Effects of Embedded Morphological Instruction on Children's Morpho-Syntactic Production

Mary Northey

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2017

Reading Committee:

Deborah McCutchen, Chair

Elizabeth Sanders

Heather Hebard

Program Authorized to Offer Degree:

Education

University of Washington

Abstract

Effects of embedded morphological awareness on children's morpho-syntactic production

Mary Northey

Chair of the Supervisory Committee:
Dr. Deborah McCutchen
College of Education

Being able to use the words that are learned and encountered in school is an important skill for communicating and displaying understanding. Students are increasingly expected to be able to appropriately and effectively use complex academic words in their writing, yet we know little about this productive aspect of word knowledge. This study explores the effects of relevant word characteristics, student characteristics, and an embedded morphologically-focused instructional intervention for 499 fourth and fifth grade students randomly assigned to treatment and control classrooms (n = 233 and 266, respectively) on their morpho-syntactic production. Considering word-level SFI, word-level phonological and orthographic transparency (for Morphological Production outcome only), student-level comprehension and morphological knowledge at pretest, and classroom level instructional condition (treatment or business-as-usual control), item-level responses were predicted using cross-classified mixed models with random effects for two more-or-less constrained measures of morpho-syntactic production: Morphological Production and Sentence Combining Derivations. Results suggest that instruction has a unique and positive effect on students' control of morpho-syntactic word forms produced in sentence-level syntax activities. Students' pretest comprehension, as a proxy for general literacy skill, and students' pretest morphological skill were both uniquely predictive of both outcomes, for instructed and non-instructed words. Word-level frequency (SFI) was predictive only for deriving non-instructed words on sentence combining, while phonological and orthographic shift type was predictive only for deriving instructed words on morphological production. Implications for word-level instruction are discussed.

TABLE OF CONTENTS

Page

List of Figures…………………………………………………………………………………...i

List of Tables………………………………………………………………………………...ii

Introduction……………………………………………………………………………......1

Chapter 1: Supporting Students' Navigation of Complex Language……………………………..3

Chapter 2: Exploring Student's Morpho-Syntactic Production…….…………………………....17

Chapter 3: Modeling the Complexities of Language Processes…………………………………35

Chapter 4: Method……………………………………………………………………………49

Chapter 5: Results……………………………………………………………………………..62

Chapter 6: Discussion………………………………………………………………………78

References……………………………………………………………………………………..89

Appendix: Morphological Production and Sentence Combining Tasks ………………………...96

LIST OF FIGURES

Figure 1. Predicted probability of correct posttest MP response given pretest comprehension....72

Figure 2. Predicted probability of correct posttest MP response given pretest MP...….............72

Figure 3. Predicted probability of posttest MP response given treatment condition.…...……….73

Figure 4. Predicted probability of posttest SCD derivation given pretest comprehension...…….76

Figure 5. Predicted probability of posttest SCD derivation given pretest SCD….……...……….77

Figure 6. Predicted probability of posttest SCD derivation given treatment condition....……….77

LIST OF TABLES

Table 1. Teacher characteristics……………………………………………………….……49

Table 2. Student characteristics…………………………………………………...………..50

Table 3. Typical weekly lesson plan……………………………………………………..53

Table 4. Instruction overview………………………………………...……………………54

Table 5. Student-level descriptives for Morphological Production…………………...………62

Table 6. Student-level descriptives for Sentence Combining Derivations…………………..….63

Table 7. Zero-order correlations between student-level measures…………………..………..63

Table 8. Word characteristics……………………………………………………...………65

Table 9. Item difficulty: Morphological Production…………………………..………………67

Table 10. Item difficulty: Sentence Combining Derivations……………………………………68

Table 11. Morphological production: Instructed words…………………………...…………71

Table 12. Morphological production: Non-Instructed words……………………...…………71

Table 13. Sentence Combining Derivations: Instructed words…………………….………75

Table 14. Sentence Combining Derivations: Non-instructed words………………………......76

INTRODUCTION

Over the course of their education, children are expected to encounter an estimated 88,533 distinct word families in school (Nagy & Anderson, 1984). Practically, spending time to provide word-specific instruction is unthinkable. And yet, students are increasingly expected to read complex texts that include longer, more abstract, and less frequent words, and their comprehension of complex texts partially depends on familiarity with those words. Word-level instruction, therefore, must be designed with a focus on the right words and the right word-level attributes to help students become independent word learners.

