplemental screening measure to the oral reading fluency task. The dynamic assessment reduced the number of false positives, and in some cases, predicted reading outcomes as well as a combination of the two measures.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I wish to thank Dr. Naomi Zigmond for her expert guidance through this process. Her ability to view problems from multiple angles enriched this experience from the initial design phase through the interpretation and discussion of results. I am deeply grateful for all the opportunities she placed before me.

For their genuine interest and helpful suggestions, I want to thank my committee members, Dr. Charles Perfetti, Dr. Kimberley Gomez and Dr. Christopher Lemons. Each of these individuals devoted considerable time and energy to advance this dissertation. The project benefited greatly from their knowledge and experience. I am also grateful to Dr. Karen Wieland for insightful comments on an earlier draft of my manuscript.

I also extend my thanks to the school district in which this research was conducted. The support of the administration, board, staff, parents and students made this research possible.

A number of others guided me and helped me to navigate the hurdles along the way. I would especially like to thank Dr. Amanda Kloo, Dr. Eileen St. John, and Maria Barron. Their friendship made the journey that much more meaningful.

Finally, I want to thank my daughters, Rachel, Hanna, and Danielle, for their patience and unwavering support.

1.0 INTRODUCTION

In recent years, a large and convincing body of experimental evidence has accumulated which demonstrates the importance of phonemic awareness in the early stages of reading acquisition (Adams, 1990; National Reading Panel, 2000; Snow, Burns, & Griffin, 1998). However, as children move beyond the first few years of school, other aspects of linguistic awareness also become critical to literacy development. Among these, and far less well understood, is the role of morphological awareness. Morphological awareness, which refers to the ability to recognize and manipulate the smallest units of meaning in language (morphemes), is important in developing reading and reading-related skills (Carlisle, 2000; Carlisle & Fleming, 2003; Carlisle & Nomanbhoy, 1993; Deacon & Kirby, 2004; Mahony, Singson, & Mann, 2000; Nagy, Berninger, Abbott, & Vaughn, 2003; Tyler & Nagy, 1989, 1990). Moreover, the role of morphological awareness in literacy development appears to strengthen as children grow, even as the impact of phonemic awareness wanes.

Research indicates that, like phonemic awareness, morphological awareness supports the process of learning to decode written words (Carlisle & Stone, 2005; Deacon & Kirby, 2004; Singson, Mahony, & Mann, 2000). Further, because morphemes serve phonological, orthographic, semantic and syntactic functions, morphological knowledge has been shown to facilitate a wide range of literacy skills beyond word recognition, including spelling (Leong, 2000; Nunes, Bryant, & Bindman, 1997; Tsesmeli & Seymour, 2006), vocabulary learning

(Anglin, 1993; Carlisle, 2000), and reading comprehension (Carlisle, 2000; Mahony, 1994). In the field of cognitive psychology, current models of language processing hypothesize that the mental lexicon is organized morphologically, with access to polymorphous words based on activation of their constituent morphemes (Taft, 2003; Schreuder & Baayen, 1995; Verhoeven & Perfetti, 2003). Altogether, these findings suggest that learners with well-developed morphological knowledge may have an advantage in acquiring and recalling morphologically complex words.

Morphological awareness is believed to follow a developmental trajectory beginning in early childhood. While morphological knowledge continues to develop throughout school and into adulthood, it is most sharply evidenced in the vocabulary growth of upper elementary students (Tyler & Nagy, 1989; Nagy, Berninger, & Abbott, 2006). When Anglin (1993) asked children in grades 1, 3, and 5 to explain the meanings of morphologically complex words, among the most important findings were that older children more frequently referred to the morphemes in the words than did the younger children. Subsequent research has found significant interactions between word knowledge, morphological awareness, and reading comprehension for students in the intermediate grades (Carlisle, 2000; Nagy, Berninger, & Abbott, 2006). Contextual factors also play a large role in this phenomenon: literacy demands increase sharply after the fourth grade. In particular, the number of words students are expected to know expands dramatically. Text materials at these levels contain higher proportions of morphologically complex words. In fact, in the upper elementary and intermediate-level grades, about 60% of the unfamiliar words encountered by students in texts are morphologically complex (Nagy et al., 1989).

Many children, particularly those with reading problems, have difficulty understanding morphologically complex words. Successful readers who encounter words which they do not know engage in strategies that enable them to generate word meanings. A number of theories have been put forth in an attempt to identify the strategy that best yields accurate generation of word knowledge (Anderson & Freebody, 1983; Baumann, Kame'enui, & Ash, 2003; Beck, McKeown, & McCaslin, 1983; Graves & Hammond, 1980; McKeown, Beck, Omanson, & Perfetti, 1983; Nagy, Anderson, & Herman, 1987; Nagy, Anderson, Schommer, Scott, & Stallman, 1989; Nagy, Berninger, Abbott, Vaughan, & Vermeuleun, 2003; Nagy & Scott, 2000; Reichle & Perfetti, 2003; Rosenblatt, 2003; Stahl & Fairbanks, 1986; Verhoven & Perfetti, 2003; White, Power, & White, 1989). However, much of the research regarding successful word acquisition focuses on typically-developing readers.

Estimates are that typical children learn nearly 3,000 words a year between third and twelfth grades (Nagy & Herman, 1984). Because it would be impossible to teach directly the meanings of all these words, some argue that vocabulary instruction should teach skills and strategies that enable students to become independent word learners (Nagy & Anderson, 1984). These strategies include contextual analysis and morphemic analysis. Morphemic analysis, alternatively referred to as morphological problem-solving, involves deriving the meaning of a word by combining the meaningful parts (morphemes) of the word (Spencer, 2001). These parts with meaning include prefixes, suffixes, and roots. Morphemic analysis is thought to be a useful vocabulary learning tool and has recently been promoted as a strategy to increase the vocabulary of English language learners (Carlo et al., 2004; Kieffer & Lesaux, 2008).

Yet, for some children, word knowledge does not develop as effectively as it does for typical children. Frequently referred to as the "fourth grade slump," previous research has revealed a pattern whereby some children, many of whom were successful readers of narrative texts in the early grades, now struggle with the differing demands of content-area textbooks. These reading demands involve increasing word length & complexity (Juel, 1988). Simultaneously, differences in exposure to print between good and poor readers begin to grow (Cunningham & Stanovich, 1991). As the word gap widens, a pressing need for strategic vocabulary instruction for students who are lagging behind is created. Proposed solutions to narrow the word gap include wide reading (Anderson & Nagy, 1992) and explicit teaching of low frequency vocabulary (Beck, McKeown, & Kucan, 2002). Increasingly, researchers are advocating the use of morphological problem-solving to teach vocabulary.

Some authors (Nation, 1990; Sirles, 1997) have suggested that learning the meanings of Greek and Latin roots is critical because estimates regarding the number of words that are derived from Greek and Latin roots range from approximately half (Moore et al., 1997) to as much as 65% of what is termed our academic vocabulary (Sirles, 1997). Thus, teaching one word part, like a root (e.g., "*dict*" or "*port"*) has the potential to enable students to determine the meaning of many words since word families comprised of as many as 20-30 words tend to be organized around a given root (Baumann & Kame'enui; 2004; Graves, 2006; Marzano, 2004; Nagy & Anderson, 1984; Nation, 1990; Stahl & Nagy, 2006; White, Power, & White, 1989). Nation (1990) describes the skills needed for morphemic analysis: (1) the ability to break new words into their morphological parts, (2) an ability to connect meanings to those parts, and (3) an ability to combine the meaning of the parts to determine the definition of the whole word. Adolescents with reading deficits not only engage in little independent reading, but because of their learning problems, they have difficulty in applying specific strategies for independently learning new words (Pany, Jenkins, & Schreck, 1982).

1.1 SIGNFICANCE OF THE PROBLEM

Significant numbers of adolescents in the US read below grade level. The most recent NAEP data indicate that 36% of fourth graders and 27% of eighth graders score at the *Below Basic* level in reading (Perie, Grigg, & Donahue, 2005; Grigg, Donahue, & Dion, 2007). Secondary-level students with learning disabilities (LD) experience the largest deficits, with 21% of these students estimated to be five or more grade levels below in reading (National Longitudinal Transition Study II, 2003). It is critical that targeted intervention begins as soon as possible for children identified with reading problems. Academic demands increase sharply in middle and high school. Reading materials become increasingly complex beginning in fourth grade as children are expected to be able to read to learn. High stakes testing and graduation requirements compound the problem. The impact of insufficient literacy skills for adults in the 21st century is enormous and the consequences are profound. As workplace literacy demands have increased, greater unemployment and underemployment plagues adults with poor reading skills. Fortunately, high quality literacy instruction at the secondary level has been shown to be effective in remediating severe reading deficits (Biancarosa & Snow, 2004).

The importance of vocabulary knowledge to overall academic success, especially in the area of reading and oral comprehension, is well documented (e.g., Catts & Kamhi, 1999; Nagy & Scott, 2000; Snow, Porche, Tabors, & Harris, 2007; Stahl, 1999). Recent reading reports (e.g., National Reading Panel, 2000; RAND Reading Study Group, 2002) have reinforced the central role of vocabulary in student achievement but have also pointed out that there has been relatively little attention given to determining the impact of various instructional strategies in raising student outcomes related to vocabulary acquisition. In light of the large number of words students need to learn annually (Nagy & Anderson, 1984), the limited time that teachers have available to teach struggling readers at the secondary level (Deshler, Schumaker, Bui, & Vernon, 2006), and the deficits students with disabilities experience, identification and validation of approaches for teaching vocabulary are needed to close the gap. It is important to understand the strategies students use and the knowledge they possess about the process of word learning. If more words can be learned through independent reading than from direct instruction, an effort should be made to investigate the process of how students derive word meanings through morphological analysis.

The *Reading Next* report, released in 2004 and targeted to the literacy needs of secondary-level students, also highlights the importance of vocabulary to reading comprehension. This report makes a number of recommendations to improve literacy learning in adolescents. Among these is a call for assessment which is tied to instructional planning and part of a comprehensive progress monitoring system. Assessments should yield clear, specific, and meaningful information that teachers can use to inform instruction. More comprehensive measures should be designed to gather both qualitative and quantitative data. Furthermore, these ought to be able to capture growth and gauge student progress (Biancarosa & Snow, 2006).

At the same time, a broader understanding of the process by which individuals with reading disabilities (RD) gain word knowledge is needed. The strategies these individuals employ to generate meanings for unknown words is important to understand. Further investigation into the morphological knowledge of struggling readers will likely increase our understanding of (1) the ways in which impairments associated with RD impact performance on word knowledge tasks compared to the performances of typically-developing students; (2) the specific strategies used by students with RD when tackling unknown words in text; (3) and the types of interventions and supports that may be most effective for these students. What is needed is the development of an assessment model which can be utilized to identify students with specific deficits in morphological knowledge, to make diagnostic decisions and plan instruction, and to measure the effectiveness of these interventions.

Dynamic assessment (DA) presents an alternative to traditional assessment that can be designed to incorporate current knowledge about reading processes. DA can yield detailed information about an individual's level of functioning that is not attainable through traditional assessment (Kletzien & Bednar, 1990; Perkins, 1988; Spector, 1992). The Dynamic Assessment Task of Morphological Analysis (DATMA; Larsen & Nippold, 2007) was developed as a measure of morphological analysis skill that can be used to better understand the underlying thought processes of students as they encounter unfamiliar morphologically complex words. The DATMA is potentially useful as a diagnostic tool to help plan instruction and as a progress monitoring tool for use in data-based decision making. The DATMA's utility and validity to assess students with reading problems has not been investigated and is worthy of attention.

1.2 PURPOSE OF THE STUDY

Morphologically complex words occur with increasing frequency in school texts as children progress through grade levels. In fact, it is largely through the learning of morphologically complex words that the curriculum becomes accessible at the upper elementary and secondary-school levels. The existing literature indicates that children benefit from the use of morphological problem-solving strategies to learn the meanings of unfamiliar words. Research is needed to design and evaluate developmentally appropriate, systematic and explicit intervention programs which promote the use of morphological analysis as a word learning strategy for older students with significant reading problems. Establishing valid and reliable measures of morphological knowledge is critical. Dynamic assessment is one potentially useful measurement technique and may offer important information not available through conventional techniques alone. Because dynamic assessment provides multiple opportunities for success it may be better able to differentiate among various low achieving students than static measures.

The purpose of this study is to add to the literature which explores morphological problem-solving skills in relation to broader language abilities in vocabulary and reading comprehension. More specifically, the study is designed to achieve the following research objectives: first, to investigate whether the Dynamic Assessment Task of Morphological Analysis (DATMA) can differentiate among students with similar reading levels and, in particular, whether this instrument can discriminate among students with the most significant reading problems. The second objective is to examine whether the DATMA contributes new information over and above that obtained with conventional measures of reading fluency, vocabulary knowledge and morphological awareness.

1.3 RESEARCH QUESTIONS

The following research questions were addressed in this study:

1. Do results obtained from the DATMA demonstrate concurrent validity with other measures presumed to be assessing the same construct or closely related skills?
    a. What is the correlation between the DATMA and a static measure of morphological knowledge?

    b. What are the correlations between the DATMA and standardized measures of vocabulary knowledge and reading fluency?
2. To what extent does the DATMA predict performance on a subsequent measure of reading proficiency? How does this compare to the predictive power of a static measure of morphological knowledge using the same criterion measure of reading performance?
3. Does the DATMA identify clusters of students who would be selected for different levels of reading intervention (i.e., benchmark, strategic, or intensive) based on their performance on an oral reading fluency test?
4. As a screening measure, how accurately does the DATMA classify students as at-risk for reading difficulty?
    a. What are the sensitivity and specificity levels of the DATMA to classify students as at-risk as defined by oral reading fluency cut scores?
    b. How does the classification accuracy of the DATMA compare to that of the oral reading fluency task in predicting proficiency on a standardized reading outcome measure?

1.4 RELEVANCE OF THE STUDY

Given the close associations between derivational morphology, word knowledge, and reading comprehension (Anglin, 1993; Carlisle, 2000; Mahony, 1994; Nagy et al., 2006) research which focuses on this aspect of literacy development is warranted. Little is known about students' comprehension of morphologically complex words that occur in textbooks that are used to teach challenging subjects such as science, mathematics, social studies, and literature. Which aspects of morphological knowledge should teachers stress in order to successfully teach morphological analysis as a cognitive strategy for vocabulary learning? Some investigators have emphasized the teaching of common prefixes and suffixes as a method of helping children learn the meanings of complex words (Henry, 1993). However, this alone is insufficient (Anglin, 1993). In recent literature reviews on this topic, Baumann and Kame'enui (2003) and Ebbers and Denton (2008) have suggested that the relationship between morphological knowledge and vocabulary learning should be viewed as a priority area for new research.

In today's classrooms, students with disabilities are expected to complete grade-level work and be held to grade-level accountability standards. In order to effectively access the general education curriculum, it is essential that students are able to develop the vocabulary needed to understand the texts used. Academic discourse--which is required to learn about and discuss complex topics--is largely comprised of morphologically complex words. It is important, therefore, that educators identify appropriate assessment measures which target vocabulary learning strategies. Timely and effective intervention depends on valid and reliable assessment tools which can guide the development of high quality instruction that addresses students' individual learning needs. Remediating vocabulary deficits is critical to improving students' academic success at the secondary level. Evaluation procedures are necessary to screen students for failure to learn new vocabulary through morphological problem-solving. These procedures can further help identify specific deficits, monitor student progress, and inform instructional practice.

What is needed is a suitable instrument that can probe students' ability to use morphemic analysis and document the use or non-use of this word learning strategy. Utilizing dynamic measurement tools offers added benefits because these procedures reveal the degree of knowledge a student has as well as the student's ability to respond to intervention. However, in order to further the use of a dynamic measure of morphological analysis skill, research must first be conducted which validates the assessment instrument itself, particularly for use with students with disabilities. The use of valid dynamic assessment procedures has the potential to substantially reduce the time teachers need to identify at-risk children. In turn, these children will be likely to spend less time in instructional programs which are not meeting their needs.

The current study adds to the research base in a number of ways. First, this study will provide further examination of the DATMA measure as a diagnostic tool of word-learning ability. The DATMA has demonstrated adequacy to evaluate word-learning potential in a typically developing population (Larsen & Nippold, 2007). Further research is necessary to explore use of this tool for other populations and to refine the procedures. Secondly, this study adds to the literature on dynamic testing and the relationship of morphological knowledge to reading proficiency.

1.5 DEFINTION OF TERMS

Dynamic Assessment (DA): an interactive approach to conducting assessments within the domains of psychology, speech/language, and education that focuses on the ability of the learner to respond to intervention. DA is not a single procedure, but both a model and philosophical orientation.

Graduated Prompt Procedure: a method of dynamic assessment which utilizes a series of hints or prompts presented in a graduated sequence of increasing explicitness.

Instructional Scaffolding: the provision of supports which allow learners to complete a task.

Morpheme: the smallest linguistic unit which has semantic meaning.

Morphological Analysis/Morphological Problem-Solving: these two terms are used interchangeably throughout the literature. They refer to the process of recognizing the morphemic constituents in complex words, accessing the meaning of those morphemes, and recombining them to determine the meaning of the larger word.

Morphological Awareness: the awareness of and the ability to manipulate the smallest units of meaning in language (morphemes).

Morphology: branch of linguistics that deals with the structure and content of word forms.

Test-Teach-Test Procedure: a method of dynamic assessment which embeds intervention within the assessment procedure. Most typically, there is a pretest then an intervention and then a posttest. This allows the assessor to determine the response of the student to the intervention.

Zone of Proximal Development (ZPD): Vygotsky's term for "the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance, or in collaboration with more capable peers" (Vygotsky, 1978, p. 86). Vygotsky argued that rather than examining what a student knows to determine performance level, it is better to examine their ability to solve problems independently and their ability to solve problems with the assistance of an adult (Berk & Winsler, 1995). When learners are not able to perform a task independently, the ZPD marks that which they can do with assistance and presumably that which they will be able to do in the future on their own, after appropriate instruction has been provided.

2.0 REVIEW OF THE LITERATURE

This review of the literature is organized as follows: first, an explanation of morphology and morphologically complex words will be offered; second, a look at the role of morphology in literacy development will be explored--including the role of morphology in vocabulary growth and reading comprehension; and finally an overview of dynamic assessment will be presented.

2.1 MORPHOLOGY

Morphology refers to the structure of words in terms of morphemes, or "minimal meaningful elements" (Bloomfield, 1933). These smallest units of meaning include roots, prefixes, and suffixes. For example, the word *unkindness* is made up of three morphemes: the prefix *-un*, the root *-kind*, and the suffix *–ness*. The meaning of a word is derived from the combined meanings of the morphemes that make up that particular word. For example, the *er* in *worker* denotes "one who works" and the *un* in *unsafe* indicates "not safe." In addition, morphemes can also serve as syntactic markers, indicating grammatical class. Morphemes combine in systematic ways, conforming to specific rules and pairing meaning and word class in a relatively fixed manner.

Morphemes are divided into two types: free forms which can stand alone, such as the morpheme *love* in *lovely*. Others, such as the suffix*–ly* in the previous example, are bound forms and must appear together with other morphemes. Bound morphemes include prefixes, which change the meaning of the root, but not its grammatical class; suffixes, which change both the meaning and the grammatical class of a root; and inflections, which indicate number, person, tense or case (Arnbak & Elbro, 2000). Morphology is further divided into three categories: inflections, derivations, and compounds. Inflections attach to a base word and mark categories such as number, person, tense, and case (e.g., the past tense marker *–ed*). Compounds involve two root words which when joined together make a new word (e.g., *doghouse*). Derivation refers to the formation of new words from existing words (e.g., *teacher* from *teach*.) Derived words can also be inflected: *teachers* from *teacher*.

A further distinction can be made between types of derivational suffixes. Suffixes attaching to base words to form derivations can be described as neutral or non-neutral. Neutral suffixes have a wide range of applicability. The primary restriction on these suffixes is their subcategorization for the part of speech of the morpheme to which they can attach (e.g., *-er* can attach to virtually any verb to form an agentive). Non-neutral suffixes do not have the same broad range of applicability. For example, root words containing *–ceive* take *–tion* (e.g., *receive/reception*), and root words ending in *–fer* take *–ence* (e.g., *prefer/preference*).

It is generally accepted that derivational morphology involves more complex skills than inflectional morphology (for a differing perspective on this issue, see Rabin & Deacon, 2008). Three major aspects of awareness of derivational morphology have been identified in previous research (Tyler & Nagy, 1989). First, relational knowledge refers to the ability to recognize the root, or stem, of polymorphic words and to understand the relationship between the stem and the affix(es). Second, syntactic knowledge involves understanding the alteration of part of speech produced by various derivational affixes. Finally, distributional knowledge is concerned with insight into how affixes are constrained by the syntactic category of the stems they attach to. For example, -*ly* attaches to adjectives, but not to nouns, so that *carefully* is correct in English while *carely* is not. Distributional knowledge is thought to be more challenging than the other aspects of derivational morphology and likely among the last to develop for learners (Ku & Anderson, 2003; Tyler & Nagy, 1989).

Relational knowledge is a prerequisite for the other forms of derivational knowledge. Children appear to develop a rudimentary knowledge of derivational morphology, or the ability to recognize a familiar stem in a derivative, before fourth grade. Knowledge of the syntactic properties of derivational suffixes appears to increase through eighth grade. Knowledge of the distributional properties of suffixes also increases, with sixth-grade students showing an increase in overgeneralization errors parallel to those found for inflectional suffixes in younger children (Tyler & Nagy, 1989).

2.1.1 Features of Morphologically Complex Words

Empirical evidence has identified three key features of morphologically complex words that impact how these words are processed: semantic transparency, productivity, and phonological shift effects. Sandra (1993) studied the effects of semantic transparency on acquiring morphologically complex words. Single morpheme and polymorphous nonsense words were taught to two groups of undergraduates, an experimental group received semantic explication of the words, and a control group did not. Both groups were faster to identify the polymorphous words than single morpheme nonsense words. No interaction effect was found for the semantic explication condition. Sandra concluded that the undergraduates were able to spontaneously recognize morphemic patterns and attributed this ability to the students' well-developed morphological knowledge. Children, in contrast, are still developing morphological knowledge and may benefit from further semantic instruction that adults do not require.

In another study of semantic transparency, this time using priming techniques, Marslen-Wilson, Tyler, Waksler, & Older (1994) found that stems effectively primed the recognition of their derivations but did not prime morphologically complex words that appeared to be relatives based on orthographic and phonological similarity. Thus, *govern* primed *government* but *depart* did not prime *department*. These researchers concluded that semantically transparent morphologically complex words are stored as separate lexical items while semantically opaque polymorphous words are stored as single entries.

A morpheme's productivity refers to the extent to which it can be combined with a variety of stems or affixes. Languages place constraints on the ways in which morphemes can be combined. For example, in English, affixes are attached to stems in a prescribed order. Hence, *wonderfully* is correct but *wonderlyful* is not. Also constrained is the type of affix that can be utilized with a particular stem. For example, in English, -*er* is a highly productive suffix to indicate the doer of an action. For this reason, although a magician is someone who performs magic, someone who uses a microwave is more likely to be called a microwaver than a microwavician. It is likely that frequency effects of highly productive affixes influence their acquisition by learners (Carlisle, 2003).

In English, the phonological relationship between stem words and their derived forms can vary. For example, *electric/electricity* exhibits a stress shift and consonant alteration. Derived words that undergo phonological shifts are more difficult to acquire (Jarmulowicz, Taran & Hay, 2008), especially for poor readers (Carlisle, Stone, & Katz, 2001). In a study comparing skilled and less-skilled secondary-level students, speed and accuracy of naming derived words of two types—transparent base form (e.g., *culture/cultural*) versus opaque form (e.g., *nature/natural*), morphologically complex words which did not alter pronunciation were easier to name (Singson, Mahony, & Mann, 2000).

2.1.2 Morphological Awareness

Morphological awareness refers both to the sensitivity to the internal morphological structure of words and the ability to manipulate these smaller intraword segments. Because morphemes have semantic as well as phonological and syntactic properties, and because they pair sound and meaning in systematic ways, facility with the morphological constituents of language involves both word knowledge and metalinguistic skill. In addition to needing semantic knowledge of root words, learners must also be able to apply syntactic knowledge—or how a derivational affix marks a word for a new syntactic category (e.g. the noun *mountain* becomes an adjective when suffixed with *–ous*) to accurately analyze morphologically complex words. Knowledge of morphological relationships among words enables students to expand their vocabulary by applying morphological principles. For learners, recognizing and understanding word families, stems, and affixes can provide students with clues to the meanings of literally thousands of words and enhance their comprehension of connected text.

Important questions regarding the nature of the construct of morphological awareness remain unanswered. Is morphological awareness a single construct? Or, is morphological awareness comprised of multiple dimensions? One key issue that has been raised in previous research attempts to separate implicit versus explicit knowledge of morphology and conscious versus unconscious control of morphological analysis. Knowledge of morphology, like other types of linguistic knowledge, can be tacit, meaning we use it when we need it without necessarily consciously reflecting upon it. In fact, it is possible to be able to use morphemes correctly without being able to explain how it's done. On the other hand, some tasks designed to measure morphological knowledge require explicit understanding and manipulation of morphemes. It appears very likely that morphological awareness is reciprocal with other literacy skills, but which specific dimensions of morphological awareness contribute to which aspects of literacy development have not been fully explored.

There are a variety of tasks that have been designed to measure morphological awareness. Some utilize items from different levels of phonological transparency: no change (*person-personal*), stress shift/vowel change (*parent-parental*), consonant change (*relate-relation*), vowel change (*deep-depth*), and silent letter (*sign-signature*). Administration of tasks also varies. Some are written; participants silently read the items before marking their answers. Others include both an oral-plus-written component; the experimenter reads each item aloud while simultaneously presenting the task in written form. This type of administration is designed to decrease demands on decoding ability, which theoretically should result in a more accurate measure of morphological knowledge. Other times, tasks are administered entirely orally; participants respond to the tester's question without having a written form of the task available. Some tasks require only recognition, others production of morphemes. Some tasks require participants to decompose morphologically complex words (or pseudowords); others require composition.