Interventions that focus on supporting different aspects of students' word-knowledge are one common way researchers have chosen to address the recently increased emphasis on students' interaction with complex texts. By helping to demystify the more complex aspects of vocabulary that is common in complex academic texts—difficult orthographic patterns, nuanced usage, multiple affixes, and abstract meanings, for example—researchers hope to ultimately give students the tools they need to independently comprehend the complex texts that are increasingly encountered in school. Instructional interventions focusing on morphological awareness and vocabulary aim to help students to recognize and understand the complex academic words that they are likely to encounter.

The vast majority of word-level instructional interventions focus on reading outcomes, such as growth in vocabulary definitional knowledge, in reading complex words, and in reading comprehension. Writing outcomes, whether at the word, sentence, or text-level of writing production, are not often considered, despite the high standards that are placed upon students' writing skills. More broadly, measures of productive use of words (as opposed to receptive), as a measure of word learning, are not often considered.

This study addresses three areas of this research that require more attention in the field. First, this study adds to our understanding of morphological awareness by specifically investigating morphological production as a distinct aspect of morphological skill. In morphological awareness research, production of morphological derivations is not often the specific focus of study, nor is it investigated as a skill distinct from the more broadly-assessed receptive measures of morphological skill. Second, this study is among the first to consider language production outcomes from morphological awareness instruction by investigating students' sentence-writing skill. Though the last decade has seen many instructional interventions seeking to increase students' insight into the morphological structure of words, few studies have explored students' ability to produce complex morphological forms in writing tasks, instead focusing on more receptive measures of morphological awareness. Thirdly, this study addresses the complex nature of word learning by considering both word- and student-level features in the analysis of these writing outcomes. It is common for intervention effects to be interpreted at the group level without paying explicit attention to the inherent variation found among students, as well as among the words that are tested. This study accounts for some aspects of the continuous variation of both the words and students to better reflect the complex nature of the relationship: instruction is likely to affect students differentially, and it may be more easily applied to some words than others.

CHAPTER 1

Supporting Students' Navigation of Complex Language

As children progress through school, they encounter increasingly complex academic language. Academic language, according to Nagy and Townsend (2012), is the "specialized language, both oral and written, of academic settings that facilitates communication and thinking about disciplinary content" (p. 92). Academic language is *specialized* for the purpose of conveying "abstract, technical, and nuanced ideas and phenomena that are not typically examined in settings that are characterized by social and/or casual conversation" (Nagy & Townsend, 2012, p. 92). Compared to conversational English, academic language has more grammatical metaphor (e.g., abstract nouns that take a verb), and proportionally more nouns, adjectives, and prepositions than spoken language. It is also more informationally dense, and it deals with more abstract words and ideas than conversational English (Nagy & Townsend, 2012). These features can make the language of schooling somewhat inaccessible to students without explicit instruction. The increased emphasis on student exposure to genuine complex texts encouraged by the Common Core (National Governors Association and Council of Chief School Officers, 2010) has created a need to explore the skills necessary to gain access to and participate in academic language.

Academic language, put simply, is the language of schooling. The words, language patterns, and ideas that are used to communicate in school and in school texts are often not the same as those used at home or with friends. Children learn from an early age to control different registers, depending on their experiences outside of school, and continue to develop a more nuanced understanding of the characteristics of different registers over time (Schleppegrell, 2012). Heightened understanding of what type of language is valued in different situations helps students to switch between registers, depending on what a situation calls for. For example, in their study of "textisms" (like *LOL*, *u* instead of *you*, etc.), Plester, Wood, and Joshi (2009) found that use of different types of textisms (created in response to different imagined texting situations with peers) were significantly related to students' literacy skills. Use of accent stylization (*wiv*, *elp*, *anuva*), for example, was significantly and positively correlated with word reading and spelling. The authors noted that despite common misconceptions, these 10-year-olds were "aware of the different registers of language, and the boundary crossings surrounding the appropriateness of each" (p. 157). Though it is not uncommon for these types of spellings and phrasings to be written off as "poor English" or "improper grammar", the ability to design and craft language to reach certain social goals can be a marker of heightened control over language. Though stylized language has value of its own, the more formal language of schooling enjoys a privileged position within Western society, and proficiency in the register may be a prerequisite for success.

Academic Vocabulary

The linguistic features of academic language create a considerable obstacle to school achievement for struggling readers (Snow & Uccelli, 2009). Much of the research on students' academic language proficiency has focused on word-level features: specifically, academic vocabulary. To understand the increasingly complex texts encountered in school, students must know something of the words within the text: comprehension of complex texts relies heavily on familiarity with the words in a text and knowledge of their meaning. Academic language is full of words that children are less likely to encounter outside of school, and therefore less likely to learn without explicit instruction. The difficulty of this register is increased by the morphologically complex Greek and Latinate (as opposed to German) words that are so common in academic language: these words often have multiple affixes ({funct--ion}>al>>ism>), bound stems (that is, root morphemes that must be combined with another morpheme to form a word, like *path* in {sym--*path*--y}), phonological and/or orthographic shifts among family members (*visual*, *visible*, *vision*), and more abstract meanings (e.g., the word *analysis*) (Nagy & Townsend, 2012; Goodwin, Gilbert, Cho, & Kearns, 2014).

What researchers generally refer to as academic vocabulary—especially concerning students in upper-elementary grades—is what Beck and McKeown (1985) have termed "Tier Two" vocabulary. Tier Two vocabulary words are "words that are of high frequency for mature language learners and are found across a variety of domains" (Beck, McKeown, & Kucan 2002, p. 8). Tier One words are the most common words of the language; that is, the words which students frequently encounter on their own. Vadasy, Sanders, and Logan Herrera (2015) clarify the distinction between Tier Two and Tier Three words by describing Tier Three words as "labels for unfamiliar content [that are often] associated with domain-specific knowledge (e.g., photosynthesis, plebiscite, treble)" and Tier Two words as words that "allow for making finer distinctions about commonly known concepts, like sadness (e.g., grief, despair, depression, sorrow, bereft)" (p. 327). These words, which occur more frequently in written than in spoken language (McKeown, Beck, & Sandora, 2012), are not so infrequently used and conceptually obscure as to be considered academic jibberish (Krashen, 2012); instead, they are words that have high utility for expressing more nuanced and exact meaning to match authorial intent, and are less likely to be encountered by students outside of school.

Two fields that have investigated complex vocabulary as it relates to academic language proficiency are vocabulary research and morphological awareness research. While the fields overlap—it is well documented that much of the variance in morphological awareness outcomes can be attributable to vocabulary knowledge (Spencer, Muse, Wagner, Foorman, Petscher, Schatschneider, et al., 2015)—both have played an important role in contributing to our understanding of how word-level knowledge contributes to comprehension of more complex texts. Sufficient breadth and depth of vocabulary knowledge provides support for comprehension of complex texts (Perfetti, 2007). A large body of research (somewhat tentatively) links morphological awareness (that is, a "conscious awareness of the morphemic structure of words and [the] ability to reflect on and manipulate that structure" (Carlisle, 1995, p. 194)) to reading comprehension. While the underlying mechanisms of this relationship are not entirely clear, it is thought that the relationship between morphological awareness and comprehension is partially mediated by word reading (Deacon, Kieffer & Laroche, 2014; Gilbert, Goodwin, Compton & Kearns, 2014) and vocabulary (Nagy, Berninger & Abbott, 2006; Kieffer, Biancarosa & Mancilla-Martinez, 2013). Through the successful reading of morphologically complex words (Goodwin, Gilbert, & Cho, 2013; Kearns Steacy, Compton, Gilbert, Goodwin, Cho, Lindstrom, & Collins, 2014) and morphological analysis of words to infer word meaning (McCutchen & Logan, 2011; Deacon, Tong & Francis, 2017), varying aspects of morphological knowledge provide support in navigating complex academic language in the texts students encounter in school.