Measures of relational knowledge include those employed by Berko (1958), Carlisle (1995), Derwing (1976), Mahony (1994) and Mahony, Singson, and Mann (2000). These measures asks participants to decide if a second word is derived from the first word (e.g., q*uick/quickly*; *moth/mother)*. Generally, in this type of assessment, items are presented visually while the experimenter reads the items aloud. A similar measure of morphological relatedness knowledge is the "comes from task:" here participants are asked if a larger word contains a smaller word (e.g., Rubin, 1988). Morphological relatedness judgment tasks are also used (Derwing, Smith, &Wiebe, 1995). In these tasks, participants identify whether pairs of words are morphologically related or not.

In addition to these recognition tasks, some measures designed to assess students' knowledge of the relations between base and derived forms include production tasks (Carlisle, 2000). The following is an example of a task which requires decomposition of a derived form: given a derived word such as *driver*, children are asked to complete a sentence such as "children are too young to __________." Production of a morphologically complex form from a simple form may use a real word (e.g., Carlisle, 1988; 2000) or a nonsense word (e.g., Berko, 1958; Brittain, 1970; Elbro, 1989). In the reverse task, the root word is given and children are asked to produce the derivation (composition) in order to finish a sentence (e.g., the word is *farm*. The sentence is: "my uncle is a __________"). The sentence analogy task assesses the productive ability to manipulate morphemes. In the sentence analogy task, participants are given two example sentences such as "Tom helps Mary" and "Tom helped Mary." They are asked to carry out the same transformation on another sentence, such as "Tom sees Mary" (Deacon & Kirby, 2004).

Another frequently used assessment is the derivation suffix choice test [e.g., Mahony (1994), Singson, Mahony, and Mann (2000), Nagy, Diakidoy, and Anderson (1993), and Tyler and Nagy (1989; 1990)].   Given four options of a base with derivational suffixes, participants must choose the one that fits a particular sentence. Variations include use of real words with plausible but novel suffixes attached (e.g *dogless*): "when he got a new puppy, he was no longer *dogless*s." A third type uses nonwords—where knowledge of derivational suffixes can be measured without interference of semantic content. These tasks measure syntactic and distributive knowledge of morphemes.

To measure semantic knowledge of morphemes, the morpheme identification task assesses the ability to distinguish different meanings across homophones. For each item, two different pictures are presented simultaneously to the child and each of the pictures is labeled orally by the experimenter. The child is then given a word or phrase containing the target morpheme and asked to choose from between the two pictures the one that best corresponds to the meaning of that morpheme. For example, a picture of "my son" with a picture of "the sun" shown; child asked to select the picture with the same meaning as "son" in "grandson." The morphological completion task asks participants to create new meanings by combining morphemes.

Other measures of morphological awareness do not require participants to manipulate morphological units. A morphological variant of the repetition priming paradigm presents a target word preceded by a morphologically related word; accuracy and/or response times are then compared to priming by non-morphologically related words. The fragment completion task is a variant of this type of task where participants are required to complete a word with missing letters (e.g., *t_rn*), first primed by a morphological relative (*turning*) and then by an unrelated word which is matched for length and orthographic similarity to the target word (*turnip*). (see Feldman, Rueckl, DiLiberto, Pastizzo, & Vellutino, 2002) These types of tasks are believed to measure implicit morphological knowledge rather than explicit knowledge.

A variety of methodological issues have been raised in measuring morphological awareness. Isolating the construct itself is difficult. Confounding variables include vocabulary knowledge, phonological awareness, and decoding ability. Due to the semantic content of morphemes, it isn't surprising that vocabulary knowledge is highly correlated with morphological awareness (Fowler & Liberman, 1995; Ku & Anderson, 2003; McBride-Chang et al., 2005). Because vocabulary size is a potential confound when measuring morphological awareness, research designs frequently incorporate some method to control for it. One approach used to address this issue is to include a vocabulary assessment as a covariate. An alternate approach utilizes pseudowords in order to remove the word familiarity effect. Morphological awareness is also highly correlated to phonological awareness. As such, various methods have been used to account for any shared variance between these two variables. Another potential confound when measuring morphological awareness is decoding ability. For this reason, morphological awareness tasks are frequently administered orally. One drawback to the use of orally administered measures is the increased demand placed on short-term memory that occurs as a result. Employing grammaticality judgment tasks is one way to lower the demand on short term memory. Another solution is to provide a written version of the test and read items aloud to the participants.

2.1.3 Development of Morphological Knowledge

How do children become morphologically aware? The prevailing theory is that the acquisition of morphological knowledge follows a specific developmental course beginning in the early elementary grades and continuing through adulthood, with the most rapid growth occurring between fourth and eighth grade (Jarmulowicz, Hay, Taran, & Ethington, 2007; Nagy, Berninger, Abbott, & Vaughan, 2003). Unlike phonology, morphology draws on other aspects of linguistic structure beyond the alphabetic principle. Therefore, its influence on reading and writing ability may be important at different stages than phonological awareness. Research suggests that awareness of morphological structure and the skill with combining morphemes becomes particularly important in the upper elementary and intermediate-level grades as estimates indicate that about 60% of unfamiliar words encountered by students in these grades are morphologically complex (Nagy, Anderson, Schommer, Scott, & Stallman, 1989).

2.1.4 Morphological Processing

Written English is a morphophonemic language: that is, the semantic relationships among individual morphemes are preserved in derived forms, even at the expense of orthographic changes (e.g., *active*/*activity*) and/or phonological shifts (e.g., *decide*/*decision* and *sign/signature*). Consequently, English is considered a relatively deep orthography. In contrast, shallow orthographies, such as Spanish, offer consistent relations between phonemes and letters. English, instead of being strictly alphabetic, is a deeper transcription that depends on morphemes as well as phonemes to encode words (Chomsky, 1970). Deep alphabets transcribe spoken words at a lexical level of representation rather than as a string of phonemes; for example, *health* makes evident, through its spelling but not its pronunciation, its relationship with the root *heal*. Because English orthography represents both phonology and morphology, it is reasonable to expect that an explicit awareness of both types of language structures would be advantageous for mastery of written language. In fact, contemporary psycholinguistic research suggests that morphological awareness plays an important role in reading in deep orthographies. Models of cognitive architecture which seek to explicate how humans process language are looking toward morphology to lend explanation.

Debate continues about how morphologically complex words are dealt with in the mental lexicon. Do affixes and stems have their own entries or are morphologically complex words stored independently? Several models of morphological processing have been proposed, ranging from full listing for all words to total decomposition. The earliest models of morphological processing posited that only morphemic units are stored in the lexicon, without any whole word listings (Taft & Forster, 1975). Taft and Forster (1975) concluded that single morphemes constitute the access code to polymorphemic words in the mental lexicon. These researchers found that morphemic non-word stems like *trieve* (*retrieve*) took longer to recognize than non-morphemic letter strings with the same degree of orthographic similarity to real words (e.g., *pertoire*/*repertoire*).

More recently, others have argued that morphological information is represented through links between whole word representations of related words (Bybee, 1995; Cole, Beauvillain, & Segui, 1989). Hence, words like *sweetness* and *sweetly* would have their own whole word representations, but would also be linked to each other and to the representation of their stem *sweet* and other members of their morphological family (Meaunier & Longtin, 2007). Interactive models propose a dual lexical route: both a full-form representation and a parallel parsing route in which complex words are segmented and their meaning computed from word parts are employed in parallel (Schreuder & Baayen, 1995). Which route will activate the relevant lexical entry is determined by the linguistic and distributional properties of the word. Dual route models theorize that both whole word storage and composition from morphological units can occur. For adults, lexical frequency may determine representation, with whole words represented when they reach a certain level of frequency (Jarmulowicz, Taran, Hay, & Ethington, 2007). How the lexical system develops in children is still not well understood.

Today, most widely-accepted models of reading assert that morphemes are represented in memory, including bound morphemes such as the suffix *–ly*. A convincing body of psycholinguistic research supports this theory of morphological processing. Two types of findings are particularly cogent. First, the frequency of the stem has been shown to affect the processing speed of morphologically complex words (Burani & Caramazza, 1987; Niswander, Pollatsek, & Rayner, 2000). Derived words with high frequency stems are recognized faster than those with low frequency stems when word length is controlled. The second class of findings show that recognition of stem words is facilitated by priming with morphologically complex relatives but not facilitated by orthographically similar but morphologically unrelated words (Feldman & Bentin, 1994; Fowler, Naps, & Feldman, 1985). This robust facilitatory effect strongly suggests that morphological structure is one of the organizing principles of the mental lexicon.

Taft (2003) proposes that mental representations of morphemes develop with language experience. Research supports this theory. For example, Nagy and colleagues (1989) found that the number of words in a word family as well as the frequency of derived words in a word family influenced the speed of recognition of the base word. Thus, frequent encounters with words which have the same base morpheme, such as *nice*, *nicely*, *niceness*, reinforce the mental representation of the morphemes in those words and speeds identification of other words containing those morphemes.

Although the representational architecture underlying the observed effects remains debated, the literature provides ample evidence of a role for morphology in representing polymorphous words. Support for a morphologically-organized mental lexicon implies that learners who have morphological knowledge may have an advantage in acquiring, retaining, and retrieving morphologically complex words, and would, in turn, achieve better literacy outcomes.

2.2 MORPHOLOGY AND LITERACY ACQUISITION

A large number of studies have been devoted to studying relations between morphological awareness and reading-related constructs. Empirical support for the role of morphological knowledge in literacy acquisition is well-established. In a seminal study, Berko (1958) demonstrated that very young children use morphological rules to derive and inflect nonsense words. Several recent studies have demonstrated a unique contribution for morphological awareness to a variety of literacy outcomes. In a longitudinal study of English-speaking upper-elementary students, Fowler and Liberman (1995) identified a positive effect of morphological awareness on the development of word decoding skills. Subsequent studies replicated these results and inferred that morphological knowledge plays a valuable role in learning to read. (Deacon & Kirby, 2004; Mann & Singson, 2003). Nagy, Berniger, and Abbot (2006) concluded that morphological awareness contributes to reading comprehension, reading vocabulary and spelling for students in grades 4-9, over and above shared variance with phonological decoding. Additional evidence of the importance of morphological awareness comes from studies which show that poor readers make more errors by omitting morphemes in writing (Duques, 1989;) and speaking (Henderson & Shores, 1982). Likewise, research has demonstrated that morphological awareness distinguishes skilled from less skilled readers (Champion, 1997; Leong, 1989a; Tyler & Nagy, 1989).

Given the extent of reading deficits among students with learning disabilities it is plausible that deficits in morphological knowledge could play a role in the problems these students experience in acquiring reading and reading-related skills. Studies show inflectional performance correlates with reading ability in grades 1 and 2 (Brittian, 1970) and deficits in the ability to apply inflectional suffixes is associated with reading disability as late as adolescence

(Elbro, 1989). Similarly, research suggests students with learning disabilities experience delays in acquiring derivational morphology (Vogel, 1983; Wiig, Semel, & Crouse, 1973; Windsor & Hwang, 1999).

Word experience is another significant difference potentially impairing students' morphological knowledge. Poor readers would likely have limited exposure to derived forms in contrast to their typically-developing counterparts. This may have special significance in terms of learning derivational morphology. Awareness of morphological structure and the skill with combining morphemes becomes particularly important in the intermediate grades. Nagy et al. (1989) estimated that about 60% of unfamiliar words encountered by students in these grades are morphologically complex. Comparing the difference in vocabulary knowledge between English speakers in fifth grade and third grade, Anglin (1993) found that nearly three-fourths of the words learned between third and fifth grade consisted of derived forms. Students who have limited morphological knowledge of English may struggle with word learning and, in turn, experience greater difficulty with text comprehension. Mann and Singson (2003) found that by fifth grade, students' ability to decode written words was better predicted by morphological than phonological skills. Together, these findings support a model of reading that takes into consideration morphological patterns.

2.2.1 The Role of Morphological Awareness in Vocabulary Learning

It may be that the most direct contribution of morphological skill to literacy outcomes is in the area of vocabulary learning. Because a substantial number of words in English have meanings which are predictable from the meanings of their parts, knowledge of morphology is believed to enhance vocabulary learning. In fact, more than half the words in English are morphologically complex (Anglin, 1993). Furthermore, polymorphous words are more common in written language than in spoken language--affixed words outnumber root words in printed school English by a factor of almost four to one (Nagy & Anderson, 1984). Moreover, Nagy and Anderson (1984) estimate that 60% of the unfamiliar words a reader encounters have meanings that can be inferred based upon the meanings of their component morphemes. A reader who understands word structure will likely be able to better predict the meaning of an unfamiliar word and thereby learn more words and consequently comprehend more of what he reads (Nagy, Berninger, Abbott, & Vaughn, 2003).

School-aged children encounter a great number of words they have never seen before. One strategy children can apply to make sense of unfamiliar words is to make use of the context to determine a word's meaning. Several studies indicate that this strategy is indeed useful (Graves, 1986; Wysocki & Jenkins, 1987). Nevertheless, the context does not always provide sufficient clues to determine the meaning of an unfamiliar word. One alternative strategy is to make use of a word's morphological structure. Carlisle (2000) found a link between morphological awareness and the ability of children in third and fifth grade to define morphologically complex words. In a detailed study, Anglin (1993) demonstrates that vocabulary growth is related to morphological knowledge. In this study, children in grades 1, 3, and 5 were capable of determining the meaning of unfamiliar words through the morphological structure of those words, a method Anglin refers to as "morphological problem-solving." Morphological problem-solving, sometimes called morphological analysis, is the process of breaking apart an unfamiliar word into parts and then recombining the parts into a meaningful whole. The children in Anglin's study participated in word definition tasks where they were asked to explain the meanings of morphologically complex words. Anglin found that students' knowledge of derived words increased sharply between grades 1 and 5. Older students referred to constituent morphemes more often than younger children did. As a result, there is a dramatic increase in the number of derived words known during elementary school. He estimates that students in grade 1 know about 3,000 root words and about 7,500 root words in grade five, a growth rate of about 1,100 words per year. During the same time span, students' knowledge of derived words increased from 1,800 to 16,000, a growth rate of 3,500 words per year.

Is it worth spending time directly teaching children how to derive the meanings of words through morphological problem-solving? White, Power, and White (1989) attempt to answer this question by quantifying the effects of morphological problem-solving on vocabulary learning. They examined the characteristics of affixed words, their composition (number and type), their analyzability, and the frequency of their occurrence in texts written for students at different grade levels. These researchers speculate that middle-school aged children successfully analyze at least 3000 prefixed words per year. White, et al. postulate a three-stage process of morphological analysis takes place when a reader encounters an unfamiliar word: first, the reader looks for prefixes and suffixes and removes them in that order; second, the reader retrieves the base from memory; and third, the reader combines the meaning of the base and affix(es). By calculating the proportion of words that are analyzable and estimating the likelihood that these words would be encountered by elementary children, White et al. estimates the probability that children will be able to derive the meaning. In this way we can determine the amount of vocabulary learning that occurs from morphological analysis.

First, White et al. calculated the prevalence of words with one of four frequent prefixes, the proportion of these words that could result in "misleading analysis," and the proportion which exhibit spelling and pronunciation changes in suffixation. The four major prefixes were *un*-, *re*-, *dis*-, *in*- (meaning "not" and also written as *im-, ir*-, and *il*-). These are the most common prefixes in the *American Heritage Word Frequency Book* (Carroll, Davies, & Richman, 1971) and account for approximately 58% of prefixed words in printed school English in grades 3-9. After a search of the American Heritage corpus for words formed by these four major prefixes, a stratified random sample of 257 prefixed words was chosen (about 70% of these words also had one or more suffixes as well). The root word was extracted from each and two adult raters independently generated from 1-3 familiar meanings and selected one as most familiar. These were cross-checked with dictionary definitions. The most familiar meaning of the root was compared to the most familiar meaning of its corresponding target (prefixed word). The most familiar meaning of the root was judged relevant and sufficient for understanding the target word in 60% or 155 of the 257 cases. Another 43 words were deemed understandable given knowledge of the second most familiar meaning of the root. Nine target words could be understood from the third most familiar meaning.

On the other hand, 19% or 50 cases were considered not analyzable and were excluded. Five words with disproportionately high frequencies were excluded as well. The remaining sample was then examined to determine the total number of occurrences in each grade subset (3-9) of target prefixed words per million words of running text in the American Heritage corpus. The occurrences for all of the words were summed and multiplied by 6.667 because the sample represented 15% of the major prefixed words in the American Heritage corpus. This provides an estimate of the number of analyzable target words per million words of text.

Study 2 turned to student factors and examined students' knowledge of prefixes, suffixes, and roots. To successfully use morphological problem-solving requires knowledge of prefix meanings, knowledge of suffix meanings, and knowledge of the base word's meaning. With a sample of 46 third graders and 45 fourth graders, the researchers first estimated children's root word knowledge. Root word knowledge was tested with a random sample of 84 words from the list of 194 analyzable words with all affixes stripped off. For third grade, the mean percentage of root meanings known was 35.9%. For fourth graders, it was 53.9%, a statistically significant difference. On a measure of prefix knowledge, the mean percentage of prefix meanings known was 50% for third graders and 60% for fourth graders. Next, a suffix recognition test with 15 suffixed words presented in random order was administered. Five items contained inflectional suffixes, five neutral derivational suffixes, and five non-neutral derivational suffixes. Each item presented the suffixed word and five multiple choice alternatives comprised of the last one, two, three, four, or five letters of the word. The words were read aloud and the students were told to circle the suffix. After correcting the raw data for guessing, the researchers were able to calculate a percentage of correctly identified suffixes of each type for students in grades 3 and 4.

The obtained estimates of morphological knowledge were then combined into a single number that represents the overall probability that a student can analyze a word successfully. For third graders, the probability of successfully analyzing target words is .12; for fourth graders, that probability is .27. White and colleagues admit these estimates are quite low. The researchers explain that these figures represent certain assumptions of the stage model they propose. As such, they recommend the estimates be regarded as illustrative, not definitive.

According to estimates, third graders ought to see 144 target words each year. For fourth graders, that estimate climbs to 502 target words per year. Fifth grade students would encounter 1,325 target words per year, more than twice that of fourth graders. The number of opportunities to analyze target words is likely to double again by seventh grade. If the other prefixed words in the American Heritage corpus are included, the researchers speculate that seventh graders could successfully analyze as many as 9,000 prefixed words a year. Based on these figures, White et al. support direct instruction in morphological instruction for students in grades 4 and above. They further outline potentially important features of that instruction based on their findings. As a general principle they urge teachers to base morphological instruction on frequently occurring affixes and to focus on providing students with strategies to use morphological clues to deal with unknown words.

Vocabulary and morphological awareness are very highly correlated and probably exert a reciprocal influence on one another (Carlisle & Fleming, 2003; Nagy, Berninger, Abbott, Vaughn, & Vermeulen, 2003; Singson et al., 2000). According to at least one study, morphological awareness may be the strongest predictor of vocabulary growth in the upper elementary grades (Nagy, Berninger, & Abbott, 2006). Students with larger vocabularies tend to have a superior understanding of morphology. The relationship is likely bidirectional: understanding morphology leads to vocabulary growth and vocabulary knowledge improves students' understanding of morphology. If so, then teaching morphology in the context of general vocabulary instruction is warranted. This is consistent with findings from Tyler and Nagy (1987) who demonstrated that on a multiple choice test, sixth and seventh graders could recognize the meaning of low frequency derivations that had high frequency stems. This, they conclude, indicates students use morphological analysis. Other studies have similar findings, demonstrating that good vocabulary learners regularly engage in morphological problem-solving. For example, Wysocki and Jenkins (1987) concluded that use of morphological analysis helps superior students to excel at learning new vocabulary. In a study which looked at individual differences in acquiring morphological knowledge, advanced fifth graders were better than average eighth graders at learning the meanings of derived words (Freyd & Baron, 1982).

Some research concludes that morphological awareness has an important role in vocabulary acquisition that is distinct from the role of phonological awareness. Using hierarchical regression analysis, McBride-Chang and colleagues (2005) found that morphological awareness predicted an additional 10% of variance in vocabulary knowledge above phonological variables. Kindergarten and Grade 2 students completed measures of speeded naming, phonological awareness, word identification, nonsense word repetition, and vocabulary knowledge along with two measures of morphological awareness. Two aspects of morphological awareness were examined: morphological identification --the ability to distinguish different meanings across homophones--and morphological structure awareness or the ability to create new meanings by making use of familiar morphemes.

Phonological skills are especially important for vocabulary learning (Bowey, 2001; Gathercole et al., 1992; 1999). Results from the McBride-Chang et al. study concur: 48% of variance in vocabulary knowledge was predicted by the phonological processing and reading variables. They postulate a bidirectional relationship between morphological awareness and vocabulary learning. Morphological awareness is a separate cognitive component of vocabulary knowledge and contributes unique variance to vocabulary knowledge with all other reading and language-related measures statistically controlled. The combined measures of morphological awareness were good predictors of vocabulary knowledge, even when phonological processing, word reading skill, & age were controlled. In sum, they conclude that morphological awareness is a separate construct from phonological processing and word reading skills and further, both morpheme identification and awareness of morphological structure are potentially unique features of vocabulary development.

2.2.2 Morphological Generalization Hypothesis

Wysocki and Jenkins (1987) propose a mechanism by which students learn the meanings of complex words by analyzing constituent morphemes, what they call the *morphological generalization hypothesis*. As defined by Wysocki and Jenkins, children engage in morphological generalization when they use their knowledge of a familiar word to help them infer the meaning of an unfamiliar, but related word. In their study, 131 children across Grades 4, 6, and 8 learned a set of words and were later tested on a matched set of transfer words with the same root. Researchers selected 12 word pairs (stimulus plus transfer words) based on the words' level of difficulty as measured by a pretest. The design required training students on one of two stimulus word sets and testing them on both sets of morphologically related words. Thus three groups were created: Set A, Set B, and no training (control). Stimulus word training was provided in three sessions of 15-20 minutes each, and delivered over a two-week period in whole- group format.

Three vocabulary assessments were created, one which measured the learning of the taught stimulus words, and two which measured transfer effects (performance on related words). First, stimulus words were tested in a "weak" sentence context. Next, transfer words were tested in a "weak" sentence context. And finally, transfer words were tested in a "strong" sentence context. Transfer word tests were given two weeks after the test of stimulus words. Measurement effects were a concern because the repeated testing may have led students to use a morphological strategy when they wouldn't have spontaneously done so. Two scoring systems were utilized: one lenient and one strict. Lenient scoring allowed for definitions which did not provide the correct grammatical form as long as the general meaning of the word was understood. The strict criterion required both meaning and syntax to be correct. For example, under the lenient scoring procedure, *wise* would be an acceptable definition for *sapience*. Under the strict criterion, *wisdom*, but not *wise*, would be acceptable.

First, an ANOVA was conducted to determine whether the students learned the taught words. Both experimental groups did better on taught words than on untaught words with no effects for grade level. To determine the effects of context, a separate ANOVA was performed. Both experimental groups performed better in a strong context, particularly the older students. A third ANOVA was conducted to examine the difference in performance between stimulus words and transfer words which could indicate whether morphological generalization occurred. Support for the morphological generalization hypothesis was weak when a strict scoring procedure was employed, but stronger when a lenient criterion was used. Furthermore, an age of acquisition pattern was demonstrated as older children were better at analyzing the meanings of complex words than younger children.

Freyd and Baron (1982) extended the research on morphological generalization by addressing the following questions in their study: first, do children differ in the mechanisms used in acquiring derived words? And, are advanced students more likely than average students to use derivational rules? Results suggest that in fact children do differ in how they learn the meanings of derived words. Findings reveal that those who use morphological generalization are faster learners of novel words. To examine individual differences in learning derivational rules, the study compared average middle school-age children with precocious upper-elementary students. Participants were 48 average-ability eighth graders and 32 high-ability fifth graders. Measures included a vocabulary test of derived and simple words and the Paired Associates Learning Task which utilized pseudowords. Target words for the vocabulary measure consisted of 60 real words (30 simple and 30 derived) which were selected from Carroll, Davies, and

Richman (1971) by taking one simple and one derived word from each thousand words in the rank list. This method of selection ensured a range of familiarity for the test words as well as a rough matching of derived and simple words on frequency of occurrence. To administer, the examiner read the word out loud and the student wrote a definition. Scoring procedures required definitions for the derived words to include the meaning of the affix(es).

The Paired Associates Learning Task was developed to measure the participants' ability to learn the "meanings" of ten pairs of pseudowords. Of the ten pairs of pseudowords, half were related by consistent derivational rules and the other half were not. In the case of unrelated pseudowords, different roots were used for the two members of the stimulus pair. For example, one derivationally related pair utilized is *flur/flurment*; an unrelated pair was *torb/meckable*. The design of the study provided that for each stimulus pseudoword, two responses (i.e., definitions) were created, Form A responses and Form B responses. Pairs of words associated with the same root in one list were associated with different roots in the other list. For example in Form A, the response for *flur* was *play*; in form B, it was *say*. Similarly, in Form A *flurment* "meant" *game*; in Form B *flurment* "meant" *word*. Since half of the participants in each grade-level group used each list, each stimulus pair occurred equally often in each condition within each group. The task was administered individually to each participant. Each stimulus pseudoword was printed on a 3 x 5 index card. The experimenter went through the cards, telling the student the "meaning" of each nonsense word. Each student was encouraged to go through the cards independently as well. For the testing procedure, the anticipation method was used. The student attempted to give the response to each stimulus, card by card, with the experimenter providing the response when the student wasn't able to correctly do so. The cards were presented in the same order on each trial, but the order was chosen at random for each participant by shuffling the cards at the start of the testing procedure. The procedure continued for 12 trials, 45 minutes, or two correct trials on all words, whichever came first.