Between the less frequent, more obscure words, and the conceptually abstract and dense syntax, academic language is hard to comprehend while reading and control while writing. Because it is intrinsically so difficult to learn, Snow and Uccelli (2009) argue, efforts must be made to develop instructional programs supporting its understanding.

Word-Level Instruction

Partially due to the many demands placed on teachers by academic learning standards, teachers are often unable to provide the instructional time on vocabulary that students need, as compared to other areas of reading and writing instruction. And yet, vocabulary instruction is very important for student literacy development. Research indicates that persisting achievement gaps are in part due to differences in early vocabulary knowledge (Cunningham & Stanovich, 1997; Hart & Risley, 2003), but good vocabulary instruction is linked to growth in student reading skill. Carlisle, Kelcey, and Berebitsky (2013) provide evidence that teachers' support of student vocabulary learning (as measured by a model used to describe teacher discourse actions and the consistency with which those actions are used) is positively linked to student growth in reading comprehension above other relevant factors including teacher reading knowledge, mean classroom achievement level, and proportion of class with limited English proficiency. However, there is also evidence that teachers spend little time throughout the day teaching vocabulary (Scott, Jamieson-Noel, & Asselin, 2003), they often don't feel prepared to teach their students vocabulary in efficient, meaningful ways—including ways to support students' productive control of vocabulary (Miller, Gage-Serio, & Scott, 2010), and the textbooks used to prepare teachers largely do not provide support for teachers' understanding of effective vocabulary instruction (Wright & Peltier, 2016). Though research has provided us with a good idea of what effective vocabulary instruction should look like, there are still large gaps in our understanding of how to provide the most practical support for teachers. We know, for example, that vocabulary instruction should start developmentally early; it should provide multiple exposures of the words being taught, breadth of information (both definitional and contextual), and engagement of active or deep processing; and it should be focused on a carefully-selected set of words (see McKeown, Beck, & Sandora, 2012, for review). Support for teachers to implement vocabulary instruction that exemplifies these traits, however, is lacking.

Growing research suggests that word-level instruction with a focus on the morphological structure of words can provide a suitable framework for vocabulary instruction. Because of the close relationship between morphology and academic language, drawing attention to the morphological structure of English words through explicit instruction could help students manage the demands of academic language, as well as turn long lists of unknown academic words to be learned into a more manageable subset of related word families (Goodwin, 2016). By targeting the learning of the complex words and complex word features that are often unfamiliar to students, these interventions aim to improve not only familiarity with and knowledge of complex word meanings, but also to improve broader literacy outcomes, such as complex word reading, phonological awareness, morphological knowledge, vocabulary knowledge, and spelling (see Goodwin & Ahn, 2013, for a meta-analysis of morphological interventions).

The present study concerns a morphologically-focused vocabulary instructional intervention, aimed at improving reading and writing outcomes related to academic language proficiency. To address teachers' need for support in implementing effective word-level instruction, as well as to address a gap in the research concerning word-level knowledge as it applies to productive (writing) outcomes, our research group designed a teacher-implemented embedded vocabulary curriculum with a morphological focus. Most intervention studies concerning vocabulary or morphology instruction focus on broader intervention outcomes and do not necessarily consider growth by student and word characteristics. It is important, however, when investigating approaches for word-level instruction, to consider which types of words students show most growth on, and for which students the growth is strongest. This will better inform our understanding of best practices for word-level instruction by providing a more detailed account of the specific areas where student learning gains are the strongest as a result of intervention. The present study investigates the specific contribution of both word and student characteristics to students' productive control of complex morphological vocabulary in sentence writing tasks, in addition to the effects of instruction, accounting for characteristics of both words and students in their responses.

Below, I will discuss the typical goals and instructional approaches of morphology-focused instruction research. I will then address the lack of research on word-level productive outcomes (producing complex words in sentence-writing tasks) on instructional interventions, and provide initial discussion regarding its potential. Lastly, I will introduce the goals and approaches of our instructional intervention.