Freyd and Baron hypothesized that knowledge of derived words would correlate with speed of vocabulary learning. Hence they expected the high-ability fifth graders would be particularly good at learning derived words. Results from the vocabulary test support this hypothesis. The fifth graders outperformed the eighth graders on both simple words ($p < .001$) and derived words ($p < .001$). An interaction between grade and word type indicated the fifth graders' superiority over the eighth graders was even greater on derived words. In further support of the hypothesis, findings from the Paired Associates Learning Task show the fifth graders were correct more often on the related pairs than on the unrelated pairs ($p < .025$). The average-ability eighth graders were not. To assess differences in overall rate of learning, the numbers of correct responses within the first four test trials of the paired associates task were counted for each student. The eighth graders averaged 25.8 correct responses; the fifth graders averaged 32.6 correct responses, a significant difference ($p < .05$). Next, Freyd and Baron computed the proportion of trials in which participants gave correct responses in each condition (related v. unrelated). The final score was the difference between the proportions for the two conditions; a positive score indicates the student may have used derivational relations while a zero or negative score indicates no such use. The eighth graders showed no evidence of use of derivational relations in the Paired Associates Learning Task. Their mean score was -.02, not significantly different from zero. In contrast, the fifth graders' mean was .11, a significantly positive number ($p < .001$). Freyd and Baron conclude that faster learners of vocabulary are better able to use derivational relations.

In a more recent effort to test the *morphological generalization hypothesis*, Baumann and colleagues carried out two intervention studies (2002, 2003) with fifth graders. In the first study, participants were assigned to one of four groups with differing vocabulary instruction conditions: (1) instruction on prefix families; (2) instruction on the use of context clue strategies; (3) a combination of the first two methods; and (4) a control group which received no explicit vocabulary instruction. The results were mixed with little evidence of transfer effects (Baumann et al., 2002).

Participants were 88 students from four fifth grade classes. Following twelve 50-minute lessons, students were tested on their ability to recall the meanings of words used to teach the morphemic and contextual analysis skills, to infer the meanings of uninstructed words which contained taught morphemic elements or words that were embedded in text that included taught context clues. They were also tested on their comprehension of text containing transfer words. Experimental measures assessed students' ability to learn 60 low-frequency target words: 30 prefixed and 30 context words selected from the *American Heritage Word Frequency Book* (Carroll, Davies, & Richman, 1971). Morphemic and context words were matched for frequency. Each set consisted of 10 lesson words and 20 transfer words. In addition, Baumann et al. collected descriptive data on students' vocabulary learning as gathered through individual interviews with students from each treatment group.

Two pretests provided an evaluation of treatment groups for possible pre-experimental differences in vocabulary ability. They also served as covariates in posttest analyses. First, a standardized, multiple-choice vocabulary test provided a general index of students' pre-experimental vocabulary knowledge. In addition, a 40-item, experimenter-constructed multiple choice test was used to assess students' pre-experimental knowledge of a subset of 14 lesson words and 26 transfer words. Seven posttests—five immediate and two delayed—assessed students' knowledge of words presented during instruction (lesson words), students' knowledge of words whose meanings could be inferred as a result of instruction (transfer words), and students' comprehension of passages containing transfer words.

Findings show a strong immediate and delayed effect of morphemic and contextual analysis instruction for lesson words. However, the effect of morphemic and contextual analysis instruction for transfer words was not significant and confined to the immediate measures only. No effect on comprehension was demonstrated. No differences were found among students with differing levels of pre-existing vocabulary knowledge; thus, students of all ability levels benefited equally from the intervention. One possible explanation for the absence of transfer effects may have been related to measurement sensitivity. Posttests contained both production and recognition tasks, but production measures were generally more sensitive to treatment effects than recognition measures. For example, the Morphemic Transfer Words Production Test had a mean effect size of .95. The delayed measures for transfer words were both recognition measures and therefore may have underestimated the delayed effects. Thus, the lack of transfer effects may be the consequence of a degrading of the skills over time or it may be the result of measurement factors. Despite the mixed results, Baumann et al. claim there is support for teaching middle- to upper-elementary students to employ morphological analysis to infer word meanings.

A subsequent study incorporated derivational suffixes along with prefix families (Baumann, Edwards, Boland, Olejnik, & Kame'enui , 2003). Participants were 157 students in eight fifth-grade classrooms. The vocabulary interventions were integrated into the fifth grade social studies curriculum. Half the group received textbook vocabulary instruction and the other half received instruction in morphological and contextual analysis. The latter had both strategy instruction and instruction in anchor, or representative, vocabulary. Prefixes and suffixes were taught directly as well as various context clues. The textbook vocabulary group received instruction in specific, content-central textbook words, but not in any word-learning strategies. Twenty-five 15 minute vocabulary lessons were provided in a whole-class format. In contrast to the earlier intervention study, the experimental group performed significantly better than the control group on defining novel words, providing stronger support for a morphological generalization hypothesis.

2.2.3 Reading Comprehension and the Semantic Properties of Morphemes

More recent studies have focused on identifying the mechanism(s) by which morphological awareness contributes to reading comprehension. Given the existing evidence of a role for morphology in word recognition, vocabulary learning, and other prerequisites for reading, it is quite possible that it is via one or more of these routes that morphological awareness contributes to skilled text comprehension. One line of research advances the notion that the impact of morphological awareness on reading comprehension is mediated by vocabulary knowledge. Semantic information in morphemes helps readers access meanings of complex words. Thus, a reader who recognizes the *frustrate* in *frustration* will likely be able to access the meaning more accurately and quickly (Tyler & Nagy, 1990). For example, in one recent study, McCutchen, Green, and Abbott (2008) conclude that the contribution of morphological skill to reading comprehension is largely achieved through vocabulary. As evidence, they advance findings which confirm a large portion of shared variance exists between vocabulary and morphological awareness.

Carlisle (2000) attempted to clarify the relationship between morphological awareness, vocabulary knowledge, word-reading skill, and reading comprehension in a study with children in Grades 3 and 5. The results showed that awareness of morphological structure and meaning contributed significantly to comprehension at the word and text levels for both grade groups. Further, findings reveal a developmental pattern in which certain aspects of morphological awareness become increasingly important in the developing literacy of elementary school children.

Three morphological awareness tasks, each measuring different aspects of awareness of morphological structure, were administered. In order to clarify the relationships between different aspects of morphological awareness, the study included tasks of both structural analysis and definition. One task, the Test of Morphological Structure (TMS), required production of complex words and extraction of base forms to fit sentence contexts. Thus, both derivation and decomposition processes were tested. The two tasks contained equal numbers of phonologically transparent and phonologically opaque (shift) words. Suffixes judged to be familiar to third and fifth graders were used and each suffix was equally represented in the two tasks. Carlisle hypothesized that the derivation task would be more closely related to the children's ability to define morphologically complex words than the decomposition task because producing derived forms, like defining derived forms, requires knowledge of the grammatical roles and meanings of suffixes. This task was chosen because it may necessitate the same competencies students need when they encounter morphologically complex words in reading.

A second task assessed children's use of morphological problem-solving to determine the meanings of inflected, derived, and compound words. This measure, referred to as the Definition Task, was adapted from The Test of Absolute Vocabulary Knowledge (TAVK; Anglin, 1993).

The Definition Task entails interviewing the child about a list of words: asking the child to give the meaning of each word, to use it in a sentence, and (if needed) to pick the meaning from a multiple-choice set. To earn credit for an item, the child must show an understanding of both the general meaning of the word and its grammatical role in a sentence. In some cases, the Definition Task can indicate the use of morphological problem-solving. For example, during the interview process, some students will work out the meaning of a complex word they previously may not have been able to define.

A third task, The Word Reading Task (WRT), measured speed and accuracy in reading of morphologically complex words and included 25 high frequency words, of which 9 were phonologically and orthographically transparent and 16 were shift words. A second set of 20 words had low surface frequencies but high base word frequencies. All of these had transparent relations (e.g., *puzzlement*). The Vocabulary and Reading Comprehension subtests of the Comprehensive Testing Program III (Educational Records Bureau) were used as outcome measures. The Vocabulary subtest asks the child to read a word and select the best meaning from a set of choices. The Reading Comprehension subtest entails reading passages and selecting a best answer to comprehension questions.

Two regression analyses were carried out at each grade level, one for each of the two outcome measures. The three morphological awareness tasks accounted for large portions of the variance in vocabulary and reading comprehension. Closer examination of the data revealed a developmental pattern. For the third graders, the three tasks combined accounted for 43% of the variance in reading comprehension; however, of the three morphology tasks, only the word-reading task made a significant independent contribution. For the fifth graders, the three morphology tasks accounted for 55% of the variance in reading comprehension. Of the three tasks, only the Test of Morphological Structure (TMS) made a significant independent contribution. These results suggest that developmental changes occur in the process of literacy acquisition which alters how different aspects of morphological awareness relate to reading comprehension. For the third graders, the ability to read derived words was the major influence. In contrast, for the fifth graders, awareness of the structure, meaning, and grammatical roles of words played a more important role than the ability to decode derived words.

Another important finding was a significant relationship between the TMS and the Definition Task (.64 for fifth grade and .46 for third grade). A regression analysis showed that, at both grade levels, an awareness of morphological structure contributed most significantly to the students' definitions of complex forms on the Definition Task. These results support Carlisle's hypothesis that the ability to produce derived words from given bases is closely related to the ability to infer the meanings of complex words through morphological analysis. While these findings add to the literature, the results of this study are somewhat limited by small sample size and near-ceiling performance by some fifth graders.

Extending earlier research on the links between morphological awareness, vocabulary, and reading comprehension, Carlisle and Flemming (2003) focused on the semantic aspects of morphemes and their effect on comprehension in a longitudinal study. Findings demonstrated that first and third graders' analysis of complex words was predictive of vocabulary and reading comprehension two years later as measured by standardized tests. Moreover, their results suggest that third graders, more than first graders, are capable of morphological processing that is meaning driven. While breadth and depth of word knowledge clearly facilitates reading comprehension, Carlisle and Flemming suggest that morphological problem-solving itself makes a direct and important contribution to comprehension.

Two experimental measures of morphological processing were administered at the start of the study (Year 1). The Word Analysis Test (WAT) is adapted from Rubin's (1988) task in which children are asked whether a larger word contains a smaller word that means something like the larger word. Test items included both inflections and derivations. Each of the complex words was matched to a mono-morphemic word with the same final sound (e.g., *hilly /silly*). Credit was given only when the child responded correctly to both items in a pair. The WAT was used as a way to determine whether the children could distinguish morphemic segments on the basis of the meanings of familiar words. An adaption of The Test of Absolute Vocabulary Knowledge (TAVK; Anglin, 1993;) was used to measure the nature and degree of children's knowledge of word definitions. Carlisle and Flemming posit that this task, what they call the Definition Task, assesses morphological processing, not merely vocabulary knowledge, because it taps into both the semantic and syntactic aspects of morphological awareness. Additionally, one standardized measure, The Peabody Picture Vocabulary Test-Revised (Dunn & Dunn, 1981), was also administered at this time to assess receptive vocabulary skills.

Results of the WAT showed the third graders were better than the first graders at determining whether a word could be decomposed morphologically (e.g., *pinned* or *wind*). First graders responded correctly to 57% of word pairs; third graders were correct on 71% of pairs. Performance on the Definition Task also revealed a developmental pattern. Third graders were significantly better than first graders in defining morphologically complex words and using them in sentences. Three words from the Definition Task were selected for further analysis. These words--*knotless, stillness, and treelet*--each had high frequency bases but a low surface frequency. Third graders were more likely to refer to the base word in their answers, even if their definitions of the word presented revealed a lack of awareness of the meaning or grammatical role of the suffix. Few first graders recognized the familiar base word in these items. Carlisle and Flemming argue that this infers the younger children are less-skilled "morphological problem solvers."

In Year 3, the Reading Vocabulary and Reading Comprehension subtests of the Gates-MacGinitie Reading Test (MacGinitie & MacGinitie, 1989) were administered. The Reading Vocabulary subtest assesses the ability to read and select the best meaning of a word from multiple options; the Reading Comprehension subtest assess the ability to read short passages and then select the best answers to multiple-choice questions about the passage. Two standard regression analyses were carried out for each outcome measure—one for the younger group and one for the older group. Results indicate that for the younger group, only 11% of the variance in vocabulary knowledge was predicted by the two experimental measures of morphological awareness ($p = .17$). However, these measures accounted for a significant 23% of the variance in comprehension ($p < .05$). For the older group, 41% of the variance in vocabulary ($p < .01$) and 27% of the variance in reading comprehension ($p < .05$) was accounted for by the two measures. In each case, the Definition Task, not the WAT, accounted for a significant portion of the variance. Of the two morphological awareness tasks, only performance on the Definition Task significantly predicted the children's performance on the reading measures two years later. Based on these findings, Carlisle and Flemming argue it is the semantic features of morphemes which have the greatest impact on reading comprehension. The researchers propose that it is morphological problem-solving, and not vocabulary knowledge per se, that is an important part of literacy development in the elementary years. Without the capability to infer the meanings of complex words through morphological analysis, children's ability to comprehend text is limited.

The primacy of vocabulary as the mechanism by which morphological awareness contributes to comprehension also gains support, albeit in an indirect way, in one intervention study. Arnbak and Elbro (2000) found Danish children in Grades 4 and 5 with dyslexia improved their reading performance when provided with morphological awareness training which focused on the semantic features of morphemes. At the conclusion of the training period, the experimental group showed gains in reading comprehension and spelling of morphologically complex words. Because training focused on the semantic aspects of morphology, Arnbak and Elbro theorize that knowledge of morphological structure helps students with poor phonological skills to develop meaning-oriented decoding strategies which in turn enhanced comprehension.

Participants were recruited from nine schools and were selected based on a history of severe reading and writing problems. Each student was identified as reading at least 2 years below expected reading level in spite of otherwise normal cognitive development. An experimental group (n = 33) received morphological awareness training for 15 minutes, three times per week for a total of 12 weeks. A matched control group (n = 27) received traditional remediation in phonological awareness, grapheme to phoneme recoding, spelling to dictation, and oral and silent reading tasks. The study design provided for instruction to be delivered in small groups. Compounds, derivational affixes, and inflections of nouns, verbs, and adjectives were covered. Since the focus of the morphological awareness training was on the meaning of morphemes, the words used in the program were ordered according to semantic transparency. First, the students were made familiar with semantically transparent morphemes in words; then they were presented with more semantically opaque morphemic units.

The intervention had a small, positive effect on the morphological awareness of the experimental group based on pre- to post-test gains on three morphological awareness tasks.

These tasks consisted of segmenting compounds into morphemes, making morphological analogies, and compounding, deriving and inflecting pseudowords. In order to control for any variation in the students' abilities prior to the intervention, analyses of covariance were performed with the pre-test measure as a covariant. Individual gains in morphological awareness were calculated following the log odds method (Elbro, 1990). A correlation was found between group size and gain in morphological awareness. Although this pattern was seen across all three measures, only on the pseudoword task did the correlations reach significance. These findings suggest that the smaller groups of students benefited the most from the intervention.

Passage comprehension was measured via a standardized silent reading test with picture selection (Nielsen & Petersen, 1992). Scores were the number correct in 5 minutes and the number correct per minute within a 15 minute period. The experimental group had a higher gain in passage comprehension than the control group did. The gain was significant in an analysis by schools as well as by subjects (both $p < .05$). Arnbak and Elbro interpret these results to support their hypothesis that knowledge of the morphological structure of words boosts meaning-oriented strategies in reading.

2.2.4 Summary

Knowledge of morphemes and the ability to analyze them is clearly an important aspect of literacy development in school-age children and adolescents. Studies have consistently shown the ability to use morphological knowledge contributes to a host of reading and reading-related outcomes. Significant interactions between morphological awareness and a range of literacy skills such as decoding, spelling, vocabulary learning and comprehension have been found across all age groups and in various languages. Based on these findings, classroom instruction in morphological analysis is likely to be valuable. Continued progress in understanding the development of morphological awareness and its instructional implications depends upon future empirical research.

In a recent article reviewing the research base in vocabulary instruction for students with learning disabilities, Ebbers and Denton (2008) recommend that while teaching vocabulary through a morphological approach makes sense, it should be regarded as a promising practice until more research validates its effectiveness. The same should be said of utilizing morphological awareness instruction to improve any literacy outcome for struggling readers. A contemporary review of morphological intervention research on reading outcomes in grades K-12 cautions that only a limited number of studies have been conducted and calls attention to a variety of methodological concerns (Reed, 2008). In particular, there is a need for intervention studies which target the population of students with learning disabilities. Given the paucity of intervention studies in this area, the full effects of direct, explicit instruction in morphological awareness and morphological problem-solving remain unknown and deserving of attention.

2.3 DYNAMIC ASSESSMENT

Data from traditional assessment is expected to provide an accurate and objective examination of one's performance or ability because its procedures are standardized and because it is administered by a neutral observer. These assessments measure the outcome or end result of learning, the product of accumulated knowledge or skill. For this reason, traditional assessment is sometimes referred to as static assessment. Unlike traditional static measures, dynamic assessment (DA) is administered by an examiner who provides scaffolding and intervention when a student fails. Unlike traditional static measures, dynamic assessment procedures attempt to capture the process as well as the product of learning. Thus, DA measures two performances: what a child can do independently and what he can do with academic assistance from a capable adult. DA is expected to offer a measure of potential achievement. That is, if students can be taught to "coach" themselves to solve problems through the use of strategies, this can be revealed by the amount of scaffolding they require during the assessment. By indicating what level of support is needed for a student to successfully complete a task, DA may be a more accurate reflection of how well students will achieve during classroom instruction with varying degrees of instructional intensity. For this reason, dynamic measures may be useful in a Response to Intervention (RTI) framework.

DA is based on the theoretical model of the *Zone of Proximal Development* (ZPD) proposed by Vygotsky (1978). Vygotsky posits that children's knowledge develops within social interactions with more able adults. When children interact with a more knowledgeable partner, emergency development processes become evident. Thus, the ZPD is the range of learning children can achieve when provided with support from an experienced adult. As such, DA can help teachers determine when and how to intervene. For example, students who require extensive assistance during the assessment are likely to require intensive intervention in order to succeed. In addition, DA can reveal the strategies that underlie a student's performance and allow us to make inferences about student thought processes. Its primary goal is the evaluation of the processes of thinking, learning, and problem-solving: "the focus of dynamic assessment is on the assessor's ability to discover the means of facilitating the learning of the child" (Lidz, 1991, p. 9). Greater insight into these processes can provide relevant information for intervention.

Feuerstein (1979) argues that dynamic assessment can provide the best reflection of learning ability or potential for change by (1) assessing the type of intervention that causes change in a skill, (2) measuring the extent of the change that occurs, and (3) determining whether the skill learned can be transferred to other areas. There are two primary assessment methods that use the dynamic assessment model: test–teach–test and graduated prompting. Both methods focus on measuring the process of learning rather than measuring static knowledge. An additional aspect of both dynamic assessment models is that they focus on the conditions under which performance can be changed instead of only on the performance itself. In each method, the assessment focuses on the interaction of the assessor and the examinee, the influence of this interaction on performance, and the responsiveness of the examinee to the interaction (Lidz, 1991). In both models, *learning ease* is used as an index of potential. There are two ways in which DA quantifies learning ease. In a test–teach-test design, gain scores are generally examined. In models which use graduated prompts, learning ease is measured by the amount of scaffolding needed for mastery.

In comparing different methods of dynamic procedures, a fundamental distinction can be made between DA which is clinical in focus and DA which is assessment-oriented (Caffrey, Fuchs, & Fuchs, 2008). Clinical DA not only gives performance-contingent feedback but offers instruction in response to student failure with the purpose of enhancing the student's achievement. Clinically-oriented DA is intended as an educational treatment and often uses a problem-solving format with non-standardized procedures. Assessment-oriented DA, in contrast, uses a standardized protocol during which the examiner provides feedback in a single session. This form of DA is utilized to identify students who may require more intensive intervention.

The test-teach-test model has its history in Budoff’s learning potential tests. These relied on standardized coaching techniques and were reported to be an excellent predictor of future performance (Budoff, 1987). The test–teach–retest method has been used to evaluate learning of vocabulary (Peña et al., 1992, 2001) as well as learning of synonyms (Gutierrez-Clellen et al., 1998). Similar dynamic assessment tools have also been developed to evaluate categorizing skills (Ukrainetz, Harpell, Walsh, & Coyle, 2000) and narrative abilities (Miller, Gillam, & Peña, 2001). Swanson & Lussier (2001) conducted a meta-analysis of 30 DA studies focusing particularly on training forms of DA which embed intervention within the assessment tool. These measures are particularly concerned with the modifiability of children’s learning (Lidz, 1991).

The graduated prompts method, sometimes referred to as *successive cueing* or *testing through learning and transfer*, was developed by Campione & Brown (1987). A system of scaffolding is created by this series of graduated prompts as students are given progressively more explicit “hints” until they can solve a problem independently. The cues are standardized and administered in a pre-determined order. Learning ease is quantified as the number of cues necessary for success. Students who require the fewest number of hints are presumed to have the greatest learning ease (Campione, Brown, & Bryant, 1985). Ferrara, Brown, and Campione, (1986) studied the cognitive differences between students of high ability and students of low ability. Results show that students with lower overall ability require more cues to solve a problem and transfer information than do students with high ability. Finding significant relations between the number of prompts required for problem solution and IQ scores indicates some degree of concurrent validity. Dynamic assessment models using graduated prompts have also been informative for predicting language change (Bain & Olswang, 1995; Olswang & Bain, 1996).

Proponents of DA as a measurement instrument, as opposed to an instructional tool, believe standardization and technical adequacy are needed (Swanson, 1994; Bryant, Brown, & Campione, 1983). Standardization of protocols, reliability measurement, and fidelity of the testing procedures are important concerns. Ultimately, dynamic procedures must be both reliable and efficient. One frequently cited disadvantage of dynamic measures is that they are very time-intensive to administer and difficult to use with fidelity (Swanson, 1996). Most agree that DA will not replace traditional assessment, but may be used in conjunction with it (Lidz, 1987). The graduated prompts method is thought to be efficient, standardized, and have the potential to be generalizable (Caffrey, Fuchs, & Fuchs, 2008). Further, scripted prompts encourage consistent technique and therefore better reliability. If DA is to become a useful method, fidelity, reliability, and validity need to be carefully considered.

The most important question that must be addressed when selecting any assessment tool is: what unique information can this measure provide? DA may be a better discriminator of individual differences than static assessment. That is, low-achieving students can be differentiated by how easily they learn new skills. For this reason, DA is sometimes touted as an alternative to language assessments that rely on standardized static measures which may overdiagnose language disability in children from nonmainstream backgrounds (Gutierrez-Clellen et al., 1998; Washington & Craig, 1992, 1999). Dynamic measures are likely to yield information that is meaningful and relates to classroom learning and instructional planning. For example, DA can be used to identify the type and intensity of sustained intervention that is required for success as well as the necessary level of instructional explicitness to reach mastery.

For students with disabilities, specific obstacles and cognitive deficits responsible for student failure may be identified.

2.3.1 Assessing Vocabulary with Dynamic Measures

Dynamic assessment has shown promise in addressing a number of the concerns raised with standardized static measures of vocabulary. For example, a growing number of studies have demonstrated that dynamic assessment can aid in distinguishing children from nonmainstream populations who are typically developing from those who have language learning difficulties (Peña, Iglesias, &Lidz, 2001; Gutierrez-Clellen et al., 1998). Peña (1992) used a dynamic assessment procedure to differentiate Spanish-speaking preschoolers with language disabilities from non-disabled Spanish-speaking preschool students. She concluded a static measure would over identify for special education.

In contrast, Restrepo (1998) did not find that dynamic assessment using a novel vocabulary-learning task and a novel bound-morpheme generalization task was informative for differentiating between children who were typically developing and those with language disabilities. However, one very plausible explanation Restrepo offered is that the task did not stress the language learning mechanism enough resulting in a ceiling effect for both groups of children. The mediated learning experience in the Restrepo study provided a very structured procedure of model, imitate, identify, and name. This structured learning and elicitation method has been described as an explicit mechanism for teaching children with language learning difficulties by providing them with the extra support needed. It is possible that a more difficult novel vocabulary-learning task may better aid in the differentiation of children with language learning disabilities from those who are typically developing (see Rice et al., 1990; Roseberry & Connell, 1991, for discussion).

Often, it is difficult to tell the difference between children who lack underlying learning skills and those who may lack general word experience which limits their vocabulary knowledge. Previous research provides indirect evidence that differences in the static test performance of the two groups may be linked to past language exposure and/or experience. Studies have shown that scores on standardized static vocabulary assessments correlate with variables other than semantic knowledge (Gray et al., 1999). A study conducted by Peña, Iglesias, & Lidz (2001) examined the differential performance of two groups of students—high-risk and low-risk--on standardized static vocabulary measures and on a dynamic measure. The results from this study argue that while the Peabody Picture Vocabulary Test—Third Edition (PPVT-III; Dunn & Dunn, 1997) may adequately measure a current and particular vocabulary, it may not provide a measure of word-learning ability for all populations. The authors suggest it is possible that the high-risk participants hear and therefore know fewer words than the low-risk participants. It is also possible that the specific items on PPVT-III are more or less relevant in the lives of children from low-risk versus high-risk backgrounds. Although this study confirms that scores on early static vocabulary measures are not fully informative as the sole measure of word-learning ability in high-risk populations, performance on such measures can function as predictors of later reading skills and academic success (Snow et al., 1998). Thus, although the results of this study suggest that assessment to evaluate the language-learning mechanism is important, actual vocabulary knowledge is also relevant and informative. It appears that the use of dynamic assessment in conjunction with traditional static measures is likely to facilitate more accurate diagnosis of word-learning ability.

2.3.2 Dynamic Assessment as a Measure of Response to Intervention

A primary criticism of the use of static assessments is that they over-identify African American students for special education and possibly under-identify students from bilingual backgrounds (Reed, 2000; U.S. Department of Education, 2000, 2001). In addition, recent research from the National Institutes of Child Health and Human Development (Lyon, 2001) calls attention to the numbers of students who are identified with a specific learning disability but who actually may have not been given appropriate reading instruction. In response to these issues, more appropriate ways of identifying students with a specific learning disability continue to be examined.