Morphological Instruction

Because English is a morphophonemic orthography, awareness of the morphological structure of words can provide insight into the semantic, phonological, orthographic, and syntactic features of English words. The insight afforded from awareness of the morphological structure of English words (and other morphology-based languages) can help with reading unfamiliar words, analyzing unfamiliar words for their meaning, and recognizing spelling patterns and their relationship to phonology and meaning. For example, the relationship between *genealogy* and *gene* (and not *jean*) becomes more apparent from the orthographic consistency between the two words, made explicit in written language. Developmentally, awareness of the morphological structure of words begins at an early age and continues to develop through the school years. Control over inflectional forms (forms that change tense and number, like *walk* →

walked or *bird* → *birds*) is typically achieved by third or fourth grade (Menyuk, 1988). Control over lexical morphology (Jarmulowicz & Taran, 2013), which concerns knowledge of derivational forms (forms that change grammatical category, like *walk* → *walker*), continues to develop after elementary school (Tyler & Nagy, 1989; Berninger, Abbott, Nagy, & Carlisle, 2010). Typically, studies investigating morphological instruction target students in upper-elementary school's understanding of and control over lexical morphology.

Explicit instruction on the morphological features of words is implemented with the goal of directly improving word-level knowledge. At a higher level, research in morphologically-focused instructional interventions aims to investigate whether this word-level instruction can have positive effects on higher-order reading skills (e.g., reading comprehension; Bowers, Kirby, & Deacon, 2010). Additionally, as Kuo and Anderson (2006) suggest, the relationship between morphological awareness and reading is not unidirectional. While across languages, morphological awareness appears to cause growth in reading, the relationship is likely reciprocal, and grows stronger throughout development (Kuo & Anderson, 2006). Thus, many researchers recommend morphological instruction as an important strategy for supporting students' reading comprehension as they increasingly encounter complex academic language.

Instructional approaches

Morphology-focused instructional interventions have typically fallen into four basic instructional approaches (as identified by Carlisle, 2010): 1) heighten student awareness of the morphological structure of words, 2) teach the meanings of affixes and base words, 3) foster morphological problem solving, and 4) instruction in and application of morphological analysis strategies. Though instructional goals and approaches across the field have differed, there is still strong evidence that explicit instruction on the morphological nature of English words improves students' learning of specifically-taught words, as well as transfer of knowledge to similar, untaught words (though generally with a smaller effect). In their meta-analysis reporting weighted effect sizes from morphological interventions (both direct and embedded instruction), Goodwin and Ahn (2013) reported significant growth of intervention students (relative to their comparison group peers) in decoding ($\bar{d}$=0.59), phonological awareness ($\bar{d}$ = 0.48), morphological knowledge ($\bar{d}$ = 0.44), vocabulary ($\bar{d}$ = 0.34), and spelling ($\bar{d}$ = 0.30).[1]

Many morphology instructional interventions emphasize teaching lexical patterns across many words rather than focusing on a list of previously-identified words or word families. Bowers and Kirby (2010), for example, taught 430 words over 20 sessions, using "structured word inquiry" aimed to teach core concepts of morphology and orthography to fourth and fifth graders. Apel, Brimo, Diehm, and Apel (2013) taught affix lessons to 2nd and 3rd grade students to directly heighten awareness of affixes and the relations between base words and their inflected and derived forms. More broadly, they aimed to encourage students to think about and manipulate morphemes in a broad sense and for a variety of purposes. These studies, like most other studies on morphology-focused instruction, focus specifically on bringing students' direct attention to the morphology of words to foster a conscious awareness of word structure to be transferred to other complex words they will encounter.