These criticisms have also led to changes in the eligibility criteria for specific learning disability (SLD) in the reauthorization of the Individuals with Disabilities Education Improvement Act (IDEA) of 2004. These changes prevent local education agencies from requiring the use of a discrepancy formula for determination of SLD. Additionally, IDEA 2004 permits schools to consider whether the student has responded to scientific intervention as part of the determination process (Moore-Brown & Montgomery, 2005).

Response to Intervention (RTI), a primary approach to assessment in use today, purports to distinguish between underachievement from poor instruction versus low achievement which is the result of a disability (Fuchs, Fuchs, & Compton, 2004). Similarly, DA is an instrument tool which shows change in performance in response to intervention. As a mechanism for identifying non-responders, DA may provide useful information.

A major purpose of educational assessment is to predict future achievement and identify students early who are at risk for poor academic outcomes so that intervention can begin before their problems become intractable (Fuchs et al., 2008). Research indicates that DA might have particularly important implications for discriminating students at the lowest end of the distribution. For example, two children may obtain the same score on an assessment while one may have a greater potential to learn. One may succeed with only minimal assistance; another may require extensive scaffolding (Fuchs et al., 2008).

Caffrey, Fuchs and Fuchs (2008) conducted a mixed-methods review of 24 studies to explore the predictive validity of DA. In a five-phase study, the researchers: (1) compared the predictive validity of DA to traditional tests, (2) compared two forms of DA, (3) examined predictive validity by student population, (4) investigated various outcome measures; and (5) assessed the value added of DA over traditional testing.

The predictive validity of DA was highest when the instruments were standardized and feedback was not contingent on the student's response. It was higher for students with disabilities than for at risk or typically-achieving students. DA correlated highest with independent DA tests and criterion-referenced tests. When norm-referenced tests and teacher judgment were used, the predictive validity of DA was weaker. Within a subset of 10 studies which conducted multiple regression analysis, two studies (Bryne et al., 2000; Meijer, 1993) investigated the unique contribution of DA after traditional achievement tests had already been entered in the equation. The other eight studies examined the unique contribution of DA after traditional cognitive tests (i.e., IQ tests) had been entered in the equation. These researchers conclude there is evidence that DA can predict unique achievement not indexed by traditional static assessments. When dynamic measures follow entry of traditional scores in forced-entry multiple regressions, they explain significant variance in predicting future performance in reasoning, reading achievement, and math achievement.

The child's responses to the dynamic testing procedures can be used to make predictions about the child's response to intervention. Spector (1992) administered a DA of phonemic awareness in the fall of kindergarten and then assessed word-level reading skill in the spring. DA explained an additional 21% of the variance in end of kindergarten word reading beyond static phonological awareness measures. In this study, the number of prompts required to reach criterion was utilized as the predictor.

In a more recent study, Fuchs et al. (2008) examine the predictive value of DA to measure third grade students' potential to learn algebra. Introducing novel academic content, which third graders would have been extremely unlikely to encounter, increases the probability that students' performance on the dynamic measure reflects their potential to learn new content rather than reflecting existing abilities. Fuchs et al. suggest that novel academic content may be the most appropriate approach to take in designing DA for other domains. They hypothesize that future learning is better indexed by DA than with static assessments which are designed to index what has already been learned.

Within the RTI framework, a student's responsiveness to secondary interventions is considered the mechanism by which teachers can differentiate between students whose skills are low due to insufficient instruction and those which are low due to disability. Secondary preventative intervention requires at least 10 weeks and perhaps longer to identify a student with a learning disability. Fuchs et al. propose that DA may be useful in predicting responsiveness to intervention much sooner. They base this on findings which indicate DA is better at forecasting distal, or far off, aspects of learning which are not relevant to the student's proximal instructional needs. It may be possible that DA could be used productively to identify early-predictors of late-emerging reading problems (Fuchs et al., 2008).

2.3.3 Summary

Observation of children learning a new skill or strategy focuses on not only what is learned but also how learning occurs. In this way DA can measure a unique dimension of student abilities. As such, DA has the potential to be a valuable tool for a variety of assessment purposes, including identifying students in need of intensive intervention, monitoring progress, and predicting later achievement. Because DA allows for systematic observation of areas affecting student learning that typically are not considered with standardized assessment measures, the dynamic assessment process can uniquely contribute to the overall understanding of a student and his or her learning challenges. Given the difficulty of assessing vocabulary knowledge linked as it is to culture, socioeconomic status and language background, a dynamic measure of word learning that could distinguish between typical and atypical word learners would have considerable utility.

3.0 METHODS

Morphological analysis is a word learning strategy with promising applications. Given the importance of complex vocabulary at the secondary level, further understanding of how to teach word learning strategies to students with reading problems is needed. In summary, the objectives of this study are to evaluate the value added of a dynamic measure of morphological analysis to differentiate among students with similar reading levels. It is hoped this will yield information not attainable through traditional assessment alone. Such an outcome would provide further support for the usefulness and validity of dynamic assessment of morphological knowledge.

3.1 PARTICIPANTS AND SETTING

Twenty-seven grade six students participated in this study. The school district from where this sample was drawn is located in a primarily working-class community of approximately 12,000 residents located in the southwestern region of Pennsylvania. The school district enrolls 3,500 students in grades K-12 across nine schools. The student body across these schools is approximately 84% white, 14% African American, and 2% Asian, Hispanic, or other race/ethnicity. Approximately 49% of students in the district receive free or reduced-price lunch.

Informed consent was obtained for students enrolled in the sixth grade (see Appendix A for a copy of the consent letter). Students at this grade level were selected based on empirical evidence which suggests it is a period of rapid developmental growth in morphemic analysis skill (Carlisle, 2000; Larsen & Nippold, 2007; Nagy et al., 2006). Students with sensory or neurological deficits were excluded from participation as were children whose native language is not English. The resulting sample was comprised of 17 boys and 10 girls. The mean age of participants was161.28 months (*SD* = .393) or 12 years 9 months.

3.2 PROCEDURES

Each participant individually completed a series of tasks. These tasks include the Peabody Picture Vocabulary Test-4 (Dunn & Dunn, 2007), a curriculum-based measure of reading fluency (CBM), and two measures of morphological knowledge: the Dynamic Assessment of Morphological Analysis (DATMA) and The Test of Morphological Structure (TMS). Testing was conducted in an unused classroom at the school site and was split into two sessions, each of about 30-40 minutes duration. In the first session, participants completed the PPVT-4, the three sixth grade oral reading fluency passages, and an abbreviated five-item version of the DATMA. Order effects were controlled by randomly varying the sequence in which the measures were administered. In the second testing session, the TMS and the full DATMA was administered. The static measure, the TMS, was given before the DATMA to avoid any confound caused by potential learning effects from the dynamic assessment instrument. Testing was conducted by the investigator and two research associates from the University. Each testing session was recorded on audiotape. In addition, scaled scores from the sixth grade reading subtest of *Pennsylvania State System Assessment* (PSSA) were obtained for each participant.

3.3 DESCRIPTION OF MEASURES

In all, data gathered from six different assessments was utilized: three standardized measures, a curriculum-based measure, and two instruments which measure morphological knowledge.

3.3.1 Standardized measures

Peabody Picture Vocabulary Test—Fourth Edition (PPVT-4; Dunn & Dunn, 2007) is a standardized, untimed, individually administered receptive vocabulary test. The test format is an easel with stimulus items consisting of four full-color pictures as response options on a page. For each item, the examiner says a word, and the examinee responds by selecting the picture that best illustrates that word's meaning. The age-norm and grade-norm samples were designed to resemble the English-proficient population from ages 2:6 to 90+, and closely match 2004 Census data for demographic variables. The median internal consistency (split-half) reliability coefficient of the PPVT-4 is reported to be .94. Test-retest reliability is reported to be .93 (Dunn & Dunn, 2007). The PPVT–4 instrument was correlated with the EVT–2, CASL, CELF-4, GRADE, and PPVT-III instruments to demonstrate the extent to which the observed pattern of correlations agrees with the pattern expected of a valid vocabulary measure. The PPVT-4 yields standard scores based on a mean of 100 and a standard deviation of 15. This assessment was chosen as a measure of general vocabulary knowledge which is likely to be related to derivational morphology skills.

Pennsylvania State Standards Assessment (PSSA) for reading consists of both narrative and expository passages, followed by a series of multiple choice comprehension questions. Both recall of factual information and the ability to make inferences on the passages are tested. The PSSA reading assessment employs two types of test items: multiple choice and performance tasks. They are designed to measure students' comprehension of the information contained in the reading passages. Multiple choice or "selected response" items present four answer choices from which the student selects the correct response. The performance tasks are open-ended questions or "constructed response" items which require students to construct an answer which is scored on a scale of 0 – 3 based on both content and completeness. These items focus on the meaning of the texts and require students to "summarize, identify, explain, and analyze" the information in the reading passages (PSSA Assessment Handbook, 2009). According to the PSSA Assessment Blueprint (2009), the PSSA reading test for sixth grade contains questions across two categories: (a) Comprehension and Reading Skills and (b) Interpretation and Analysis of Fictional and Nonfictional Text, with approximately 50% - 70% of the items falling within the first category and approximately 30% - 50% of the items falling within the latter.

Pennsylvania's General Performance Level Descriptors are Advanced, Proficient, Basic, and Below Basic. The Advanced level reflects superior academic performance, indicating an in-depth understanding and an exemplary display of skills. The Proficient Level reflects satisfactory academic performance, indicating a solid understanding and adequate display of skills. The Basic Level reflects marginal academic performance, indicating a partial understanding and limited display of skills. The Below Basic Level reflects inadequate academic performance, indicating little understanding and minimal display of skills. This assessment yields scaled scores. For sixth grade students, the Pennsylvania Department of Education defines a score of 1456 and above as "Advanced," a score of 1278-1455 as "Proficient," a score of 1121-1277 as "Basic" and 1120 and below as "Below Basic." All skills assessed are included in the Pennsylvania Academic Content Standards. Both *Basic* and *Below Basic* levels indicate a need for additional instructional opportunities. Students proficient and above are considered to be making adequate progress. PSSA reading subtest scores were obtained to provide information about participants' reading proficiency.

3.3.2 Curriculum-based measure (CBM)

Oral Reading Fluency Passages – A set of three standardized sixth grade passages representative of the curriculum being taught to the students were administered during the first testing session. The participant was provided standard directions and asked to read from each passage orally. The participant is given 1 minute to read each passage. The number of words read correctly is counted and recorded. The median score from the three passages wase utilized in the data analysis. This task was chosen to provide further information about participants' reading fluency, an aspect of reading proficiency shown to be critically related to reading comprehension (Fuchs, Fuchs, Hosp, & Jenkins, 2001). Administration of this measure was consistent with the procedure generally used in most clinical and experimental investigations of reading fluency (Shinn, 1989). Copies of the reading passages that will be used are included as Appendix C.

3.3.3 Measures of Morphological Awareness

The Test of Morphological Structure (TMS), developed by Carlisle (2000), is made up of two tasks and requires both production of complex words as well as extraction of base forms to fit sentence contexts. The first task, the Production task, requires production of a derived word in order to finish a sentence. The other task, the Decomposition task, requires the decomposition of derived words in order to finish sentences. The two tasks contain equal numbers of word relations that are transparent (i.e., the sound of the base form is intact in the derived form, as in *reason* and *reasonable*) and shift words (i.e., the phonological representation shifts from base to derived form, as in *produce* and *production*). Suffixes judged to be familiar to upper elementary-aged students are used; these include -th (e.g., growth), -ance/ence (e.g., performance), -er (e.g., teacher), -ity (e.g., equality), -tion/sion (e.g., description), -ous (e.g., famous), -able (e.g., variable). Each suffix was equally represented on the two tasks. The author of the TMS reports that the base and derived forms were equivalent in word frequency on the two tasks: for Decomposition, the mean standard frequency index (SFI) for the base forms was 56.7 (5.0 SD) and for the derived forms was 50.2 (5.9 SD), while for Production, the mean SFI for the base forms was 55.8 (6.2 SD) and for the derived forms was 51.0 (5.5 SD).

Carlisle hypothesizes that the derivation task, in particular, is closely related to children's ability to define morphologically complex words because producing derived forms, like defining derived forms, requires knowledge of the grammatical roles and meanings of suffixes. This assessment instrument has been chosen because it may demand the same problem-solving competencies students need when they encounter unfamiliar morphologically complex words. Appendix D contains a copy of this assessment.

The Production task was administered before the Decomposition task. The examiner read the directions to the participants, followed by 2 practice items, for which participants were provided with corrective feedback. The examiner presented all the items orally, and the participants did not see the written items.

Dynamic Assessment Task of Morphological Analysis (DATMA; Larsen & Nippold, 2007) is designed to measure differing levels of morphological ability. Participants are presented with a series of derived words and asked to define each one; a series of cues are given until either all the cues are exhausted or the child defines the word correctly. In its original form, this assessment instrument originally consisted of 15 test items plus two practice items. One potential limitation of the DATMA is its exclusion of compound words among test items. In order to avoid this limitation, additional items were added by the investigator for use in the current study. A complete explanation of the changes that were made is provided in the description of stimulus words below.

3.3.3.1 Task design

The DATMA utilizes the graduated prompts format. The cueing structure was based on previous research suggesting that when an individual encounters an unfamiliar morphologically complex word a certain process occurs (Anglin, 1993; Carlisle & Nomanbhoy, 1993; Rastle, Davis, Marslen-Wilson, & Tyler, 2000). First, the constituent morphemes are identified and their meanings accessed in the mental lexicon. Next, these meanings are combined to reveal the meaning of the morphologically complex word. Unlike many other measures of morphological awareness, this instrument specifically targets the ability to *use* morphological knowledge to determine word meanings. After each cue, the examiner pauses for up to 10 seconds; if no response is given or the child responds incorrectly, the next clue is provided. Once the child correctly defines the target word, cueing for that word stops and the next word is presented. The following script will be used when administering this task:

Examiner: I am going to say some words, and your job will be to tell me what you think they mean. I'll also show you a written copy of each word. If some of the words are hard, I'll give you some help. Are you ready?

1. Tell me what the word *beastly* means. (pause 10 seconds)

2. If child answers correctly, the examiner will say: "How did you know that?" Then proceed to step 3 (unless the child has referred to the individual morphemes). If the child responds incorrectly, the examiner will proceed directly to step 3.

3. Does the word *beastly* have any smaller parts? What are those parts? (pause 10 seconds; if child does not respond or is incorrect, proceed to step 4. If child is correct, the examiner will ask: Now can you tell what the word means?)

4. The smaller parts in this word are *beast* and *ly*. Now can you tell what the word means?

5. Listen to this sentence and then tell me what *beastly* means: "Jan tried to scare her brother by dressing up and acting *beastly*."

6. Which of these gives the meaning of the word *beastly*? (examiner will present three choices)

If the child provides an acceptable definition of the target word in step 1, he or she is questioned in order to reveal an awareness of the constituent morphemes and how they contribute to the target word's meaning. The cues mirror the process discussed above: first, identify the constituent morphemes and then use that information to determine the meaning of the derived form. The fifth cue gives a sentence context to help determine word meaning and the final clue offers three choices from which to choose the meaning. The stimulus words, sentences, and choices are presented in written form as well as auditorily, with each word presented on an index card throughout the cueing hierarchy. Allowing students access to the orthographic representation of the target word has two purposes: to reduce the load on working memory and to encourage the visual detection of the root word. Two practice words are presented first and not scored.

3.3.3.2 Stimulus words

As explained by Larsen & Nippold (2007), the stimulus words used were adapted from Carlisle (2000). Fifteen (not including the two practice words) low frequency derivatives of high frequency root words were chosen. Frequencies were determined using *The American Heritage Word Frequency Book* (Carroll, Davies, & Richman, 1971). High frequency words were those occurring at least once in 100,000 words. Low frequency words are those occurring fewer than once per 1 million words of text. High frequency words should be familiar to the average sixth grader, while low frequency words are unlikely to be familiar to a sixth grader (Carlisle, 2000). The stimulus words all contained a high frequency root form, but a low frequency derived form. Using unfamiliar words with a familiar root form was meant to encourage students to use morphological analysis to determine the meaning of the derived form.

Using a method similar to that utilized by Larsen and Nippold, the DATMA was augmented with the addition of five new items: one noun compound, one compound adjective, two derived verbs, two derived adjectives and one derived nominal. Six of the original words were excluded from the final administration of the DATMA; these items—five test items and one practice set--were used to pilot the instrument during the first assessment session. The final version consists of a total of 15 test items: 2 compound words, 3 derived verbs, 5 derived nominals, and 5 derived adjectives. Two additional items are included in a practice set. Updated frequencies values were obtained for each polymorphous word and for each root based on *The Corpus of Contemporary American Words* (www.americancorpus.org; Davies, 2008-). See Appendix B for a complete list, including root and derived form frequency values, dictionary definitions of the derived forms, and the sentence contexts and choices used in the dynamic assessment protocol.

3.3.3.3 Scoring criteria

Two types of scores were obtained from the dynamic assessment task. The first involves determining the acceptable definitions of the target derived form. In general, the conventions established in Anglin (1993) and later adapted by Carlisle (2000) were followed. Definitions could include references to the root form, but must include evidence of understanding the change in meaning caused by the suffix. The second score obtained was based on how far the student progressed through the hierarchy of cues before providing an accurate definition. The following system is used here: For each word, a maximum of five points could be earned. For each step in the hierarchy presented before the student defines the word, one point is deducted.

5 points = child answers #1 and #2 correctly and completely

4 points = child explains the word correctly after prompt #3

3 points = child explains the word correctly after prompt #4

1 point = child explains the word correctly after prompt #6

0 points = child does not explain the word correctly

Thus the maximum score per word is five and the minimum is zero. This scoring system is designed to capture individual differences among the participants based on the amount of assistance needed to successfully define a target word.

It is hoped that dynamic assessment will reveal the zone of proximal development for morphological analysis and thus provide different information than that provided by previous studies using traditional, static tasks. In this case, the zone of proximal development would be the area between the child's ability to provide an adequate definition of a morphologically complex word without any support and the child's ability to provide a definition of the word with high levels of support. If children are not able to make use of the cues, it is likely that most scores will be either a five, suggesting the child was able to define the word with no assistance; or a zero, indicating that the child was unable to define the word.

3.3.3.4 Inter-rater reliability

To examine inter-rater reliability, two raters independently scored 20% of the protocols using audio recordings made during the administration of the DATMA. Inter-rater reliability as assessed by Pearson correlation was strong ($r = .85$). Disagreements were resolved through discussion.

4.0 RESULTS

This chapter will report the findings of this research. First, descriptive statistics for each of the measures are provided. Then, each of the research questions posed in this study is individually addressed. All statistical calculations were carried out with SPSS 17. Data are reported for 27 participants who completed the study. Participants included 17 boys and 10 girls drawn from three different sixth grade classrooms. Mean age of participants was 161.28 months (SD = .393) or 12 years 9 months.

4.1 DESCRIPTIVE STATISTICS

Means and standard deviations for each of the measures used in this study are presented in Table 1. As described in the Methods section, standard scores were used for the PPVT-4. The number of words read correctly per minute was used as the score for the oral reading fluency task and scaled scores were used for the PSSA. Raw scores, obtained as described in the Methods section, were used for the two experimental measures.

In this study, a dynamic assessment was used to measure the morphological analysis skills of participating sixth graders. The primary aim of this research was to examine the utility of a dynamic assessment of morphological problem-solving and specifically to identify whether this measure can differentiate among students based on reading proficiency. Because the dynamic instrument is an experimental one, reliability of these scores was examined. In order to assess the internal consistency of the data obtained by the dynamic measure, item score on the DATMA was tested with Cronbach's alpha yielding a reliability score of .698. Reliability is expressed as a statistical index with values ranging from 0 (not at all reliable) to 1 (perfectly reliable). It is generally accepted that an internal consistency alpha of .70 is satisfactory, and that an alpha of .80 or higher is desirable. Nunnally (1967) stated that when conducting early-stage research, relatively low reliability coefficients (.5) are acceptable. Because this study is highly exploratory in nature, it was determined that reliability above .6 would be acceptable.

Table 1

Mean Scores

Measure	Mean	SD	Range
DATMA	30.74	8.332	10-48
TMS	35.59	10.210	11-49
PPVT-4	98.48	10.416	79-128
ORF	115.30	37.118	20-208
PSSA	1209.30	202.766	780-1573

Note. DATMA = Dynamic Assessment Task of Morphological Analysis; TMS = Test of Morphological Structure; PPVT–4 = Peabody Picture Vocabulary Test—4; ORF= oral reading fluency task; PSSA = Pennsylvania statewide assessment in reading

4.1.1 Morphological Analysis Skills

Dynamic Assessment Task of Morphological Analysis (DATMA; Larsen & Nippold, 2007) is designed to measure differing levels of morphological ability. Participants are presented with a series of derived words and asked to define each one; a series of cues are given until either all the cues are exhausted or the child defines the word correctly. The mean raw score for the Dynamic Assessment of Morphological Analysis was 30.741 (*SD* = 8.332) out of a possible maximum score of 75. An inspection of the data showed that the distribution of scores associated with the DATMA approximated a normal distribution, with a skewness statistic of -.46. The data are negatively skewed, meaning that the left tail is longer. In general, if skewness is between −.5 and +.5, the distribution is approximately symmetric (Bulmer, 1965). Scores ranged from a low of 10 (13%) to a high of 48 (64%). This variability indicates that some children were able to use morphological analysis to explain the meanings of unknown words while others required greater amounts of scaffolding. Inspection of the data revealed that the majority of scores (19) fell between 27 and 37. Five scores fell below this range and three fell above this range. Overall, the participants in this study had difficulty explaining the meanings of low frequency morphologically complex words. Only three students scored above 50% of the total possible points for the task.

The mean score for the Test of Morphological Structure was 36.593 (*SD* = 10.21) out of a maximum possible score of 56. Scores on this measure ranged from a low of 11 and a high of 49. The majority of scores (56%) ranged 40 – 49, a correct response rate of 71% - 86%, suggesting that this task was easier than the task required by the DATMA, where only three participants scored above 50% of the possible points correct. One part of the TMS required participants to produce a derived word to finish a sentence (Production). The second part of the task required the participants to decompose derived words in order to finish sentences (Decomposition). Each task contained 28 items. The base and derived forms were equivalent in word frequency on both the Production and Decomposition tasks. The two tasks contained both phonologically transparent relations (e.g., *discuss-discussion*) and phonologically opaque relations, where there was a phonological shift from the base to the derived form (e.g., *produce-production*).

Based on earlier research which suggested that the Production task may be more closely associated with the ability to define morphologically complex words than is the Decomposition task (Carlisle, 2000), the two subtests of the TMS were analyzed separately. For the Production task, the examiner read the target base word followed by a short sentence with its final word missing. The participants were asked to complete each sentence by orally producing another form of the target word that correctly completed the sentence (e.g., *Help. My sister is always ______. [helpful]).* For the Decomposition task, the examiner read a target derivative followed by a short sentence with its final word missing. Again, the participants orally produced another form of the target word, the base word itself (e.g., *Improvement. My teacher wants my spelling to ____. [improve]).* Scores on the TMS-Production and TMS-Decomposition were determined by the total number of correct responses produced for each task. .

The Decomposition task mean was 15.70 (*SD* = 4.410). The Production task mean was 19.33 (*SD* = 7.077). The two tasks were positively correlated with one another ($r = .385$, $p = .047$). A paired samples *t*-test calculated the difference between groups as significant ($t = -2.797$, $p = .010$,). To assess whether the TMS was a reliable measure of morphological awareness, estimates of internal consistency were calculated. The TMS had a Cronbach's alpha of .87, with similar results when the two tasks were analyzed separately (.86 for Production and .90 for Decomposition). The DATMA score correlated significantly with TMS score

(r = .589, p = .001), suggesting that the dynamic assessment is measuring a similar construct as the TMS. When examined separately, the DATMA was significantly correlated with the Production task only (r = .536, p = .004).

One central purpose for utilizing a dynamic assessment was to determine whether in fact this instrument would reveal various levels of morphological analysis ability. It was hypothesized that a dynamic measure would do this better than would a static measure. If sixth graders did make use of the cues to determine the meanings of unfamiliar words we would expect items to reveal a range of scores. This is in fact what was shown. Mean item scores ranged from 1.407 to 3.037. If participants did not make use of the prompts, the majority of item scores would be either five, indicating that the child knew the meaning of the word independently or zero, suggesting the student did not benefit from the clues provided by the assessor. Table 3 contains the mean score for each item on the DATMA.

To further explore the performance on the DATMA, the frequency with which each of the six different scores (5, 4, 3, 2, 1, 0) had been assigned to the 15 target words was tabulated for each participant. The most common scores were 1 and 2, followed by 4 and 0; scores of 3 and 5 were least common. A score of one indicates that the student was only able to provide the definition when presented with three choices. A score of two indicates that the student was able to successfully define the word when a sentence provided context clues to the meaning. Five-point responses indicated a student knew the word with no additional prompting needed and mentioned the constituent morphemes in the definition produced. These patterns suggest that

Table 2

Means and Standard Deviations for Items on DATMA

Item	Mean	Standard Deviation
Boathouse	2.52	1.50
Cookery	1.70	1.07
Equalize	1.78	.70
Fearsome	2.30	1.24
Hospitalize	2.74	1.83
Idealize	1.67	1.14
Moonlit	2.48	1.56
Oddity	1.33	1.39
Odorous	1.48	1.19
Preventative	2.33	1.30
Puzzlement	1.51	1.34
Reliance	1.41	1.28
Secretive	2.96	1.40
Sparkly `	3.04	1.48
Stardom	1.44	.51

*maximum possible score = 5

although some variation was present, the majority of participants needed several prompts before they were able to successfully define the stimulus words. Of the 432 individual item scores, 144 (33%) were assigned a score of one. Participants benefited from the contextual clue prompt 138 times. Together the scores one and two, which reflect the need for the greatest amount of scaffolding, were assigned to 65% of the total responses. Only 20% of the responses show students able to provide the correct definition after minimal prompting (scores of three and four). Twice as many responses (42) were given a score of zero, revealing that a child was unable to produce the definition after exhausting all of the prompts, than were given a score of five (21). Only 5% of all responses were assigned a score of 5. Table 3 presents frequency data for the scores.

Table 3

Frequency of Each Score Assigned on the DATMA

Score	Frequency	Percent	Cumulative Percent
0	42	9.7	9.7
1	144	33.3	43.1
2	138	31.9	75.0
3	7	1.6	76.6
4	80	18.5	95.1
5	21	4.9	100.0
Total	432	100.0	

4.1.2 Vocabulary Knowledge

This study utilized the PPVT-4 to measure general vocabulary knowledge. This assessment was chosen based on its high reliability and validity for research purposes. The participants' mean standard score on the PPVT-4 was 98.48 (*SD* = 10.416). Participants' scores fell between 79 and 128, revealing a wide range of word knowledge. As a standardized measure of receptive vocabulary, the PPVT-4 score was significantly and positively correlated with the Production task subtest of the TMS ($r = .387$, $p = .046$), but not with the overall TMS or the Decomposition task. The scores on the PPVT-4 were positively and significantly correlated with the ORF ($r = .624$, $p = .001$), the PSSA ($r = .515$, $p = .006$), and the DATMA ($r = .408$, $p = .034$)

4.1.3 Reading Proficiency

Two measures were utilized to assess the participants' reading skills: an oral reading fluency task was administered by the researcher and scores from the statewide reading proficiency exam were obtained. The mean score on the oral reading fluency task was 115.30 (*SD* = 37.118). The goals and cut-points utilized in this study are based on CBM normative data from fourth and fifth grade students from Hasbrouck and Tindal (1992) and the average slope of reading progress information from Fuchs, Fuchs, Hamlett, Walz, and Germann (1993). These norms for end of year sixth grade, based on words read correctly (WRC), are as follows: (a) 125 WRC and above = no risk; (b) 104 – 124 WRC = some risk; (c) 103 WRC and below = at risk. The majority of students who participated in this study are below average readers. Only five participants scored at the grade level benchmark on the ORF. Eleven students scored "at some risk" and 11 more scored "at risk."

The mean scaled score on the statewide reading assessment was 1209.30 (*SD* = 202.766). As reported in the Methods section above, Pennsylvania's General Performance Level Descriptors are Advanced, Proficient, Basic, and Below Basic. The Advanced level reflects superior academic performance, indicating an in-depth understanding and an exemplary display of skills. The Proficient Level reflects satisfactory academic performance, indicating a solid understanding and adequate display of skills. The Basic Level reflects marginal academic performance, indicating a partial understanding and limited display of skills. The Below Basic Level reflects inadequate academic performance, indicating little understanding and minimal display of skills. For sixth grade students, the Pennsylvania Department of Education defines a score of 1456 and above as "Advanced," a score of 1278-1455 as "Proficient," a score of 1121-1277 as "Basic" and 1120 and below as "Below Basic." In this study's sample, the majority of students (67%) scored "Basic" or "Below Basic." Nine students scored "Proficient" or above. This large number of scores at the lower end of the distribution indicates that most of the children who participated in this study demonstrated below average reading proficiency.

4.2 RESEARCH QUESTION ONE

The first research question posited by this study concerned the validity of the DATMA to measure morphological analysis skills. To examine the concurrent validity of the dynamic measure, raw scores on the DATMA were compared to a static measure of morphological awareness (TMS) used in previous research (e.g., Carlisle, 1995, 2000; Katz & Carlisle, 2002). In addition, the DATMA was compared to other measures tapping skills important to reading: specifically, accepted measures of vocabulary knowledge and reading fluency. In particular, the following sub-questions were addressed: what is the relationship between the DATMA and the TMS, a static measure of morphological awareness? What is the relationship between the DATMA and the PPVT-4, a standardized measure of receptive vocabulary? What is the relationship between the DATMA and an oral reading fluency task? A correlation matrix is provided in Table 4, reporting Pearson correlation coefficients for all predictor measures used in the study. Moderate to large correlations were found between the DATMA and the other measures. An explanation of these relationships follows.

Table 4

Correlations for Predictor Variables

Measure	DATMA	TMS	PPVT-4	ORF
1. DATMA	___	.589**	.408*	.622**
2. TMS		___	.365	.490**
3. PPVT-4			___	.624**
4. ORF				___

Note. DATMA = Dynamic Assessment Task of Morphological Analysis; TMS = Test of Morphological Structure; PPVT–4 = Peabody Picture Vocabulary Test—4; ORF= oral reading fluency task.

* $p < .05$

**$p < .01$

4.2.1 The DATMA and Morphological Awareness

Performance on the DATMA was positively correlated with the total TMS score (r =.589, p = .001). This association suggests that the DATMA is measuring a construct similar to that measured by the TMS. Earlier research by Carlisle reported that the TMS Production task was more closely related to a child's ability to define an unknown word than was the Decomposition task. Based on this finding, the two parts of the TMS were analyzed separately. Results of this study support these earlier findings. Only the Production task was significantly correlated with the DATMA (r = .536, p = .004).

4.2.2 The DATMA and Receptive Vocabulary Knowledge

The DATMA was positively correlated with the PPVT-4 score (r = .408, p = .034). No such significant relationship was revealed between the Test of Morphological Skill and the PPVT-4 (r = .365, p > .05). This interesting finding suggests that the DATMA may be capturing something unique with regards to vocabulary knowledge that the static measure of morphological knowledge does not.

4.2.3 DATMA and Reading Fluency

Scores on the DATMA were significantly correlated with a measure of reading fluency. The relationship between the DATMA and the oral reading fluency assessment was positive and significant (r = .662, p = .001). The DATMA score was the only measure to demonstrate significant correlations with all of the other predictor variables. That is, the dynamic assessment score was highly correlated with the TMS overall, TMS Production, PPVT-4, and the ORF. This finding is worthy of note because it speaks to the potentially robust nature of the dynamic assessment in accounting for several reading-related skills.

In sum, concurrent validity was established and is illustrated by the DATMA showing a positive and significant correlation with another instrument presumed to be measuring a common underlying construct (i.e., the Test of Morphological Skill). Content-relevancy was supported by the strong correlations between the DATMA and other measures which assess the critical literacy skills of vocabulary (PPVT-4) and reading fluency (ORF). A considerable body of evidence has demonstrated that morphological knowledge is associated with vocabulary (Carlisle & Fleming, 2003; Nagy, Berninger, Abbott, Vaughn, & Vermeulen, 2003; White, Power, and White, 1989), word reading (Carlisle, 2000; Fowler & Lieberman, 1995; Singson, Mahoney & Mann, 2000) and comprehension (Carlisle, 1995; Carlisle & Flemming, 2003; Deacon & Kirby, 2004; Windsor, 2000).

4.3 RESEARCH QUESTION TWO

The next question posed examined the extent to which the DATMA score predicted reading achievement. In order to explore this question a regression analysis was conducted with the state reading proficiency test (PSSA) as a criterion variable. By the method of least squares, simple linear regression calculates the coefficient of determination, R^2, a measure of the proportion of variability explained by, or due to, the linear relationship in a sample of paired data. It is a number between zero and one and a value close to zero suggests a poor model. It was hypothesized that the DATMA score would be a significant predictor of reading achievement as measured by the PSSA. To test whether this was the case, a simple regression analysis was run using the raw score from the DATMA as the predictor variable and PSSA score as the criterion variable. Analysis revealed that the DATMA significantly predicts reading achievement on the PSSA. As reported in Table 5, the DATMA explained 37% (R^2= .371; adjusted R^2 = .346) of the variance in PSSA scaled scores. This indicates that better performance on the morphological analysis task was associated with higher reading achievement scores.

Table 5

Predicting PSSA Performance with DATMA

	B	*SE*	*β*	*t*-ratio	*p*
Constant	753.749	122.797		6.138	.000
DATMA	14.819	3.860	.609	3.839	.001

R^2 = .371
Adjusted R^2 = .346

It was anticipated that the DATMA would demonstrate predictive validity in the regression analysis. It was of interest to compare the strength of the prediction to that of the Test of Morphological Structure. The results of a regression analysis of the TMS are presented in Table 6. The TMS was a significant predictor of reading achievement in this study, supporting previous research findings (Carlisle, 1995, 2000; Katz & Carlisle, 2002). The TMS accounted for 16% of the variance in PSSA scores. Figure 1 provides the scatter plots for both predictor variables.

When doing regression, the cases-to-Independent Variables (IVs) ratio should ideally be 20:1; that is 20 cases for every IV in the model. The lowest your ratio should be is 5:1, i.e., 5 cases for every IV in the model (Gelman & Hill, 2007). Based on this study's low *n*, simple linear regression offered adequate protection from Type II errors. However, to further compare

Table 6

Predicting PSSA Performance with TMS

	B	*SE*	β	*t*-ratio	*p*
Constant	916.743	137.937		6.646	.000
TMS	7.995	3.636	.403	2.199	.037

R^2 = .162
Adjusted R^2 = .129

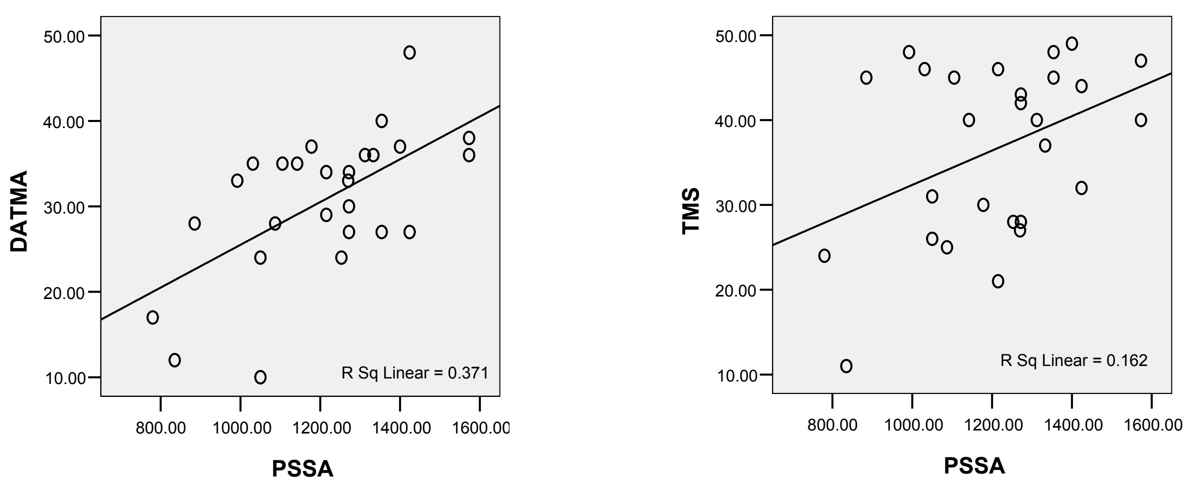

Figure 1

Scatterplots

the power of both the DATMA and the TMS to predict reading achievement, exploratory multiple regression analyses were performed. Both predictor variables were entered first together and then secondly in a stepwise fashion, with the TMS entered in the analysis first. The TMS was entered in the analysis first to see whether the DATMA would account for unique variance. Table 7 presents the results when the variables were entered together. Table 8 presents the results when the variables were entered sequentially.

Table 7

Multiple Regression Analysis with Predictor Variables Entered Simultaneously

	B	*SE*	*β*	*t*-ratio	*p*
Constant	734.554	137.486		5.343	.000
DATMA	13.857	4.865	.569	2.849	.009
TMS	1.333	3.970	.067	.336	.740

R^2 = .374
Adjusted R^2 = .322

When the predictor variables were entered in a stepwise fashion, the TMS dropped out of the model (see Table 8). While these results should be interpreted very cautiously, some preliminary evidence suggests that the dynamic measure outperforms the static measure in predicting later reading achievement as measured by the PSSA.

Table 8

Stepwise Regression Analysis

	B	*SE*	β	*t*-ratio	*p*
DATMA	14.749	122.797	.609	3.839	.001
Constant	753.749	3.860		6.138	.000

R^2 = .371
Adjusted R^2 = .346

4.4 RESEARCH QUESTION THREE

One of the main interests of this study was to determine whether the DATMA would be able to differentiate between students with reading deficits. To explore this hypothesis, the third research question posed herein asks, how well does the DATMA identify students with the most intensive needs? Also, how does this compare with how well the TMS identifies students with the most significant reading problems? Several types of data analyses were enlisted. First, it was hypothesized that students with markedly different levels of need for reading intervention would represent significantly different DATMA scores. A one-way ANOVA was calculated to test the following null hypothesis: students with differing levels of reading difficulty would not have significantly different mean DATMA scores. Participants were divided into three groups based on ORF scores. Oral reading fluency tasks are frequently used in schools to sort students for different levels of reading intervention. As explained above, the cut points used to sort students were those reported by and Hasbrouck and Tindal (1992). Thus, in this study's sample, 11 participants would have been selected for Tier 3 (intensive) instruction, 11 would have been selected for Tier 2 (strategic) instruction, and five would be selected for Tier 1 (benchmark) instruction. Results of the analysis are presented in Table 9.

Table 9

Analysis of Variance

Source	Sum of Squares	df	Mean Square	*F*	Sig.
Between Groups	600.531	2	300.264	4.982	.008**
Within Groups	1204.655	24	50.194		
Total	1805.185	26			

Note. **p < .01

Post hoc Tukey tests reveal significant differences in mean DATMA scores between students selected for Tier Two and Tier Three (p = .044) and for Tier One and Tier Three (p = .012), but not for Tier One and Tier Two (p = .507). The significant difference observed between the intensive and strategic groups on the DATMA suggests that the dynamic measure might be useful to screen students for Tier Two and Tier Three instruction, potentially identifying those struggling readers who are in need of the most intensive intervention and are less likely to respond to interventions that lack sufficient intensity, such as those more likely used in Tier Two instruction. In contrast, no significant differences were detected in the DATMA scores of students whose ORF scores assign them to strategic interventions and those whose ORF scores place them at grade level, or benchmark. This provides preliminary support for the hypothesis that a dynamic assessment would be most sensitive to differences at the lowest end of the distribution.

It was also of interest to compare the ability of the DATMA to differentiate between varying levels of reading difficulty with that of the TMS. A one-way ANOVA was calculated with the TMS as dependent variable and tier assignment as the within subjects factor. The findings from this analysis indicate no such significant differences were found on TMS scores when comparing the three groups. These results are reported in Table 10.

Table 10

Analysis of Variance for TMS

Source	Sum of Squares	df	Mean Square	*F*	Sig.
Between Groups	306.773	2	153.387	1.531	.237
Within Groups	2403.745	24	100.156		
Total	2710.519	26			

Similarly, when the two parts of the TMS, the Production task and the Decomposition task, were analyzed separately no significant differences were found among the three groups. Tables 11 and 12 report these results. These findings suggest that something in the dynamic nature of the DATMA is capturing differences a static measure is unable to detect.

Table 11

Analysis of Variance for TMS-Production

Source	Sum of Squares	df	Mean Square	*F*	Sig.
Between Groups	58.830	2	29.415	1.580	.227
Within Groups	446.800	24	18.617		
Total	505.630	26			

Table 12

Analysis of Variance for TMS-Decomposition

Source	Sum of Squares	df	Mean Square	*F*	Sig.
Between Groups	88.473	2	44.236	.875	.430
Within Groups	1213.527	24	50.564		
Total	1302.000	26			

4.5 RESEARCH QUESTION FOUR

If the DATMA is indeed sensitive to smaller differences among struggling readers, then it may prove useful as a screening measure in an identification process for levels of intervention such as within an RTI framework. To assess its usefulness in this regard, additional analyses were carried out. Two types of classification analyses were used to determine the sensitivity and specificity of the measures. Research question 4(a) concerned the sensitivity of the DATMA to identify students for various tiers of intervention as set by scores on the oral reading fluency task. Thus, in the first of these analyses, a dichotomous variable was created for risk status. Students identified as needing intensive, Tier Three instruction based on their ORF scores were compared to those identified as strategic or benchmark. Those students identified as needing intensive instruction (ORF $\leq$ 104) were assigned to the "at risk" of becoming a non-responder group and those not identified as in need of Tier Three (ORF >104) were assigned to the "likely to respond" group. The DATMA was then used as a supplemental screening measure to re-classify students deemed to be intensive. The cutoff score used for the DATMA was 35. Other cutoff scores on the DATMA were also investigated but this value provided the best trade-off between sensitivity and specificity. Participants who were below criterion (35) on the DATMA were considered to be at risk of becoming non-responders. Those above criterion were considered to be not at risk or likely to respond to intervention. Classifications were categorized in the following manner: (1) True Positives- below average morphological analysis skills and identified as at risk; (2) True Negatives- above average morphological analysis skills and identified as not at risk or likely to respond; (3) False Positives- above average morphological skills and identified as at risk; and (4) False Negatives- below average morphological skills and identified as not at risk or likely to respond. Sensitivity and specificity of prediction were then used as indices in the evaluation of this screening approach.

Sensitivity, the ability of the dynamic measure to identify those sixth grade students who scored at or below the criterion on the ORF, was calculated by dividing the number of true positives by the sum of the true positives and false negatives. Specificity, the ability of the dynamic measure to identify students who scored above the criterion on the ORF, was calculated by dividing the number of true negatives by the sum of the true negatives and false positives. The results are presented in Table 13. The classification analyses indicated that the dynamic assessment measure resulted in excellent sensitivity (89%) and acceptable specificity (56%) for risk status according to the oral reading fluency score.

Table 13

Classification Analysis

	TP	TN	FP	FN	Sensitivity	Specificity
DATMA < 35	8	10	8	1	89%	56%

Note. TP = true positives; TN = true negatives; FP = false positives; FN = false negatives

While it may be useful to locate a cut score on the DATMA which accurately places students in either at-risk or no-risk categories to match those identified by the ORF, this analysis does not enable a comparison between the DATMA and ORF as screening measures for reading achievement outcomes. One of the primary concerns of districts today is to increase their pass rates on state accountability tests such as the PSSA. As such, a comparison of the ORF and the DATMA to predict proficiency status on the PSSA was undertaken. Research question 4b of this study asked: does the addition of the DATMA offer a value-added over the ORF score to predict performance on a reading outcome measure? In order to complete this analysis, logistic regression was employed to examine the extent to which the predictors (ORF and DATMA), alone or in combination, predicted students' risk of scoring below proficiency on the Spring PSSA.

Simple logistic regression is analogous to linear regression, except that the dependent variable is nominal, not a measurement. One goal of logistic regression is to see whether the probability of getting a particular value of the nominal variable (i.e., PSSA proficiency) is associated with the measurement variable; the other goal is to predict the probability of getting a particular value of the nominal variable, given the measurement variable. In logistic regression, the predicted dependent variable is a function of the probability that a particular subject will be in one of the categories (for example, the probability that a student will meet proficiency on an achievement test, given his or her set of scores on the predictor variables).

Three separate logistic regression analyses were run. The first two logistic regressions included the ORF (Model 1) and the DATMA (Model 2) as individual predictors of reading proficiency status. The next logistic regression (Model 3) included ORF then DATMA, entered sequentially, to predict reading proficiency status. Table 14 reports the results of this analysis. We can now use this model to predict the odds that a subject with a given score will meet proficiency on the PSSA. The .898 odds ratio for the ORF indicates that the odds of meeting proficiency on the PSSA increase by a multiplicative factor of .898 for each word read correctly increase on the ORF. The DATMA's effect is slightly smaller. A one point increase on the DATMA is associated with the odds of meeting proficiency by a factor of .801.

Chi-square tests were calculated to test the significance of adding predictor variables to the model. Chi-square is a test of the null hypothesis that adding the predictor variable(s) to the model has not significantly increased our ability to predict a particular nominal value in the outcome measure.

Table 14

Logistic Regression Analysis for Predicting PSSA Proficiency Status

	B	*SE*	*Wald*	*df*	*p*	Odds Ratio
Model 1						
ORF	-.106	.054	3.804	1	.051	.898
Model 2						
DATMA	-.221	.109	4.119	1	.042	.801
Model 3						
ORF	-.088	.047	3.572	1	.059	.916
DATMA	-.164	.152	1.167	1	.280	.849

Note. ORF = oral reading fluency; DATMA = Dynamic Assessment of Morphological Analysis

As shown in Table 15, each of the models significantly improved our ability to predict proficiency on the PSSA for this study's sixth graders. The -2 Log Likelihood statistic measures how poorly the model predicts the decisions -- the smaller the statistic the better the model. The findings therefore indicate that of the two predictor variables, the oral reading fluency is more accurate than the DATMA. However, we also see that the model in which the two measures are combined offers the best predictive ability, surpassing either the ORF or the DATMA used alone.

Table 15

Chi-Square Tests of Significance

Predictor Variable	Chi-square	df	*p*	-2 Log Likelihood
ORF	15.747	1	.000	18.625
DATMA	7.888	1	.005	26.483
ORF+ DATMA	17.201	2	.000	17.170

Note. ORF = oral reading fluency task; DATMA = Dynamic Assessment Task of Morphological Analysis

Ultimately, the results of our logistic regression can be used to classify participants with respect to whether we expect them to meet proficiency on the PSSA. The odds ratios obtained in our logistic regression analysis of the combined ORF/DATMA model (Model 3) can be converted to probabilities for each of the participants. However, before participants can be classified, a decision rule must be set. Thus, if the probability of the event is greater than or equal to some threshold, we shall predict that the event will take place. By default, most statistical packages set this threshold to .5. Using the default threshold, SPSS will classify a participant into the "proficiency" category if the estimated probability is .5 or more. SPSS will classify a participant into the "below proficiency" category if the estimated probability is less than .5. Table 16 indicates that this rule allows us to correctly classify students a total of 85.2% of the time when we utilize both the ORF and DATMA to predict proficiency status on the PSSA. Moreover, as shown in the classification table, using the DATMA as a supplemental screening measure along with the ORF results in 94.4% sensitivity and 66.7% specificity.

Table 16

Classification Table with Combination of ORF and DATMA to Predict Proficiency

Observed		Predicted		
		Proficient	Not Proficient	Percentage Correct
Proficient	9	6	3	66.7%
Not Proficient	18	1	17	94.4%
Overall Percentage				85.2%

Note. Cut value is .5

Using a model that includes only the intercept (or constant) and no predictor variables results in correctly predicting PSSA proficiency status only a total of 66.7 % of the time. Given the observed rates of the two outcomes--that is 9/27 or 33.3% met proficiency and 18/27 or 66.7% failed to reach proficiency--and no other information, the best strategy is to predict, for every case, that the participant will fail to meet proficiency on the PSSA. Using that strategy, you would be correct 66.7% of the time. Therefore we can quantify the improved odds of prediction by comparing the overall percentage correct using no other information versus the overall percentage correct when utilizing the two predictors together. The multi-measure screening model improved the odds of correctly identifying students at risk for not meeting proficiency on the statewide test by 18.5%. These findings have perhaps more practical significance because the level of sensitivity was improved to almost perfect sensitivity (94.4%). In school contexts, it is preferable to increase the number of true positives so that students who are most at risk can receive intervention services and perhaps alter their performance trajectory.

5.0 DISCUSSION

The purpose of Chapter 5 is to discuss and interpret the findings of this study. The chapter will address the research questions previously identified, explain the results, and recommend future directions. First, a summary of the study's results will be provided. A discussion of these results and the conclusions drawn from them will follow. Next, the limitations of this study will be outlined, the educational implications of the findings will be discussed and recommendations for further research will be offered. Finally, the overall conclusions formed as a result of the study will be provided.

5.1 STUDY OVERVIEW

The purpose of this research was to investigate the validity and usefulness of a dynamic assessment of morphological problem-solving with sixth grade students. Twenty-seven participants, including three identified as needing special education services, underwent two testing sessions. For the purposes of this research, four assessments were administered: a standardized measure of receptive vocabulary (PPVT-4), an oral reading fluency task (ORF), and two measures of morphological knowledge, one static in nature (TMS) and the other a dynamic task which utilizes a standardized graduated prompting approach (DATMA). Four research questions were posed and a quantitative methodology was applied to address them. The following suppositions were put forward: first, in order to support the concurrent validity of the DATMA, it was expected that the dynamic task would need to demonstrate a strong correlation with the TMS, a static measure of morphological analysis skill used in previous studies (e.g., Carlisle, 1995; 2000; Carlisle & Katz, 2006). Secondly, to support the content-relevancy of the DATMA, it was expected that results from the dynamic task would converge with existent research findings and exhibit positive and significant associations with well-validated measures of vocabulary knowledge and reading skill. Third, to establish predictive validity of the dynamic measure, it was further hypothesized that the DATMA would account for significant variance in a criterion measure of reading achievement. The criterion measure utilized in this research was the statewide reading proficiency assessment (PSSA), administered in the Spring following this study. Finally, to explore the potential utility of the DATMA as a screening tool, the dynamic measure was analyzed to determine if it offered acceptable specificity and sensitivity in screening students for more intensive reading interventions. Correlational analyses (question 1) were used to examine the relationships among the various measures. The adjusted R^2 statistic offered information regarding the degree to which the significant linear relationship between the DATMA and the PSSA was meaningful (question 2). The extent to which the dynamic task was able to discriminate between at-risk and grade-level readers was tested with an ANOVA (question 3). Finally, evidence of validity according to classification accuracy was also examined (question 4).

5.2 MAJOR FINDINGS

Several significant results can be drawn from this study. Evidence was generated that the DATMA was both internally consistent and valid. The task had high estimates of internal consistency and seemed to perform similarly to other measures of morphological problem-solving found in the literature. Correlations between the DATMA and measures of vocabulary knowledge and reading fluency were moderate to large in size and positive, also making it consistent with other studies and suggesting that the dynamic measure captures skills related to these critical literacy constructs. Given the research to date on the relationships between morphological awareness and reading, these results were expected. A preponderance of evidence identifies an significant role for morphological knowledge in reading, beginning in the early to middle elementary grades (e.g., Carlisle, 1995; Carlisle, 2000; Carlisle & Stone, 2003; Windsor, 2000) and growing in importance during the upper elementary and middle school years (Mann & Singson , 2003; Nagy, Berninger, & Abbott, 2006). Additionally, scores from the dynamic assessment task indicated that the participating sixth graders experienced difficulty in defining morphologically complex words. These results too are consistent with previous research.