Other interventions adhere more closely to the ideal instructional approaches as outlined by vocabulary research. Lesaux, Kieffer, Faller, and Kelley (2010) citing the need for more comprehensive, classroom-based vocabulary instruction, designed an intervention to be

---

[1] Calculated effect sizes are the standardized mean difference (g_i) between comparison and treatment group, corrected for the small sample bias.

implemented by classroom teachers that emphasized strategic and explicit vocabulary instruction especially aimed at supporting linguistically-diverse classrooms. Across eight 8-day instructional units, teachers taught 8-9 complex academic words, for a total of 72 specific words taught in the program. Throughout each 8-day unit, students encountered and practiced with the new words in a variety of ways, including reading the words in context, practicing using context clues to decipher word meaning, building definitions, answering text-based questions using the target words, using target words to write novel sentences, analyzing the roots and affixes of the target words, practicing forming new words using constituent morphemes, and using the target words in longer writing exercises. By providing many varied opportunities for students to encounter and use the target words, and by teaching to many different aspects of word knowledge (definitions, morphology, syntax, etc.), the authors attempted to create a rich curriculum with which linguistically-diverse classrooms can improve in vocabulary knowledge and reading comprehension.

Regarding student learning outcomes, the primary focus of the investigation, Lesaux et al. (2010) found significant treatment effects (as compared to control classrooms) on target word mastery (multiple-choice definition task), morphological decomposition (complexity: The problem is _______. *complex*), and word-meanings-in-context comprehension (a contextualized vocabulary measure using target words from the intervention), though did not find significant effects of treatment on norm-referenced reading measures (Gates-Macginitie Reading Comprehension and SAT-10 Reading Vocabulary). A secondary goal of the investigation was to examine the ecological validity of more comprehensive vocabulary instruction as implemented by classroom teachers. Many instructional interventions are taught by researchers, are not easily implemented within the classroom, or are appropriate only for certain learners. Overall, they found that the participating teachers were mostly successful in implementing the curriculum, though perhaps not unexpectedly, had less success with students below grade level on vocabulary.

By considering the differential effects of their intervention on students of varying skill levels, Lesaux et al. (2010) provide valuable information concerning the students for which instruction like this has the highest effect. They note in their discussion that their findings "highlight and reinforce the challenge of bolstering vocabulary and reading comprehension skills for students who are not on grade level" (p. 219). Most studies of morphological interventions report overall effects between groups, though few provide detailed investigations of differential outcomes across students or words within the curriculum. More nuanced analyses like those provided by Lesaux et al. help to construct a more complete understanding of the effects morphological instruction might have on learning outcomes across different words and students. To add to this understanding, the current study also considers more nuanced effects of instruction.

Additionally, despite including many instructional activities where students practiced producing the instructed target words in their oral and written language output, Lesaux et al. (2010) did not investigate any language production outcomes. Few word-level instruction studies include production outcomes as a focus of investigation, despite the importance of students being able to use the words that they learn through instruction.

The potential of production outcomes for word-level instruction

Much of the research on word-level instructional interventions—both in vocabulary and in morphological awareness—focuses on receptive lexical knowledge. Studies investigate the processes of learning new words and their structural features, best practices for instruction, and how both lead to gains in reading comprehension. Fewer studies examine the processes involved in the ability to successfully use complex vocabulary in spoken and written expression, yet students are consistently called upon to display their understanding and opinions through production in the academic language register. By fifth grade, for example, under the Common Core Standards, students are expected to both *acquire* and *use* age-appropriate general and domain-specific academic vocabulary (National Governors Association and Council of Chief School Officers, 2010). They are also expected to be able to manipulate sentence-level syntax, with the intent of adjusting to specific meaning, audience interest, and style purposes. In the Common Core Standards, these skills are considered important in ensuring that students become independent, critical thinkers who can effectively produce language to communicate ideas and participate in discussion. The ability to both write and speak in the academic language register, therefore, is valuable in school.

Beyond the essential value of effective language production, there may be unexplored benefits of practicing production on general literacy skills. Nagy and Townsend (2012), for example, suggest that the ability to skillfully use academic vocabulary in written and oral language production may have reciprocal effects with reading comprehension. McKeown et al. (2012) hypothesize that a student's ability to produce a word may be indicative of their ease of access to that word, and ease of access is important for comprehending more complex language. This presents a potentially valuable skill area for investigation.

Appropriate lexical production requires using words within a broader context (sentences and texts), so assessing lexical production also necessitates considering how a word is used semantically and syntactically within the sentence or text. Controlling developing syntax and revising existing syntax can be difficult for young writers—especially when trying to use a more complex word in a sentence. Academic language proficiency requires understanding beyond word recognition and word meanings, and despite the potential of morphology and vocabulary research to provide insight into children's attempts at academic language production, there has been little attention paid to it. Investigating students' morpho-syntactic production (that is, using complex words in context in a syntactically and semantically appropriate way) could provide important insight into students' maneuvering of complex academic language at the level of the sentence (syntax) and the word (morpho-syntax) (Berninger, Nagy, & Beers, 2011).

Embedded Morphologically-Focused Vocabulary Instruction

To explore the effects of morphological instruction on language production outcomes, as well address the needs of classroom teachers, our research group designed a twelve-week intervention curriculum for teachers to implement in their own fourth and fifth grade classrooms. The purpose of our larger study was to investigate whether we can improve both student reading and writing using instruction that 1) focuses on students' morphological knowledge and skills 2) uses authentic activities and 3) supports teachers' overall instructional goals. Like more typical vocabulary interventions, our curriculum taught a specific subset of complex words (Baumann, Kame'enui, & Ash, 2003). Students in treatment classrooms learned the meanings of 40 target words and their morphological relatives, averaging 4.5 new word families per week over the course of the intervention. They practiced reading, hearing, speaking, and writing the words across three academic topics and the three school genres emphasized in the CCSS (narrative, informative, opinion; National Governors Association and Council of Chief School Officers, 2010).

As part of the broader study, we were interested in exploring the ecological validity of our embedded morphology-focused intervention, as well as in exploring its potential effect on various reading and writing outcomes. The current study concerns a smaller subset of those outcomes: specifically, it concerns students' productive use of complex lexical morphology. Two of our tested outcomes—Morphological Production and Sentence Combining Derivations—involved morphological manipulating a word so that it fits appropriately within a sentence. These measures assess students' ability to flexibly manipulate a given base word to fit within existing syntax (Morphological Production; *memory*: The party was *memorable*.) or to revise existing syntax (Sentence Combining Derivations; The boy described his plan. He often imagines. The way it was described was vivid. → *The imaginative boy vividly described his plan*.). Few studies have addressed these skills, and, consequently, our understanding of morpho-syntactic production—including what it is and how to measure it—is limited. In the following chapter, I will discuss how these tests are used, what they may be measuring, and what the demands of controlling complex syntax and of controlling production of appropriate morpho-syntactic forms within syntactic structure are.

CHAPTER 2

Exploring Student's Morpho-Syntactic Production

In school, students display their understanding of a topic through their writing and through their participation in class discussions. This sharing of knowledge often requires the use of more complex words, especially when exploring more complex topics or disciplinary topics that have more specific vocabularies. When writing about space, for example, students might need to use the word family *gravity* (*gravitational*, *gravitate*). When describing the events and tensions that led to the American Civil War, they might need to use the word family *abolition* (*abolitionist*, *abolitionary*). Though there are certainly ways to get around using these specific words (e.g., *the people who were against slavery*), the ability to use the words that are learned in class could show 1) more nuanced understanding of material, and 2) greater ownership over the new words learned (i.e., more developed word knowledge). In addition to showing understanding of meaning, productive use of words could provide information about the student's knowledge of the word's grammatical function, by building a syntactically-appropriate sentence around the word. Typically, when students use words in their writing and speech, it is referred to as productive language use. However, what is considered to be production is not very well-defined, nor is it well-understood. In the following sections, I will review the literature on word-level production from vocabulary research and from morphological awareness research. Then, I will discuss sentence-level language production, and how morphological skill may provide a testable link between word-level knowledge and flexibility, and control over language production.