Other interesting findings occurred in this study which extend our understanding of the role of morphological problem-solving in reading performance. First, the DATMA emerged from the analyses as a potentially useful screening measure, capable of differentiating sixth graders at risk for reading difficulty and including procedures and content more sensitive than a static measure of morphological knowledge. The results indicated that the DATMA significantly improved the prediction of reading outcomes over and above the static measure, suggesting that the dynamic nature of the former contributed to the prediction accuracy. The predictive validity of the DATMA was also compared to the oral reading fluency task, a commonly used reading screening measure. Results showed that the dynamic measure added significantly to the prediction of reading outcomes. Exploratory analyses examined the use of the DATMA as a supplemental screening measure to the oral reading fluency task. The DATMA reduced the number of false positives, and in some cases, predicted reading outcomes as well as a combination of the two measures. As such, the results of this study provide preliminary support for the usefulness of a dynamic assessment of morphological awareness within an RTI framework for sixth graders.

5.2.1 Research Question One

The first question posed concerned the reliability and validity of the DATMA, an experimental measure with limited use in previous research. In an RTI framework, assessments are used to inform placement of students into supplemental reading instruction, and thus, it is critical that assessments used in this process have adequate reliability. The extant literature on dynamic assessment has not typically reported the reliability of the measures (i.e., Caffrey, 2006; Ferrara et al., 1986). This is likely due to the nature of treatment-oriented dynamic assessments, which are often lengthy and lack standardization (e.g., Budoff, 1974; Feuerstein, 1979). However, because this study relied on the use of a standardized graduated-prompts methodology, the measurement of reliability estimates was possible. The inter-rater (.85) and internal consistency alpha reliabilities (.70) were moderate to high and met the acceptability levels proposed by

Nunnally (1967). Therefore, this study contributes to the literature by showing that a dynamic measure of morphological awareness can be reliable.

The concurrent validity of the DATMA was tested by examining the correlation coefficient between the dynamic measure and a static measure of morphological analysis skill, the Test of Morphological Structure (TMS). A significant and positive association between these two measures was found, supporting the concurrent validity of the DATMA. The content relevancy was further tested by examining the correlation coefficients between the dynamic measure and two well-validated measures of reading-related skills, one of vocabulary knowledge (PPVT-4), the other a commonly used screening measure, an oral reading fluency task (ORF). Previous research with the DATMA has documented a significant and positive relationship with measures of word knowledge and reading fluency (Larsen & Nippold, 2007). In a longitudinal study, Carlisle and Fleming (2003) found that first and third graders' analysis of complex words was predictive of vocabulary and reading comprehension two years later as measured by standardized tests. Results from the present study converge with previous research: facility with morphological analysis is associated with higher levels of vocabulary knowledge and reading proficiency.

The DATMA was designed to examine children's morphological problem solving abilities by asking them to define low-frequency morphologically complex words with high frequency bases. The results of other studies have suggested that vocabulary knowledge is significantly related to and/or predictive of morphological awareness (e.g., Carlisle, 1995; Carlisle & Nomanbhoy, 1993; Mann & Singson, 2003; Singson, et al., 2000; Windsor, 2000). Similarly, scores from the measure of vocabulary knowledge used in this study (PPVT-4) did correlate significantly with performance on the DATMA. It makes sense when one considers how the meaning of an unfamiliar morphologically complex word might be processed (see Schreuder & Baayen, 1995). It is reasonable to believe that one's vocabulary knowledge is related to the ability to use morphological analysis to derive meanings of unfamiliar morphologically complex words. Wysocki and Jenkins (1987) argued that much of the vocabulary growth in children that occurs in the middle to late elementary grades occurs because of morphological analysis. White, et al. (1989) recommended that children in grades four and above receive instruction in how to use morphological problem solving to derive meanings of unfamiliar words because this would lead to greater incidental word learning during reading.

Further, vocabulary knowledge has been found to be a strong predictor of reading comprehension (Freebody & Anderson, 1983; Kuhn & Stahl, 1998). Anglin (1993) found that children use morphological analysis to make meaning out of already known constituent morphemes. Given that the proportion of unfamiliar morphologically complex words found in grade-level texts is substantial in the upper elementary grades and beyond, the ability to use morphological analysis for making sense of texts may become critical. Moreover, it is possible that it is not strictly one's vocabulary knowledge *per se*, but one's ability to continue to build his/her vocabulary knowledge independently through morphological analysis that is most important for reading growth.

Further, separate from vocabulary knowledge, morphological analysis is also associated with word reading and comprehension. It seems likely that the ability to attend to the semantic and syntactic aspects of morphology is important for reading fluency. In this study, the DATMA correlated significantly with an oral reading fluency task. These findings are in line with Larsen & Nippold (2007). The sixth graders participating in this study had considerable difficulty capturing both the meaning and the grammatical roles of the suffixes in their definitions. For example, many student responses indicated knowledge of the base word but a lack of knowledge about the suffix. A representative response which illustrates this follows: when asked to define *equalize*, one participant's first response was "it means the same as." When prompted to identify the word's constituent morphemes, the student responded, "*equal*" and "*ize*." After asked if this helped him to define the word, the child responded, "when two things are the same, equal." The student was unable to correctly identify the grammatical function of the suffix in his definition. In none of his attempts to define *equalize* did he seem to be aware that the *–ize* suffix had changed the most frequent use of the word *equal* as an adjective to a verb. After the following sentence context prompt was provided: *At the track meet, Josh ran extra fast so he could equalize his time with his opponent's*, the child was still unable to make use of the word's function in the sentence to accurately define *equalize* as a verb. In fact, he was unable to select the correct definition (*to match*) from three choices, despite have a firm understanding of the meaning of the base word.

One interesting result from this research was that when analyzed by word type, derived nominals were the most difficult for participants to define. Derived nominals are morphologically-complex words whose derivational ending creates a noun from a verb or adjective (e.g., *reliance-rely, oddity-odd*) or creates a new noun from a noun whose meaning has somehow altered such as in the case of the agentive (e.g., *organist-organ*). In still other cases, the derivational ending creates a new noun from a word whose grammatical function fluctuates (e.g., *cookery-cook*) where *cook* can be both a noun (person) and a verb (to cook). In this study, participants' three most difficult items were all derived nominals: *oddity, reliance*, and *stardom*. In fact, five of seven of the lowest scored items were derived nominals. Verbs were the second most difficult word type.

Previous studies have demonstrated that children frequently ignore the grammatical function of suffixes. Only 10% of first graders and 37% of third graders took the meaning and grammatical role of the suffix into account (Carlisle & Fleming, 2003). Carlisle and Fleming concluded that children's access to representations of full forms, base forms, and affixes contributed to their reading ability in the late-elementary grades. The present study does suggest that for semantically transparent words with a stem that is familiar, students do seem to be sensitive to the stem and likely have a good sense of its meaning. However, a more strategic approach might be more important for recognizing stems that are less familiar or when the stem is obscured in sound and/or spelling. Beyond recognizing the stem, students need to learn to coordinate the use of affix meaning to help them determine the meaning of an unfamiliar morphologically complex word.

The idea that it is the combination of semantic and syntactic knowledge that makes a difference in reading comprehension is reasonable given that semantics and syntax seem to work together to help readers understand text (Perfetti, 1990). For example, Perfetti suggests that reading comprehension involves cooperation between semantics and syntax. He argues that semantic clues help readers bring meaning to ambiguous syntactic structures in text which leads to accurate syntactic parsing and processing of text. Nation and Snowling (2000) found that syntactic awareness was predictive of reading comprehension in good and poor readers. It seems likely that failing to be sensitive to the grammatical influence of affixes would interfere with comprehension and disrupt reading fluency. This is one potential conclusion that can be drawn from the results of the current study.

5.2.2 Research Question Two

The second research question concerned the predictive validity of the DATMA to account for variance in scores on a measure of reading achievement. The TMS has significantly predicted reading comprehension in previous studies (e.g., Carlisle, 1995; 2000; Katz & Carlisle, 2002). In Carlisle (1995), first graders' performance on the TMS-Production measure accounted for a significant 10% of the variance in their second grade performance on a reading comprehension measure. In Carlisle (2000), fifth graders' performance on the TMS (Production and Derivation combined) explained a significant 14% of the variability in reading comprehension performance. One would expect that a valid dynamic measure of morphological problem-solving would, like its static counterparts, also demonstrate moderate to strong correlations with reading outcome measures.

Results of simple linear regression analyses showed that the DATMA accounted for a significant amount of variance when predicting reading achievement as measured by the Pennsylvania statewide assessment in reading (PSSA). Likewise, as in previous studies, the TMS also correlated significantly with the PSSA. Furthermore, when the DATMA and the TMS were sequentially entered into a stepwise multiple regression model, the TMS dropped out as a significant predictor. These findings provide preliminary support for the idea that the dynamic nature of the DATMA improved the predictive ability of a static morphological analysis task.

5.2.3 Research Question Three

The third research question explored the DATMA's utility as a screening measure to identify students with the most significant reading difficulties. One of the central purposes of this study was to examine how well the DATMA could identify students with the most intensive needs. Of secondary interest was how the DATMA compares with the TMS to identify students with the most significant reading problems. To answer these questions, students were sorted into three intervention levels according to their oral reading fluency scores. As explained above, the cut points used to sort students were those reported by and Hasbrouck and Tindal (1992). The ORF was chosen because of its widespread use as a universal screening measure in RTI. It is common practice for school districts to use ORF scores to inform placement of elementary students in Tier 2 and Tier 3 reading interventions.

Results of a one-way ANOVA indicated that students sorted into the three levels of intervention based on the ORF scores did have significantly different mean DATMA scores. Post hoc tests reveal significant differences in mean DATMA scores between students selected for Tier Two and Tier Three and for Tier One and Tier Three. In contrast, no significant differences were found in the mean DATMA scores for students in Tier One and Tier Two groups. The significant difference observed between the intensive and strategic groups on the DATMA suggests that the dynamic measure might be useful to screen students for Tier Two and Tier Three instruction, potentially identifying those struggling readers who are in need of the most intensive intervention and are less likely to respond to interventions that lack sufficient intensity, such as those used in Tier Two instruction. In contrast, no significant differences were detected in the DATMA scores of students whose ORF scores assign them to strategic interventions and those whose ORF scores place them at grade level, or benchmark. This provides preliminary support for the hypothesis that a dynamic assessment would be most sensitive to differences at the lowest end of the distribution.

5.2.4 Research Question Four

To address the final research question, the usefulness of the DATMA as both an alternative to the ORF, and as a supplemental screening measure along with the ORF, was examined. While widely used in the earlier grades, the reliability and validity of the ORF as a screening measure in the upper elementary and middle school grades is not well established. It was of interest first, to see if the use of the DATMA alone would improve classification accuracy as compared to the ORF alone. Secondly, it was of interest to see if the DATMA would reduce the number of false positives when used as a supplement to the ORF as a screening measure. In the first set of analyses, the cutoff score recommended by the ORF was used to identify poor readers, and then, the DATMA was used to reclassify those students deemed to be at risk. Proficiency status on the PSSA was the criterion outcome variable.

In this study, analyses of the DATMA as a screening measure with PSSA proficiency status as the criterion measure resulted in high sensitivity but only adequate specificity. This procedure yielded excellent sensitivity (.94) when predicting proficiency status on the spring PSSA, with accompanying specificity levels of .67. Although the specificity values are lower than the preferred levels, they are quite consistent with an RTI approach. In fact, the DIBELS assessment is a frequent choice for screening and progress monitoring for RTI. Yet the classification accuracy for DIBELS is well below optimal levels. Schatschneider (2006) reported sensitivity and specificity levels of .52 and .85% for first grade students. Riedel (2007) reported .68 for both sensitivity and specificity at the same grade level. When the purpose for screening is preventative in focus, a greater emphasis should be placed on sensitivity, rather than specificity, to ensure students at-risk are not overlooked. On the other hand, when the purpose is identification for special education services, the cut points might need to be more stringent so as to reserve limited resources for those with the greatest need for them. Sensitivity is easily manipulated by adjusting cut-scores. If the cut-scores are raised sufficiently high enough, we are likely to identify all or nearly all of the children who will eventually fail the criterion test. The increased sensitivity, however, would be offset by decreased specificity because raising the cut-score means the screen will over-identify many students who are not really on the path to reading failure. A perfect screen would distinguish every student who needs intervention from every student who does not. Unfortunately, a perfect screen does not exist and schools must weigh the trade-offs of over- and under-identifying students at risk. When screening mistakenly over-identifies too many students as at risk (false positives), schools spend precious resources on those who do not need the supplemental intervention, and, as a result, may find remaining resources are inadequate to serve those who really need them. Screening accuracy is clearly very important.

Jenkins (2003) argues that an ideal screening measure in an RTI framework would offer a sensitivity level of $> .90$ and specificity of $> .80$. In practice it is difficult to obtain optimal levels with a single measure. Jenkins recommends using a combination of measures which can maximize sensitivity while maintaining manageable numbers of false positives. Compton, Fuchs, and Fuchs (2007) and Foorman et al. (1998) reported better screening accuracy using multiple measures in comparison to a single measure. There are surprisingly few studies of screening accuracy beyond the primary grades for reading. For older students, screens should include measures such as vocabulary and language comprehension as well as those which measure reading fluency (Davis, Lindo, & Compton, 2007).

In the final analyses, the classification accuracies were compared for models including combined screening, the DATMA alone, and the ORF alone. The supplemental screening approach did yield higher accuracy rates than the DATMA or the ORF alone. Although the results from this research are exploratory and only preliminary, they suggest that the DATMA alone or in combination with other measures is a valid and useful screening approach to place students in Tier 3 instruction. However, it is important to note that based on the sample in this study, three false positives (11% of the 27 participants) would have been mis-selected for intensive interventions when in reality they would demonstrate proficiency on the PSSA. Selecting as many as 11% of students for Tier 3 interventions who are not in need of those supplemental services represents a misallocation of a district‘s limited available resources--one which must be carefully considered when selecting a screening protocol for placing students within an RTI framework.

Nonetheless, my study‘s results are consistent with those of both Spector (1992) and O‘Connor and Jenkins (1999) who found that a dynamic measure improved prediction accuracy in young students over a static measure. It is likely that the feedback provided in the dynamic measure allowed students with partially developed morphological problem-solving skills to separate themselves from those with little to no knowledge. Findings from the current study, though exploratory, suggest that the DATMA may be useful as a supplemental screening measure in a two-step process to identify students at greatest risk of reading failure and one which, while not completely eliminating false positives, does reduce the rate of false positives associated with many screening measures such as the ORF.

5.2.5 The Use of Dynamic Assessment for Predicting Later Achievement

Before considering the use of the DATMA as a predictor of future reading achievement, it is critical to more closely consider the outcome measure utilized in this study. For the purposes of this research, scaled scores from the statewide accountability test in reading--administered in the spring of sixth grade--were obtained for each of the participants. The annual PSSA is a standards-based criterion-referenced assessment used to measure a student's attainment of academic standards while also determining the degree to which school programs enable students to attain proficiency standards. According to the Pennsylvania Department of Education, the PSSA in reading measures skills which "are at the heart of what students must be able to do to be good readers in today's society" (Assessment Handbook, 2008; p. 41). This assessment tool was selected as an outcome measure in part because it represents a high-stakes test of particular importance to school districts. Secondly, the PSSA is generally accepted as a valid and reliable measure of overall grade-level reading proficiency. Reliability coefficients are provided in the PSSA technical manuals produced by Data Recognition Corporation (DRC) each year (Mead & Melby, 2002; Mead & Melby, 2003; and Mead, Smith, & Swanlund, 2003). For the reading subtest, PSSA test-retest reliabilities range from 0.92 to 0.94. PSSA scores correlate positively and significantly with a variety of comparison reading achievement tests, including the Stanford Achievement Test, Terra Nova Comprehensive Test of Basic Skills, and the California Achievement Test-5. Correlations were typically from 0.7 to 0.9, indicating a high degree of similarity between the construct of reading ability measured by PSSA and other norm-referenced tests.

In Pennsylvania, the statewide reading proficiency test is designed to measure reading comprehension through two types of test items: selected response and constructed response (see Section 3.3.1 above for a further description of the question types on the PSSA). The constructed response test items are open-ended; they allow students to construct a written response in which they can reflect on what they have read, integrate background/prior knowledge with text-based information, expand meaning, and convey their ideas in written form. It is possible that the open-ended nature of this test question format may offer one possible explanation for why the DATMA—an open-ended definition task—was a stronger predictor of PSSA reading performance than was the static Test of Morphological Structure which requires a single word response to fit a sentence context as a measure of morphological knowledge. Moreover, it is not clear from the results of this study whether it was the dynamic nature of the DATMA that offered an advantage over the TMS as a predictor of future reading achievement; or, if the task required by the DATMA itself presented a more sensitive measure of morphological problem-solving than did the TMS.

Although this study was not designed to determine why a dynamic screening measure may be a more accurate predictor of later reading achievement than a static measure, the work by Vygotsky (1978) can inform this discussion. Vygotsky's idea of the zone of proximal distance defines that distance between the functioning level students demonstrate independently and the higher level at which they function with adult scaffolding. The dynamic measure utilized in this study can be considered within this framework. The students who benefited the most from the prompts/feedback might be the students who also benefit the most from classroom instruction. Therefore, the dynamic assessment wasn't just a measure of individual variations in a morphological analysis task; it was a measure of individual variations in a student's ability to respond to adult instruction.

In a comprehensive review of dynamic assessment, Grigorenko and Sternberg (1998) proposed that dynamic assessment taps student learning potential in a way that is distinct from static measures. Specifically, static measures typically assess already-developed abilities whereas dynamic measures are an indicator of a student's potential to learn new information. This research also adds to the existing dynamic assessment literature by providing initial support for its use as a screening measure for sixth grade struggling readers. The findings also converge with other research that has shown a multivariate screening approach may increase the accuracy of early identification (e.g., Catts et al., 2001; Compton et al., 2006; Foorman et al., 1998; O'Connor and Jenkins, 1999). As noted above, this research study found that a combination of two screening measures, the ORF and the DATMA, yielded acceptable sensitivity (.94) levels when predicting reading outcomes. However, the specificity levels associated with this combination model were lower than desired (.67). And while there are costs associated with false positives in an educational setting, these costs are less problematic than those associated with false negatives.

The use of a dynamic screening measure as part of universal screening might be particularly beneficial with older students. It is reasonable to suggest that a dynamic screening measure of morphological analysis skills might be particularly useful within an RTI framework. Most RTI models require students to remain in Tier 2 intervention for as many as 10-30 months before being considered a "non-responder." This means that a large part of the school year could pass without those students receiving individualized intervention (Caffrey, 2006; Caffrey, Fuchs, & Fuchs, 2008; Grigorenko, 2009).   A dynamic screening measure might serve to more quickly and/or accurately identify those students who will ultimately show poor response to Tier 2 intervention. If this is the case, students who perform poorly on a dynamic screening measure could receive individualized instruction more quickly, and thus, eliminate participating in many weeks of a Tier 2 intervention that might not be effective.

5.3 LIMITATIONS

This study has several limitations that must be addressed. The first limitation of this study is the small number of participants. Thus, while sufficient power was established for a simple linear regression analysis, multiple regression analyses should be considered exploratory and must be very cautiously interpreted due to the small sample size. A second limitation is the use of convenience sampling as opposed to random selection. The sample was taken from three classrooms across two schools in one school district. Both of these limitations decrease the generalizability and transferability of the results to the general population. Further, selection bias is also a concern. Although all of the sixth grade students in two of the district's schools were invited to participate in the study, participation was voluntary and the majority of the parents of these students did not give consent for participation. It is impossible to know if the students who did participate were somehow different from the students who did not participate.

It is important to note too that the participants in this study were from a district with a large percentage of children reading below grade level. In this study's sample, 67% of the sixth graders scored below proficiency on the statewide reading assessment and demonstrated reading difficulties. In addition, a significant number of children in the district were from families with incomes below the poverty line: approximately 49% of students in the district receive free or reduced-price lunch. The results of the current study may not hold for students coming from districts with higher reading scores and/or income levels.

In addition, the provision of intervention amongst the different schools (and classrooms) makes the interpretation of data concerning classification accuracy more difficult. As Good, Cummings, and Powell-Smith (2008) point out, intervention often improves the outcomes of at-risk children, and as a result, estimates of classification accuracy are compromised. However, this problem is unavoidable when conducting research in an educational setting. In this study, no information regarding specific classroom-based PSSA preparation was available to the researcher. Districts across Pennsylvania have implemented classroom instruction geared toward directly preparing students for the statewide accountability assessments. Future investigations could address this problem by obtaining a large enough sample to utilize multilevel modeling techniques, which could take classroom and school effects into account.

A second set of limitations is related to the design of the measure. It would have been very useful to know whether students knew the stem words used in the task. The design of this study also limits the conclusions that can be drawn about the importance of the dynamic aspect of the DATMA. The first concern is that an evaluation of the effectiveness of the prompting feedback was not conducted. We have no way to know if the prompts helped the children use morphological problem-solving to define the words. It is the possible that the pattern of responses could be an artifact of allowing multiple opportunities to give a definition.

There were several ways this study could have been improved. A longitudinal design that assessed the impact of morphological analysis skill on future vocabulary and reading growth would have added more theoretical importance to the study. Measures of vocabulary and reading proficiency were not designed specifically to contain morphologically complex words, which may have been a clearer test of the more direct role of morphological problem solving in reading. Moreover, only standardized tests were utilized as outcome measures. Good performance on such tests does not necessarily generalize to good performance in a classroom setting.

It is unlikely that one assessment tool or method will solely lead to the accurate identification of reading deficits. Identification will most likely require a battery of assessments and use of latent variable techniques so that effects due to measurement methods can be removed (Francis, Fletcher, Catts, et al., 2005). Therefore, future work in establishing the validity of the DATMA should be conducted with larger samples and more diverse measures. Latent variable models can then be used to evaluate the discriminant and convergent validity of the measure while controlling for effects due to method and test error.

In sum, future work with the DATMA needs to more comprehensively establish its predictive validity, explore the effectiveness of the graduated prompts, and more thoroughly address issues of construct validity. In addition, the amount of time and resources required for administering the test must be addressed. There is a balance that must be maintained between the extra information gained from the dynamic test and the resources required to administer it. Principally, the analyses in this study were exploratory with a small sample of sixth graders, and replication across a different, larger sample of students will render a more robust test of the validity and utility of the DATMA.

5.4 EDUCATIONAL IMPLICATIONS

Struggling readers in the district where this study was conducted continue to lose ground. Recent PSSA results for the district revealed that 54% of the district's sixth graders met grade level reading standards and only 52 % of seventh graders met standards in reading according to the statewide proficiency test. Although this is cross-sectional data, these scores suggest that the students in the district are losing ground in reading as they grow older. This suggests that the results from this investigation are educationally relevant. A clearer understanding of how the DATMA may predict PSSA performance can offer important information relevant to classroom instruction intended to prepare students for the statewide accountability test in reading. For example, further analysis of student response patterns on the DATMA may be useful in designing classroom-based PSSA preparation programs for students at-risk of scoring below proficiency. Based on this study's preliminary findings, for example, it may be fruitful for teachers to focus on the syntactic information provided by common derivational suffixes.

The findings reported in this study support results from previous research demonstrating the importance of morphological awareness for reading achievement in the upper elementary grades. The grade-level texts these children are required to read in school contain many unfamiliar morphologically complex words. Being able to read and derive meaning from these words is critical to vocabulary development and reading growth. Children who are able to use morphological problem-solving strategies to derive the meanings of unfamiliar words are at an advantage relative to children who do not use these strategies. Without morphological analysis skills for recognizing and making meaning from these words, children are at risk for continued problems with reading across the content areas.

Understanding how morphological analysis skill contributes to reading proficiency helps researchers design interventions to improve reading achievement. Those studies examining the effectiveness of instruction in morphological analysis have demonstrated success (e.g., Berninger, Nagy, Carlisle, Thomson, Hoffer, et al., 2003; Henry, 1989; Katz & Carlisle, 2002; Lovett, et al., 2000). These studies found that following instruction in morphological word reading strategies, students with and without reading difficulties showed improvements in deriving meaning from morphologically complex words. Katz and Carlisle taught morphological analysis strategies to fourth graders with language and learning difficulties that focused on both the recognition of the morphemic structure of words as well as the meanings of individual morphemes. Throughout and following the instruction, students showed significant growth in their abilities to derive meaning from morphologically complex words as well as in their reading fluency and reading comprehension. A recent meta-analysis reviewed studies with participants from preschool to eighth grade who received morphological interventions. Results indicate that morphological instruction benefits learners, and in particular, it brings significant benefits for less able readers (Bowers, Kirby, & Deacon, 2010) .These authors conclude that teaching morphological problem-solving skills is a valuable component of effective literacy instruction. The results of the present study provide additional support for these recommendations.

5.5 RECOMMENDATIONS FOR FURTHER RESEARCH

Given the findings from this study as well as evidence that instruction in morphological analysis is valuable for students with and without reading problems, further studies are certainly warranted. There are very few empirically tested measurement instruments for syntax and morphology for the population included in this study. This presents a limitation as well a recommendation for further study. Given this study's indication of the predictive importance of syntax and morphology, a valid and reliable measurement instrument takes on increased importance and should become a future focus of research.

Supportive evidence was generated that the DATMA is reliable and sensitive to differences in reading performance. This suggests that it might be appropriate for use in other studies of the influence of morphological analysis skills on reading achievement. The parameters used to develop items for the DATMA might also be used to generate additional items that would be appropriate for a wider age range of children or that cover a larger number of affixes.

While this study suggests that morphological problem-solving is associated with vocabulary and reading proficiency for upper elementary students, it does not address causal claims. Word learning is very complex. In terms of teaching words, a greater awareness of morphological structure may be an important component of instruction. More research is needed to determine whether such instruction improves students' learning of morphologically complex vocabulary or the comprehension of texts with a high proportion of morphologically complex words.

Finally, future research is needed with an increased number of participants and variables. In addition to increasing the sample size, it would be important to include a sample of more diverse participants. Because this study was conducted with children from a struggling school district, it is important to recognize that the findings may not hold for a different population of students. Therefore, it is recommended that this research continue in a different kind of school district with different kinds of children.

Although the results of this research are promising for the use of the DATMA as a screening measure, future research is warranted to both replicate and extend the results of this investigation. Specifically, predictive validity should be examined in a larger sample obtained from a more diverse population. Additionally, variables included in the analyses should be expanded to include a wider range of predictor variables, such as more general language or cognitive measures, as well as a wider range of outcome measures, including both timed and untimed measures of word identification and decoding.

Additionally, future research should explore the instructional implications of the use of dynamic assessment. As discussed by Sternberg and Grigorenko (2002), "in order for an educator to evaluate a student's ability to learn, the educator needs to teach students something and then observe their learning." This is essentially what occurs in a dynamic assessment. The prompting hierarchy provided in the DATMA may provide educators with ideas for instructional support necessary for a child to succeed. For example, one student might respond well to morpheme cues, whereas another student might need the addition of context clues in order to provide the correct definition. This type of information is important for planning supplemental instruction found in Tier 2 and Tier 3, and in fact, it is this type of information that more commonly utilized static screening measures are not able to provide. Further research could identify the prompts that are most salient to particular students and then further explore the utility of such prompts in a morphological analysis intervention study.

5.6 GENERAL CONCLUSIONS

This study represents a preliminary attempt at assessing morphological analysis skills with a dynamic measure and investigating the power of this instrument. Findings concerning the reliability and validity of the DATMA are encouraging. In this study, the dynamic measure was shown to have adequate internal consistency. When exploring the validity of the dynamic measure, findings show that the DATMA had a strong relationship to the TMS, a static measure of morphological awareness. Further, the DATMA had strong relationships with validated measures of vocabulary and reading. The DATMA explained significant variance in reading proficiency scores, suggesting it may be useful in finding students likely to be non-responders. In addition, although the classification example was exploratory, the pattern of results was interesting. The DATMA may be able to identify children across reading levels that show marked differences in their learning potential. An important strength of this study is that it adds new information to the current research in the area of prediction of chronic non-responders, that is, those presumed to have a reading disability. This is particularly true for the prediction of reading disabilities that are late identified or late-developing in older students.

This study found that morphological problem-solving as measured by the DATMA made a moderate contribution to the prediction of reading proficiency for sixth graders. According to Nagy and Anderson (1984), texts from grades 3-9 involved 139,020 transparent derived words such as *growth* compared to 49,080 semantically opaque words. The ability to recognize and estimate the meanings of these derived words from a known morpheme would apparently be an important contributor to reading fluency and comprehension of grade level texts. My study was specifically aimed at examining the components involved in reading proficiency for older students. It is likely that these components differ from those involved in reading proficiency in the earlier grades. For example, studies have shown that as students become older, phonological awareness becomes less predictive of reading proficiency when other variables are considered (Aarnoutse, et al., 2001; Roth, et al., 2002). Furthermore, research has shown that as students become older, reading demands change. When student enter the upper-elementary grades, two-thirds of the unfamiliar words in the texts they are required to read are morphologically complex (Nagy & Anderson, 1984). It was reasonable to expect that morphology would play a role in reading proficiency for sixth graders.

Constructing reliable and valid tests for the identification of children who struggle with reading will require a new consideration of how tests should be constructed and what they should measure. The DATMA, used in combination with other tests, may be helpful in differentiating upper elementary and middle school struggling readers. Although promising, more work will need to be conducted to determine the measure's predictive power, to isolate and adequately capture children's responsiveness to the instruction provided, and to determine the measure's relative utility among other tests of reading skill before any definitive recommendations can be made regarding its use.

Dynamic assessment poses an alternative method to the static procedures of standardized testing by assessing a child's ability to learn with guided instruction. Dynamic assessment gives the information that a standardized test gives plus a description of how a child actually learns in his/her academic environment. Thus, results of this type of assessment may be more accurate and readily transferable to intervention. The current study shows that dynamic assessment provided accurate results concerning reading related skills in sixth graders as compared to standardized tests. Further analysis of participant responses suggest that dynamic assessment may be used to obtain insight into the learning characteristics of a child. Thus, this research adds to the existing dynamic assessment literature by providing initial support for its use as a screening measure for students in sixth grade.

The findings from this study may be used to ask questions about the extent to which morphological analysis strategies can be taught to students as well as the extent to which this instruction affects reading achievement. It may be that the morphological structure of words needs to be made obvious to children. Further, it appears as though the understanding of both the semantic and syntactic aspects of morphology should be considered in designing instructional programs for future intervention studies. Instruction that helps students attend to units of meaning may expand their lexicon and therefore improve their ability to gain meaning from text. Currently no studies are available detailing the prevalence and quality of morphological instruction, yet instruction in units of meaning could be embedded within classroom teaching across grade levels, allowing students to focus on meaning within text. Prior research has demonstrated that for students in the intermediate grades, awareness of units of meaning and word structure play a critical role in literacy achievement (Anglin, 1993; Carlisle, 1995, 2000; Carlisle & Nomanbhoy, 1993, Deacon & Kirby, 2004; Katz & Carlisle, 2002; Kieffer & Lesaux, 2008; Leong, 2000, Nagy et al., 2006). Based on the preponderance of evidence, it would seem difficult to overestimate the importance of morphology in later reading development.

APPENDIX A

CONSENT FOR A CHILD TO PARTICIPATE IN A RESEARCH STUDY

Differentiating Among Students: The Value Added of a Dynamic
Assessment of Morphological Problem Solving

My name is Kathleen Stanfa and I am a student in the Ph.D. Program in Special Education at the University of Pittsburgh. As part of my dissertation, I am conducting a study which is designed to determine if an assessment method is useful in measuring students' word learning skills. I am interested in your child's participation in this study.

Participation is voluntary and there are no negative consequences for withdrawing at any time or declining to participate. Participation, refusal to participate, or withdrawal at any time will not affect your relationship with the University of Pittsburgh, your child's school, or your child's teacher or their grades. Subjects who complete the study will be compensated for their time and inconvenience with a $5 gift certificate. Subjects who consent but do not complete the study receive an incentive in the form of a $2 gift certificate.

The study is designed to benefit participants by providing information about their vocabulary and reading comprehension skills. Any possible benefits are not guaranteed. There are no anticipated risks involved in this study with the exception of a breach of confidentiality which exists as a risk in all data collection. There will be approximately 30 students participating. The assessment will consist of two sessions of approximately 30 minutes duration each and will be administered in an empty classroom.

Both students with and without a learning disability are eligible to take part in the study. Each student will be told about the study before being asked to participate. Study participants will be given a number of tasks designed to assess vocabulary and reading-related skills during the sessions. Each student will be given a list of words to define. If the student is unable to define a word, he/she will be provided with clues to break down the unfamiliar words in order to determine their meanings. A brief oral reading test will be given as well as a standardized measure of vocabulary knowledge. In addition, reading subtest scores from the 2010 PSSA will be obtained for each participant.

At the end of this study, any records that personally identify your child will remain stored in locked files and will be kept for a minimum of seven years. Your child's identity will not be revealed in any description or publications of this research. In unusual cases, your child's research records may be released in response to an order from a court of law. It is also possible that authorized representatives from the University of Pittsburgh Research Conduct and Compliance Office may review your child's data for the purpose of monitoring the conduct of this study. Also, if the investigators learn that your child or someone with whom your child is involved is in serious danger or potential harm, they will need to inform the appropriate agencies, as required by Pennsylvania law. Results from the assessments administered during the course of this study will be available to you in the future upon your request. If you so wish, information can be shared with your child's teacher. If you have any questions about this study please call me at (412) 648-3137 or contact me at 5157 Posvar Hall, University of Pittsburgh, Pittsburgh, PA 15260. I can also be reached by email at kmm78@pitt.edu.

If you would like your child to participate please sign and return one copy of this consent form to your child's teacher in the enclosed envelope. The second copy of this consent form is for you to keep for your records. Thank you for your consideration.

VOLUNTARY CONSENT/ PARENTAL CERTIFICATION

The above information has been explained to me and all of my current questions have been answered. I understand that I am encouraged to ask questions about any aspect of this research study during the course of this study, and that such future questions will be answered by a qualified individual or by the investigator(s)at the telephone number(s) given. I understand that I may always request that my questions, concerns or complaints be addressed by a listed investigator.

I understand that I may contact the Human Subjects Protection Advocate of the IRB Office, University of Pittsburgh (1-866-212-2668) to discuss problems, concerns, and questions; obtain information; offer input; or discuss situations in the event that the research team is unavailable.

By signing this form, I agree for my child to participate in this research study. A copy of this consent form will be given to me/my child.

____
Printed Name of Child-Subject

"I understand that, as a minor (age less than 18 years), the above-named child is not permitted to participate in this research study without my consent. Therefore, by signing this form, I give my consent for his/her participation in this research study."

______________________________ ______________________________
Parent's Name (Print) Relationship to Participant (Child)

______________________________ _______________
Parent Signature Date

CHILD ASSENT (to be used with children who are developmentally able to sign)

This research has been explained to me, and I agree to participate.

_________________________________ _______________
Signature of Child-Subject Date

____
Printed Name of Child-Subject

CERTIFICATION of INFORMED CONSENT

I certify that I have explained the nature and purpose of this research study to the above-named individual(s), and I have discussed the potential benefits and possible risks of study participation. Any questions the individual(s) have about this study have been answered, and we will always be available to address future questions as they arise. I further certify that no research component of this protocol was begun until after this consent form was signed.

____________________________________ _________________________
Printed Name of Person Obtaining Consent Role in Research Study

__________________________________ _____________
Signature of Person Obtaining Consent Date

APPENDIX B

DEFINITIONS, FREQUENCY VALUES, AND SENTENCES

Dictionary definitions were obtained from the *Oxford English Dictionary* (Simpson & Weiner, 1989). Frequencies expressed as tokens per million words and obtained from the *Corpus of Contemporary American English* (COCA; Davies, 2008-).

***Beastly** (adj.): of or pertaining to the lower animals, merely animal; bestial

Derived form frequency: .34

Root frequency (beast): 11.07

Jan's brother tried to scare her by dressing up as a bear and acting *beastly.*

a) like an animal

b) like a plant

c) like a clown

Boathouse (n.): a building or shed, usually built partly over water, for sheltering a boat or boats

Derived form frequency: .42

Root frequency (boat): 65.22 (house): 550.26

We returned to the *boathouse* after a long day on the water.

a) a building in which boats are kept

b) a boat equipped for living in

c) a room next to a swimming pool

***Celebratory** (adj.): the quality of celebrating

Derived form frequency: 1.92

Root frequency (celebrate): 21.17

After winning the prize, Alex was in a *celebratory* mood.

a) ready to party

b) ready to sleep

c) ready to study

Cookery (n.): the art or practice of cooking; a place or area for cooking.

Derived form frequency: .36

Root frequency (cook): 84.39

Grandma's *cookery* was awful so Bob hated eating at her house.

a) the art of decorating

b) the art of preparing food

c) the art of sewing

***Corrective** (adj.): having the property or function of correcting or setting right what is erroneous or faulty, or of producing amendment; tending to correct

Derived form frequency: 2.00

Root frequency (correct): 54.03

Tracy's braces were *corrective*, so her teeth would not be crooked.

a) tending to clean

b) tending to damage

c) tending to fix

***Dramatize** (v.): to convert into a drama; to put into dramatic form; adapt for representation on the stage

Derived form frequency: .67

Root frequency (drama): 28.40

For the school play, the fourth graders decided to *dramatize* the book, Harry Potter and the Sorcerer's Stone.

a) to make into a performance

b) to make into a painting

c) to make into a book report

Equalize (v.): to equal; match

Derived form frequency: .53

Root frequency (equal): 40.90

At the track meet, Josh ran extra fast so he could *equalize* his time with his opponent's.

a) to increase

b) to match

c) to decrease

Fearsome (adj.): fear inspiring; frightful; dreadful

Derived form frequency: 1.90

Root frequency (fear): 99.64

The movie was so *fearsome* that John had to cover his eyes most of the time.

a) funny

b) romantic

c) frightening

****Flowery** (adj.): abounding in or covered with flowers; producing flowers

Derived form frequency: 1.21

Root frequency (flower): 17.98

Molly hated the dress with daisies because it was too *flowery*.

a) covered with blossoms

b) covered with dots

c) covered with colors

Hospitalize (v.): to place in a hospital for medical care or observation

Derived form frequency: .07

Root word frequency (hospital): 112.06

The doctor was unsure whether to *hospitalize* the patient or not.

a) place in the hospital

b) treat kindly

c) enjoy

Idealize (v.): to make ideal; to represent something in an ideal form

Derived form frequency: .18

Root word frequency (ideal): 30.56

People often *idealize* movie stars as perfect.

a) create

b) destroy

c) put up on a pedestal

Moonlit (adj.): lit by the moon

Derived form frequency: 1.02

Root word frequency (moon): 54.90 (lit): 22.20

The *moonlit* lake glistened under the stars.

a) round

b) lit by the moon

c) shiny

Oddity (n.): the quality or character of being odd or peculiar; peculiarity; strangeness; singularity

Derived form frequency: .90

Root frequency (odd): 30.12

The singing horse with a purpose tail was an *oddity* they saw at the fair.

a) a common thing

b) an unusual thing

c) a pretty thing

Odorous (adj.): emitting a smell or scent; scented, odiferous

Derived form frequency: .16

Root frequency (odor): 7.31

Mary hated cleaning the litter box out because it was so *odorous*.

a) large

b) boring

c) strong smelling

****Organist** (n.): a person who plays the organ

Derived form frequency: .48

Root work frequency (organ): 11.02

The band hired a new *organist* to perform in their next concert.

a) hero

b) singer

c) musician

Preventive (adj.): that anticipated in order to ward against; precautionary; that keeps from coming or taking place; that acts as hindrance or obstacle

Derived form frequency: 1.06

Root frequency (prevent): 58.44

The rain coats were *preventive* and kept the children from getting all wet.

a) tending to make something happen

b) tending to make into a game

c) tending to keep something from happening

Puzzlement (n.): the fact or condition of being puzzled; perplexity; bewilderment; confusion

Derived form frequency: 1.22

Root frequency (puzzle): 11.89

Kim tried to figure out how to put the bike together and finally gave up in *puzzlement*.

a) confusion

b) happiness

c) fright

Reliance (n.): confident or trustful dependence; something or someone relied on

Derived form frequency: 6.80

Root word frequency (rely): 23.64

Her *reliance* on others to do her work was in large part caused by laziness.

a) dependence

b) failure

c) cause

Secretive (adj.): inclined to, fond of secrecy; very reticent; indicative of secrecy

Derived form frequency: 3.32

Root frequency (secret): 73.12

Tim had to be very *secretive* so Gail wouldn't find out about the surprise party.

a) silly

b) mysterious

c) angry

Sparkly (adj.): sparkling, something that sparkles

Derived form frequency: 1.19

Root frequency (sparkle): 2.61

Elizabeth got to wear a *sparkly* crown on her birthday.

a) dull

b) silky

c) glittery

***Spoilage** (n.): the state of spoiling

Derived form frequency: .41

Root frequency (spoil): 2.83

The food was kept cold to prevent *spoilage*.

a) freezing

b) rotting

c) stealing

Stardom (n.): the status of a performer or entertainer acknowledged as a star; star performers considered as a group

Derived form frequency: 1.91

Root word frequency (star): 112.82

David came close to getting the lead in several movies, but he missed his chance for *stardom*.

a) modesty

b) fame

c) luck

*piloted item
**practice item

APPENDIX C

ORAL READING FLUENCY PASSAGES – SIXTH GRADE

Oral Reading Fluency Passage #1

What is one of the world's biggest holes in the ground? It's more than a mile deep and almost three hundred miles long, and more than ten miles wide at some points. If you guessed the Grand Canyon, you're right. The Grand Canyon is an enormous gorge carved over millions of years by the Colorado River in northwestern Arizona.

Among the world's great tourist attractions, the Grand Canyon is walled by colorful strata, or layers, of rock dating back millions of years. The reds, pinks, and yellows in the rock are the result of traces of different minerals.

Most tourists visit the South Rim of the canyon, where there are hotels and many trails to explore. Bright Angel Trail is a popular hiking trail. The South Rim is open year round to visitors. The North Rim is cooler and quieter than the South Rim but is open only six months of the year.

The only ways to reach the inner canyon are by foot, on mule, or by raft on the Colorado River. Visitors can take daylong raft trips over smooth water or weeklong trips that include rolling rapids. Almost two hundred years ago, American John Wesley Powell led the first successful trip through the canyon. He and ten other men traveled down the river in four small boats, braving waters that had never been mapped.

Native Americans were the first to live and work in the canyon, more than eight hundred years ago. They lived in rock pueblos on both rims of the canyon, hunting and fishing, growing crops, making pottery, and weaving baskets.

Wildlife is abundant in the canyon. Hundreds of kinds of birds live there, as well as bighorn sheep, mule deer, beavers, bats, snakes, lizards, and frogs. There are also many types of trees, cacti, and wildflowers.

You can see that the Grand Canyon is much more than just a big hole in the ground. It is an amazing site, alive with stories of the past and present that are written on the rock, on the land, and on the river.

Oral Reading Fluency Passage #2

If you visit Fossil Butte National Monument in Wyoming, you may meet a fish that is fifty million years old. The park was established to preserve the rock formations that contain a wide variety of fossil remains of plants and animals from a lake that covered the area long ago. The fossils are so well preserved that scientists can use them to study relationships among the plants and animals. Scientists can also use the fossils to study the effects of climate change in the area.

Fossil Lake, now a dry bed, was once more than fifty miles long and twenty miles wide at its maximum. The lake and its surrounding area were alive with gars, stingrays, herring, perch, crocodiles, turtles, insects, and horses the size of dogs.

Scientists aren't sure why so much of the lake's life was preserved as fossils. One theory is that plants and animals that sank to the bottom of the lake were quickly covered with a substance in the water that protected them. Scientists can also tell that a great number of fish were killed suddenly, but no one has yet solved this mystery.

Fossil Butte is a high desert, with hot, sunny summers and cool nights and cold winters. It usually has perfect weather for hiking. It is easy to get out of breath as you hike, though, because it is so far above sea level. You're likely to see mule deer and a variety of birds in the park. If you're lucky, you might also see elk, moose, and beaver.

You can explore the park on your own or with a ranger. There are two groomed hiking trails in the park. A quarry, located on one of the trails, is open to the public. Here visitors can help the park staff excavate fossils. And here is where you might meet your fish, its skeleton, teeth, scales, and skin perfectly preserved and ready to tell you an ancient story.

Oral Reading Fluency Passage #3

Tennis is one of the most popular sports in the world. If you've discovered the game of tennis and enjoy playing, you probably know the basics by now, such as the rules of play, ways

to serve and return the ball, and, of course, tennis etiquette, or the good manners of tennis. Three other things that are just as important are warming up before and after playing, doing practice exercises, and cross training.

Stretching to warm up helps loosen the muscles of your body, which helps you avoid injuries during a tennis game. One example of a good stretching exercise is to sit on the ground with your legs straight out in front of you. While trying to keep the backs of your knees on the ground, reach for your toes or as lose to your toes as possible. Hold this stretch for about twenty seconds, relax, and then repeat the stretch three more times.

Practice exercises will help you improve your speed and reaction time for tennis. An example of a good practice exercise is to stand a few feet away from a partner and throw each other a tennis ball at the same time. You have to aim carefully and throw underarm for this exercise. If this seems too easy, try clapping or turning around before you catch the ball.

Cross training has to do with engaging in a sport other than your favorite sport. The theory behind cross training is that participation in other sports helps tune your body for your main sport.

For example, if you're a tennis player, playing soccer or basketball can improve your coordination and strengthen your leg muscles for moving quickly around the tennis court. Swimming and jogging can increase your endurance so that you don't tire out too quickly in a tennis match.

Another benefit of cross training is that it puts some variety in your sports life. If you focus all of your attention on tennis, you risk tiring of the sport too quickly. Variety in training will help you maintain your excitement about tennis or any other sport.

APPENDIX D

TEST OF MORPHOLOGICAL STRUCTURE

Part 1: Derivation

Practice a. Farm. My uncle is a _________________________. [farmer]
b. Help. My sister is always ___________________. [helpful]

1. warm. He chose the jacket for its ____________________. [warmth]
2. teach. He was a very good __________________________. [teacher]
3. permit. Father refused to give _______________________. [permission]
4. profit. Selling lemonade in summer is ________________. [profitable]
5. appear. He cared about his __________________________. [appearance]
6. express. ‘OK’ is a common _________________________. [expression]
7. four. The cyclist came in __________________________. [fourth]
8. remark. The speed of the car was ____________________. [remarkable]
9. protect. She wore glasses for ________________________. [protection]
10. perform. Tonight is the last _________________________. [performance]
11. expand. The company planned an ___________________. [expansion]
12. revise. This paper is his second ___________________. [revision]
13. reason. Her argument was quite ____________________. [reasonable]
14. major. He won the vote by a ________________________. [majority]
15. deep. The lake was well known for its _______________. [depth]
16. equal. Boys and girls are treated with ________________. [equality]
17. long. They measured the ladder’s ___________________. [length]
18. adventure. The trip sounded ________________________. [adventurous]
19. absorb. She chose the sponge for its _________________. [absorption]
20. active. He tired after so much _______________________. [activity]
21. swim. She was a strong __________________________. [swimmer]
22. human. The kind man was known for his _____________. [humanity]
23. wash. Put the laundry in the ________________________. [washer]
24. humor. The story was quite _________________________. [humorous]
25. assist. The teacher will give you ____________________. [assistance]
26. mystery. The dark glasses made the man look _________. [mysterious]

27. produce. The play was a grand ____________________. [production]
28. glory. The view from the hill top was _______________. [glorious]

Part 2: Decomposition

Practice: a. Driver. Children are too young to ___________________. [drive]
b. Improvement. My teacher wants my spelling to ________. [improve]

1. growth. She wanted her plant to ____________________. [grow]
2. dryer. Put the wash out to _________________________. [dry]
3. variable. The time of his arrival did not _______________. [vary]
4. width. The mouth of the river is very ________________. [wide]
5. density. The smoke in the room was very _____________. [dense]
6. discussion. The friends have a lot to _________________. [discuss]
7. famous. The actor would achieve much _______________. [fame]
8. description. The picture is hard to ___________________. [describe]
9. fifth. The boy counted from one to ___________________. [five]
10. election. Which person did they ___________________? [elect]
11. strength. The girl was very _________________________. [strong]
12. decision. The boy found it hard to __________________. [decide]
13. popularity. The girl wants to be ___________________. [popular]
14. runner. How fast can she __________________________? [run]
15. publicity. His views were made ___________________. [public]
16. difference. Do their opinions _______________________? [differ]
17. originality. That painting is very ___________________. [original]
18. agreeable. With that statement I could not ____________. [agree]
19. courageous. The man showed great _________________. [courage]
20. admission. How many people will they _____________? [admit]
21. dangerous. Are the children in any __________________? [danger]
22. reduction. The overweight man was trying to _________. [reduce]
23. baker. She put the bread in to ______________________. [bake]
24. division. The cake is hard to _______________________. [divide]
25. guidance. The map was her ________________________. [guide]
26. continuous. How long will the storm ________________? [continue]
27. reliable. On his friend he could always ______________. [rely]
28. acceptance. Is that an offer you can _________________? [accept]

BIBLIOGRAPHY

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*: MIT Press.

Anderson, R. C., & Freebody, P. (1983). Reading comprehension and the assessment and acquisition of word knowledge. In B. Hutson (Ed.), *Advances in Reading/Language Research* (pp. 231-256). Greenwich, CT: JAI Press.

Anderson, R. C., & Nagy, W. E. (1992). The vocabulary conundrum. *American Educator: The Professional Journal of the American Federation of Teachers, 16*, 14-18, 44-47.

Anglin, J. M. (1993). Vocabulary development: A morphological analysis. *Monographs of the Society for Research in Child Development, 58*, 1-166.

Arnbak, E., & Elbro, C. (2000). The effects of morphological awareness Training on the reading and spelling skills of young dyslexics. *Scandinavian Journal of Educational Research, 44*, 229-251.

Bain, B.A. & Olswang, L.B. (1995). Examining readiness for learning two-word utterances by children with specific expressive language impairment: Dynamic assessment validation. *American Journal of Speech-Language Pathology, 4*, 81-92.

Baumann, J. F., Edwards, E. C., Boland, E. M., Olejnik, S., & Kame'enui, E. J. (2003). Vocabulary tricks: Effects of instruction in morphology and context on fifth-grade students' ability to derive and infer word meanings. *American Educational Research Journal, 40*, 447-494.

Baumann, J. F., Edwards, E. C., Font, G., Tereshinski, C. A., Kame'enui, E. J., & Olejnik, S. (2002). Teaching morphemic and contextual analysis to fifth-grade students. *Reading Research Quarterly, 37*(2), 150-176.

Baumann, J. F., & Kame'enui, E. J. (2003). Research on vocabulary instruction: Ode to Voltaire. In J. Flood, D. Lapp, J. R. Squire, & J. M. Jensen (Eds.), *Handbook on research teaching the English language arts* (2nd ed., pp. 604-632). Mahwah, NJ: Erlbaum.

Baumann, J. F., & Kame'enui, E. J. (2004). *Vocabulary instruction: Research to practice*. The Guilford Press: New York.

Baumann, J., Kame'emui, E., & Ash, G. (2003). Research on vocabulary instruction: Voltaire Redux. In J. Flood, D. Lapp, J. R. Squire, & J. M. Jensen (Eds.), *Handbook on research teaching the English language arts*. (2nd ed., pp. 752-785). Mahwah, NJ: Erlbaum.

Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing Words to Life: Robust Vocabulary Instruction. Solving Problems in the Teaching of Literacy*: Guilford.

Beck, I. L., McKeown, M.G., & McCaslin, E. S. (1983). Vocabulary development: All contexts are not created equal. *The Elementary School Journal, 88*, 177-183.

Berk, L. E. & Winsler, A. (1995). *Scaffolding children's learning : Vygotsky and early childhood education.* Washington, DC : National Association for the Education of Young Children.

Berko, J. (1958). The child's learning of English morphology. *Word, 14*, 150-177.

Biancarosa, G., & Snow, C. (2004). *Reading next: A vision for action and research in middle and high school literacy—A report to Carnegie Corporation of New York*. Washington, D. C.: Alliance for Excellent Education.

Bloomfield, L. (1933). *Language*. New York: Holt.

Bowey, J. A. (2001). Nonword repetition and young children's receptive vocabulary: A longitudinal study. *Applied Psycholinguistics, 22*, 441-469.

Brittain, M. M. (1970). Inflectional performance and early reading achievement. *Reading Research Quarterly, 6*, 34-48.

Bryant, N. R., Brown, A. L., & Campione, J. C. (1983, April). Preschool children's learning and transfer of matrices problems: Potential for improvement. Paper presented at the meeting of the Society for Research in Child Development, Detroit.

Byrne, B., Fielding-Barnsley, R., & Ashley, L. (2000). Effects of preschool phoneme identity training after six years: Outcome level distinguished from rate of response. *Journal of Educational Psychology, 92,* 659–667.

Budoff, M. (1987). The validity of learning potential assessment. In C.S. Lidz (Ed.). *Dynamic assessment: An interactional approach to evaluating learning potential* (pp.52-81). New York: Guilford.

Bulmer, M. G. (1965). *Principles of statistics*, 2nd edition. New York: Dover Publications.

Burani, C., & Caramazza, A. (1987). Representation and processing of derived words. *Language and Cognitive Processes, 3*, 217-227

Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes, 10*, 425-455.

Caffrey, E., Fuchs, D., & Fuchs, L.S. (2008). The predictive validity of dynamic assessment: A review. *Journal of Special Education, 41*, 254-270.

Campione, J.C. & Brown, A.L. (1987). Linking dynamic assessment with school achievement. In C.S. Lidz (Ed.). *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 82-115). New York: Guilford.

Campione, J.C., Brown, A.L., & Bryant, N.R. (1985). Individual differences in learning and memory. In R.J. Sternberg (Ed.). *Human abilities: An information-processing approach* (pp. 103-126). New York: W.H. Freeman.

Carlisle, J. F. (1995). Morphological awareness and early reading achievement. In L. B. -Feldman (Ed.), *Morphological aspects of language processing* (pp. 189-209). Hillsdale, NJ: Erlbaum.

Carlisle, J. F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing, 12,* 169-190.

Carlisle, J. F. (2003). Morphology matters in learning to read: A commentary. *Reading Psychology, 24*, 291-322.

Carlisle, J. F., & Flemming, J. (2003). Lexical processing of morphologically complex words in the elementary years. *Scientific Studies of Reading, 7*, 239-253.

Carlisle, J. F., & Nomanbhoy, D. (1993). Phonological and morphological awareness in first graders. *Applied Psycholinguistics, 14*, 117-195.

Carlisle, J. F., & Stone, C. A. (2005). Exploring the role of morphemes in word reading. *Reading Research Quarterly, 40*, 428-449.

Carlisle, J. F., Stone, C. A., & Katz, L. A. (2001). The effects of phonological transparency on reading derived words. *Annals of Dyslexia, 51*, 249-274.

Carlo, M. S., August, D., McLaughlin, B., Snow, C. E., Dressier, C., Lippman, D. N., et al. (2004). Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classrooms. *Reading Research Quarterly, 39*, 188-215.

Carroll, J. B., Davies, P., & Richman, B. (1971). *The American heritage word frequency book*. Boston: Houghton-Mifflin.

Catts, H. W., & Kamhi, A. G. (1999). *Language and language disorders*. Boston: Allyn & Bacon.

Champion, A. (1997). Knowledge of suffixed words: A comparison of reading disabled and nondisabled readers. *Annals of Dyslexia, 47*, 29-55.

Chomsky, C. (1970). Reading, writing, and phonology. *Harvard Educational Review 40*, 287-309.

Cole, P., Beauvillain, C. & Segui, J. (1989). On the representation and processing of prefixed and suffixed derived words: A differential frequency effect. *Journal of Memory and Language 28*, 1-13.

Compton, D. L., Fuchs, D., & Fuchs, L. S. (2007). *The course of reading and mathematics disability in first grade: Identifying latent class trajectories and early predictors*. Manuscript submitted for publication.

Cunningham, A. E., & Stanovich, K. E. (1991). Tracking the unique effects of print exposure in children: Associations with vocabulary, general knowledge, and spelling. *Journal of Educational Psychology, 83*, 264-274.

Deacon S. H., & Kirby, J.R. (2004). Morphological awareness: Just "more phonological"? The roles of morphological awareness and phonological awareness in reading development. *Applied Psycholinguistics, 25*,223-238.

Derwing, B. L. (1976). Morpheme recognition and the learning of rules for derivational morphology. *Canadian Journal of Linguistics, 21*, 38-66.

Derwing, B.L., Smith, M.L., & Wiebe, G.E. (1995). On the role of spelling in morpheme recognition: Experimental studies with children and adults. In L.B. Feldman (Ed.), *Morphological Aspects of Language Processing* (pp. 3-27). Hillsdale, NJ: Erlbaum.

Deshler, D. D., Schumaker, J. B., Bui, Y. N., & Vernon, S. (2006). High schools and adolescents with disabilities: Challenges at every turn. In D. D. Deshler & J. B. Schumaker (Eds.), *High school students with disabilities: Strategies for accessing the curriculum* (pp. 1-34). New York: Corwin Press.

Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test-Revisited*. Circle Pines, MN: American Guidance Service.

Dunn, L., & Dunn, L. (1997*). Peabody Picture Vocabulary Test-Third Edition*. Circle Pines, MN: American Guidance Service.

Dunn, L. M., Dunn, D. M. (2007). Peabody picture vocabulary test (4th ed.). Bloomington, MN: NCS Pearson, Inc.

Duques, S. (1989). Grammatical deficiency in writing: An investigation of learning disabled college students. *Reading and Writing, 2*, 1-17.

Ebbers, S. M., & Denton, C. A. (2008). A root awakening: Vocabulary instruction for older students with reading difficulties. *Learning Disabilities Research & Practice, 23*, 90-102.

Elbro, C. (1989). *Morphological awareness in dyslexia*. Copenhagen: University of Copenhagen, Institute for General and Applied Linguistics.

Elbro, C. (1990). *Differences in dyslexia. A study of reading strategies and deficits in a linguistic perspective*. Copenhagen: Munksgaard.

Feldman, L. B., & Bentin, S. (1994). Morphological analysis of disrupted morphemes: Evidence from Hebrew. The Quarterly Journal of Experimental Psychology, 47, 407- 435.

Feldman, L. B., Rueckl, J., Diliberto, K. A., Pastizzo, M., & Vellutino, F. R. (2002). Morphological analysis by child readers as revealed by the fragment completion task. *Psychonomic Bulletin & Review, 9*, 529-535.

Ferrara, R.A., Brown, A.L., & Campione, J.C. (1986). Children's learning and transfer of inductive reasoning rules: Studies of proximal development. *Child Development, 57*, 1087-1099.

Feuerstein, R. (1979). *The dynamic assessment of retarded performers: The learning potential assessment device, theory, instruments, and techniques*. Baltimore: University Park Press.

Foorman, B. R., Fletcher, J. M., Francis, D. J., Carlson, C. D., Chen, D., Mouzake, A., et al. (1998). *Technical report: Texas Primary Reading Inventory* (1998 Edition). Houston: Center for Academic and Reading Skills, University of Texas Health Science Center at Houston and University of Houston.

Fowler A. & Liberman, I. Y. (1995). The role of phonology and orthography in morphological awareness. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 157-188). Hillsdale, NJ: Lawrence Erlbaum.

Fowler, A., Napps, S. E., & Feldman, L. B. (1985). Relations among regular and irregular morphological related words in the lexicon as revealed by repetition priming. *Memory and Cognition, 13*, 241-255.

Freyd, P., & Baron, J. (1982). Individual differences in acquisition of derivational morphology. *Journal of Verbal Learning and Verbal Behavior, 21*, 282-295.

Fuchs, D., Fuchs, L. S., & Compton, D. L. (2004). Identifying reading disability by responsiveness-to-instruction: Specifying measures and criteria. *Learning Disability Quarterly, 27*, 216–227.

Fuchs, L. S., Compton, D. L., Fuchs, D., Hollenbeck, K. N., Craddock, & C. F., Hamlett, C. L. (2008). Dynamic Assessment of Algebraic Learning in Predicting Third Graders' Development of Mathematical Problem Solving. *Journal of Educational Psychology, 100*, 829–850.

Fuchs, L., Fuchs, D., Hosp, M., & Jenkins, J. (2001). Oral reading fluency as an indicator of reading competence: a theoretical, empirical and historical analysis. *Scientific Studies of Reading, 5*, 239-256.

Gathercole, S. E., & et al. (1992). Phonological Memory and Vocabulary Development during the Early School Years: A Longitudinal Study. *Developmental Psychology, 28*, 887-898.

Gathercole, V. M., Sebastian, E., & Soto, P. (1999). The Early Acquisition of Spanish Verbal Morphology: Across-the-Board or Piecemeal Knowledge? *International Journal of Bilingualism, 3*, 133-182.

Graves, M. F. (1986). Vocabulary learning and instruction. *Review of Research in Education, 13*, 49-90.

Graves, M. F. (2006). *The vocabulary book: Learning and instruction*. New York: Teachers College Press.

Graves, M. F., & Hammond, H. K. (1980). Vocabulary learning and instruction. In E. Z. Rothkopf (Ed.), *Review of research in education* (vol. 13, pp. 49-89). Washington, DC: American Research Association.

Gray, S., Plante, E., Vance, R., & Henrichsen, M. (1999). The diagnostic accuracy of four vocabulary tests administered in preschool-age children. *Language, Speech, and Hearing Services in Schools, 30*, 196-206.

Grigg, W., Donahue, P., & Dion, G. (2007). *The nation's report card: 12th-grade reading and mathematics 2005* (NCES 2007-468). U.S. Department of Education, National Center for Education Statistics. Washington, D. C.: U.S. Government Printing Office.

Gutierrez-Clellen, V. F., Conboy, B., Brown, S., & Robinson-Zanartu, C. (1998). Modifiability: A dynamic approach to assessing immediate language change. *Journal of Children's Communication Development, 19*, 31-42.

Henderson, A. J., & Shores, R. E. (1982). How learning disabled student's failure to attend to suffixes affects their oral reading performance. *Journal of Learning Disabilities, 15*, 178-182.

Henry, M. K. (1993). Morphological structure: Latin and Greek roots and affixes as upper grade code strategies. *Reading and Writing: An Interdisciplinary Journal, 5*, 227-241.

Jarmulowicz, L., Hay, S. E., Taran, V. L., & Ethington, C. A. (2007). Fitting derivational morphophonology into a developmental model of reading. *Reading and Writing: An Interdisciplinary Journal, 21*, 275-297.

Jarmulowicz, L., Taran, V. L., & Hay, S. E. (2008). Third graders' metalinguistic skills, reading skills, and stress production in derived English words. *Journal of Speech, Language, and Hearing Research, 50*, 1593-1605.

Jenkins, J. R. (2003, December). *Candidate measures for screening at-risk students*. Paper presented at the National Research Center on Learning Disabilties Responsiveness-to-Intervention symposium, Kansas City, MO. Retrieved December 7, 2010 from http://www.nrcld.org/symposium2003/jenkins/index.html

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80*, 437-447.

Kieffer, M. J., & Lesaux, N. K. (2008). The role of derivational morphology in the reading comprehension of Spanish-speaking English Language Learners. *Reading and Writing: An Interdisciplinary Journal, 21*, 783-804.

Kletzien, S. B., & Bednar, M. R. (1990). Dynamic assessment for at-risk readers. *Journal of Reading, 33*, 528-555.

Koger, M. E., Thacker, A. A., & Dickinson, E. R. (2004). *Relationships Among the Pennsylvania System of School Assessment (PSSA) Scores, SAT Scores, and Self-Reported High School Grades for the Classes of 2002 and 2003* (HumRRO Report No. FR-04-26). Alexandria, VA: Human Resources Research Organization.

Ku, Y.-M., & Anderson, R. C. (2003). Development of morphological awareness in Chinese and English. *Reading and Writing: An Interdisciplinary Journal, 16*, 399-422.

Larsen, J. A., & Nippold, M. A. (2007). Morphological analysis in school-age children: A dynamic assessment of a word learning strategy. *Language, Speech, & Hearing Services in Schools, 38,* 201-212.

Leong, C. K. (1989a). The effects of morphological structure on reading proficiency: A developmental study. *Reading and Writing: An Interdisciplinary Journal, 12*, 277-302.

Leong, C. K. (1989b). Productive knowledge of derivational rules in poor readers. *Annals of Dyslexia, 39*, 94-115.

Leong, C. K. (2000). Rapid processing of base and derived forms of words and grades 4, 5 and 6 children's spelling. *Reading and Writing, 12*, 277-302(226).

Lidz, C.S. (1987). *Dynamic assessment: An interactional approach to evaluating learning potential*. New York: Guilford Press.

Lidz, C. S. (1991). *Practitioner's guide to dynamic assessment*. New York: Guilford Press.

MacGinitie, W. H., & MacGinitie, R. K. (1989). *Gates-MacGinitie Reading Tests* (3rd edition). Itasca, IL: Riverside.

Mann, V., & Singson, M. (2003). Linking morphological knowledge to decoding ability: Large effects of little suffixes. In E. M. H. Assink & D. Sandra (Eds.), *Reading complex words: Cross-language studies*. New York: Kluver.

Mahony, D., Singson, M., & Mann, V. (2000). Reading ability and sensitivity to morphological relations. *Reading and Writing, 12*, 191-218(128).

Mahony, D. L. (1994). Using sensitivity to word structure to explain variance in high school and college level reading ability. *Reading and Writing: An Interdisciplinary Journal, 6*, 19-44.

Marslen-Wilson, W., Tyler, L., Waksler, R., & Older, L. (1994). Morphology and meaning in the mental lexicon. *Psychological Review, 10*, 3-33.

Marzano, R. J. (2004). *Building background knowledge for academic achievement*. Alexandria, VA: Association for Curriculum and Development (ASCD).

McBride-Chang, C., Wagner, R. K., Muse, A., Chow, B. W. –Y., & Shu, H. (2005). The role of morphological awareness in children's vocabulary acquisition in English. *Applied Psycholinguistics, 26*, 415-435.

McCutchen, D., Green, L., & Abbott, R. (2008). Children's morphological knowledge: Links to literacy. *Reading Psychology, 29*, 289-314.

McKeown, M. G. Beck, I. L., Omanson, R. C., & Perfetti, C. A. (1983). The effects of long-term vocabulary instruction on reading comprehension: A replication. *Journal of Reading Behavior, 15*, 3-18.

Meijer, J. (1993). Learning potential, personality characteristics, and test performance. In J. H. M. Hamers, K. Sijtsma, & A. J. J. M. Ruijssenaars (Eds.), *Learning potential assessment: Theoretical, methodological, and practical issues* (pp. 341–362). Lisse, Netherlands: Swets & Zeitlinger B.V.

Meunier, F., & Longtin, C.-M. (2007). Morphological decomposition and semantic integration in word processing. *Journal of Memory and Language, 56,* 457–471.

Miller, L., Gillam, R.B., & Peña, E.D. (2001). *Dynamic assessment and intervention: Improving children's narrative abilities*. Austin, TX: Pro-Ed.

Moore, B., & Moore, M. (1997). *NTC's dictionary of Greek and Latin origins: A comprehensive guide to the classical origins of English words*. Chicago: NTC Publishing Group.

Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English?" *Reading Research Quarterly 19,* 304-330.

Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal, 24*, 237-270.

Nagy, W.E., Anderson, R., Schommer, M., Scott, J.A. and Stallman, A. (1989) Morphological families in the internal lexicon. *Reading Research Quarterly 24*, 263-282.

Nagy, W., Berninger, V., Abbott, R., Vaughan, K., & Vermeulen, K. (2003). Relationship of morphology and other language skills to literacy skills in at-risk second-grade readers and at-risk fourth-grade writers. *Journal of Educational Psychology, 95*, 730-742.

Nagy, W., Berninger, V. W., & Abbott, R. D. (2006). Contributions of morphology beyond phonology to literacy outcomes of upper elementary and middle-school students. *Journal of Educational Psychology, 98*, 134-147.

Nagy, W.E., Diakidoy, I.N., & Anderson, R.C. (1993) The acquisition of morphology: learning the contribution of suffixes to the meanings of derivatives. Journal of Reading Behavior 25, 155-169.

Nagy, W., & Herman, P. (1984). *Limitations of vocabulary instruction* (Tech. Rep. No. 326). Urbana: University of Illinois, Center for the Study of Reading.

Nagy, W. E., & Scott, J. A. (2000). Vocabulary processes. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & Barr (Eds.), *Handbook of reading research Volume III* (pp. 269-284). Mahwah, NJ: Erlbaum.

Nation, I. S. P. (1990). *Teaching and learning vocabulary*. New York: Newbury House/Harper & Row.

National Longitudinal Transition Study II (2003). National Center for Special Education Research at the Institute of Education Sciences, U. S. Department of Education.

National Reading Panel. (2000).*Teaching Children To Read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Bethesda, MD: National Institute of Child Health and Human Development.

Nielsen, I. & Petersen, D. K. (1992). DIAVOK. Et material der afdoekker eventuelle lease – og stavenvanskeligheder [DIAVOK. A screening test for reading and spelling disabilities in adults]. Copenhagen: Workers Educational Association.

Nunes, T., Bryant, P., & Bindman, M. (1997). Learning to spell regular and irregular verbs. *Reading and Writing: An Interdisciplinary Journal, 9*, 427-449.

Nunnally, J. C. (1967). Psychometric theory. New York: McGraw-Hill.

Olswang, L. B., & Bain, B. A. (1996). Assessment information for predicting upcoming change in language production. *Journal of Speech and Hearing Research, 39*, 414-423.

Pany, D., Jenkins, J. R., & Schreck, J. (1982). Vocabulary instruction: Effects of word knowledge and reading comprehension. *Learning Disability Quarterly, 5*, 202-215.

Peña, E. (1993). Dynamic assessment: A non-biased approach for assessing the language of young children. *Dissertation Abstracts International, 54(2-A),* 485

Peña, E., Iglesias, A., & Lidz, C. (2001). Reducing test bias through dynamic assessment of children's word learning ability. *American Journal of Speech-Language Pathology, 10*, 138-154.

Peña, E., Quinn, R., & Iglesias, A. (1992). The application of dynamic methods to language assessment: A non-biased procedure. *Journal of Special Education, 26*, 269-280.

Perie, M., Grigg, W.S., and Donahue, P.L. (2005). *The Nation's Report Card: Reading 2005* (NCES 2006–451). U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. Washington, D.C.: U.S. Government Printing Office.

Rabin, J., & Deacon., H. (2008). The representation of morphologically complex words in the developing lexicon. Journal of Child Language, 35, 453-465.

RAND Reading Study Group. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: RAND Education.

Rastle, K., Davis, M., Marslen-Wilson, W., & Tyler, L. (2000). Morphological and semantic effects in visual word recognition: A time course study. *Language and Cognitive Processes, 15*, 507-537.

Reed, D. K. (2008). A synthesis of morphology interventions and effects on reading outcomes for students in grades K-12. *Learning Disabilities Research & Practice, 23*, 36-49(14).

Reichle, E. D., & Perfetti, C. A. (2003). Morphology in word identification: A word-experience model that accounts for morpheme frequency effects. *Scientific Studies of Reading, 7*, 219-237.

Restrepo, M. A. (1998). Identifiers of predominantly Spanish-speaking children with LI. Journal of *Speech, Language, and Hearing Research, 41*, 1398-1411.

Rice, M. L., Buhr, J. C., & Nemeth, M. (1990). Fast mapping word learning abilities of language-delayed preschoolers. *Journal of Speech and Hearing Disorders, 55*, 33-42.

Roseberry, C. A., & Connell, P. J. (1991). The use of an invented language rule in the differentiation of normal and language-impaired Spanish-speaking children. *Journal of Speech and Hearing Research, 34,* 596-603.

Rubin, H. *(*1988*).* Morphological knowledge and early writing ability. *Language and Speech, 31*, 337-355.

Sandra, D. (1993). The use of lexical morphology as a natural mnemonic aid in learning foreign language vocabulary. In J. Chapelle & M. –T. Claes (Eds.), *Proceedings of the First International Congress on Memory and Memorization in Acquiring and Learning Languages*, CLL, Louvain-la-Neuve, 263-294.

Schatschneider, C. (2006). *Reading difficulties: Classification and issues of prediction*. Paper presented at the Pacific Coast Research Conference, San Diego, CA.

Scheuder, R., & Baayan, R. H. (1995). Modeling morphological processing. In L. Feldman (Ed.), *Morphological aspects of language processing* (pp. 131-157), Hildale, NJ: Lawrence Erlbaum.

Shinn, M. (1989). *Curriculum based measurement: Assessing special children*. New York: Guilford.

Singson, M., Mahony, D., & Mann, V. (2000). The relation between reading ability and morphological skills: Evidence from derivational suffixes. *Reading and Writing: An Interdisciplinary Journal, 12*(3), 219-252.

Sirles, C. (1997). *Root Awakenings: Vocabulary Development Using Classical Word Roots*. Champaign, IL: Stipes Publishing.

Snow, C. E. E., Burns, M. S. E., & Griffin, P. E. (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy of Sciences—National Research Council Commission on Behavioral and Social Sciences and Education.

Snow, C., Porche, M. V., Tabors, P. O., & Harris, S. R. (2007). Is *literacy enough? Pathways to academic success for adolescents.* Baltimore: Brookes Publishing.

Spencer, A. (2001). Morphology. In M. Aronoff & J. Rees-Miller (Eds.), *The Handbook of Linguistics* (pp. 213-237). Oxford: Blackwell.

Spector, J. E. (1992). Predicting progress in beginning reading: Dynamic assessment of phonemic awareness. *Journal of Educational Psychology, 84*, 353–363.

Stahl, S. A. (1999). *Vocabulary development*. Cambridge, MA: Brookline Books.

Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research, 56*, 72-110.

Stahl, S. A., & Nagy, W. E. (2006). *Teaching word meanings*. Mahwah, NJ: Erlbaum.

Swanson, H. L. (1994). The role of working memory and dynamic assessment in the classification of children with learning disabilities. *Learning Disabilities Research and Practice, 9,* 190–202.

Swanson, H.L. & Lussier, C.M. (2001). A selective synthesis of the experimental literature on dynamic assessment. *Review of Educational Research, 71,* 321-363.

Taft, M. (2003). Morphological representation as a correlation between form and meaning. In E. Assink &D. Sandra (Eds.), *Reading complex words* (pp. 113–137). Amsterdam: Kluwer.

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Behavior, 1*4, 638-647.

Thacker, A. A. & Dickinson, E. R. (2004). *Item Content and Difficulty Mapping by Form and Item Type for the 2001-03 Pennsylvania System of School Assessment (PSSA)* (HumRRO Report No. 04-12). Alexandria, VA: Human Resources Research Organization.

Thacker, A. A., Dickinson, E. R., & Koger, M. E. (2004). *Relationships Among the Pennsylvania System of School Assessment (PSSA) and Other Commonly Administered Assessments* (HumRRO Report No. FR-04-33). Alexandria, VA: Human Resources Research Organization.

Tsesmeli, S. N., & Seymour, P. H. K. (2006). Derivational morphology and spelling in dyslexia. *Reading and Writing: An Interdisciplinary Journal, 19*(6), 587-625.

Tyler, A., & Nagy, W. (1987). *The Acquisition of English Derivational Morphology. Technical Report No. 407*.

Tyler, A. and Nagy, W. (1989) The acquisition of English derivational morphology. *Journal of Memory and Language 28,* 649-667.

Tyler, A. and Nagy, W. (1990) Use of derivational morphology during reading. *Cognition 36*, 17-34.

Ukrainetz, T. A., Harpell, S., Walsh, C., & Coyle, C. (2000). A preliminary investigation of dynamic assessment with Native American kindergartners. *Language, Speech, and Hearing Services in the Schools, 31*, 142-154.

Verhoeven, L., & Perfetti, C.A. (2003). The role of morphology in learning to read. *Scientific Studies of Reading, 7*, 209-217.

Vogel, S. A. (1977). Morphological ability in normal and dyslexic children. *Journal of Learning Disabilities, 10,* 41-49.

Vogel, S. A. (1983). A qualitative analysis of performance on a measure of morphology in learning disabled and achieving children. *Journal of Learning Disabilities, 16*, 416-420.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Washington, J., & Craig, H. (1992). Performances of low-income African American preschool and kindergarten children on the Peabody Picture Vocabulary Test-Revised. *Language, Speech, and Hearing Services in Schools, 23*, 329-333.

Washington, J., & Craig, H. (1999). Performances of at-risk, African American preschoolers on the Peabody Picture Vocabulary Test-III. *Language, Speech, and Hearing Services in Schools, 30*, 75-82.

White, T. G., Power, M. A., & White, S. (1989). Morphological analysis: Implications for teaching and understanding vocabulary growth. *Reading Research Quarterly,24*, 283-304.

Wiig, E., Semel, E., & Crouse, M. (1973). The use of English morphology by high-risk and learning disabled children. *Journal of Learning Disabilities, 6*, 457-465.

Williams, K., & Wang, W. (1997). *Peabody Picture Vocabulary Test-Third Edition Technical Manual*. Circle Pines: American Guidance Service.

Windsor, J., & Hwang, M. (1999). Derivational suffix productivity for students with and without language-learning disabilities. *Journal of Speech, Language, and Hearing Research, 42*(1), 220-230.

Wysocki, K., & Jenkins, J. R. (1987). deriving word meanings through morphological generalization. *Reading Research Quarterly, 22*(1), 66-81.