Language Production

For vocabulary researchers, language production typically refers to the words students can spontaneously produce in their written and oral language. Researchers might measure the lexical diversity or density of spontaneous student writing to estimate their productive abilities, for example. Language production, in vocabulary research, is “widely estimated, but rarely (or not at all) defined” (Melka, 1997, p. 84). In practice, language production is often situated in opposition to receptive language skills. In vocabulary assessment, for example, vocabulary knowledge is divided into a four-category matrix: receptive vs. productive, and written vs. oral (Pearson, Hiebert, & Kamil, 2012). The four assessable areas of vocabulary knowledge, therefore, are reading vocabulary, listening vocabulary, spoken vocabulary, and written vocabulary. While these categories provide a clear object for assessment, they are not, in reality, completely separate areas of knowledge. In truth, productive and receptive vocabulary knowledge, like many other facets of word knowledge, are interrelated.

Some vocabulary researchers theorize that some threshold of word knowledge must be met for students to be able to produce that word in their oral or written language (McKeown, Beck, & Sandora, 2012). This question is not often researched, though it is often assumed in discussions of language production. To further understand how production is conceived of in word-level research, an exploration of what it actually means to know a word is necessary.

Word Knowledge. Word knowledge is not all-or-nothing—word learning is thought of as an incremental process. Addressing this incrementality, some vocabulary scholars propose that word knowledge exists on a continuum. Dale (1965) defined this continuum regarding specific word knowledge as the following steps of knowledge acquisition: 1) *I never saw it before*. 2) *I’ve heard of it, but I don’t know what it means*. 3) *I recognize it in context—it has something to do with…* 4) *I know it*. (p. 18). Paribakht and Wesche (1997) added 5) *I can use it in a sentence.* to the end of the continuum. Meara (1997), however, takes issue with the conceptualizing of the incremental nature of word learning as a continuum, arguing that conceptualizing vocabulary knowledge as existing on a continuum implies an over-simplified linear growth of word learning. Addressing this gap, he outlines a possible model for word acquisition: when we encounter new words, they are encoded by creating links between the new words and existing words in the lexicon. Knowing a word, then, is a “cumulative activity rather than an all-or-nothing affair” (Meara, 1997, p. 117). Though his argument specifically applies to L2 language learning, it can arguably be applied to L1 vocabulary acquisition as well. Though most agree that word learning is incremental, whether knowledge of a word exists on a continuum that ends in *I can use it in a sentence* seems unlikely, given that it is possible, for example, to create new words as we write, or produce legal morphological forms that we don’t necessarily “know” (McCutchen & Stull, 2015).

Vocabulary measures of production typically consider knowledge of meaning. However, it likely takes more than knowledge of meaning for students to use a word well in their language production. Nagy and Scott (2000) argue that word knowledge is primarily procedural rather than declarative, because “knowing a word means being able to do things with it” (p. 273). They identify five main aspects of the nature of word knowledge: incrementality (knowing a word is not all-or-nothing—for example, I recognize the word *transpicuous* and am able to pronounce it, but I do not know how to use it well in a sentence), multidimensionality (word knowledge encompasses many different types of knowledge, like knowing how to spell or pronounce the word, knowing a word’s general meaning, knowing appropriate contexts for use, knowing its morphological relatives or synonyms, etc.), polysemy (words often have multiple meanings; for example, *bear* can mean *a large mammal* or *to carry*), interrelatedness (knowing a specific word is not independent of knowing other words; (knowing the word *mammal* often requires knowing words like *whale*, *human*, or *lion*) and heterogeneity (what it means to know a word is dependent on the word; for example, knowing the word *table* is different from knowing the word *existentialism* or the word *about*). Lexical knowledge, therefore, is highly dependent upon word characteristics and learner characteristics, and whether a student "knows" a word is not easily measured.

But—we still try to measure it anyway. There are limitations to measuring word use in student writing, given the situated nature of any writing activity. Some have argued that the words students spontaneously produce in their writing are likely words that can be easily retrieved, and that the words that are produced are the words students have more robust knowledge of (McKeown, Beck, & Sandora, 2012). This may be true; however, because writing is a situated activity, when students respond to a question or prompt, they are participating in an ongoing conversation, which is likely to already include vocabulary relevant to the topic. They might be primed by the words surrounding the activity, including classroom conversation, words appearing in the text or in the prompt, or words recently encountered. Thus, using a specific word in their writing might signal recent priming or convenience as much as it signals their knowledge of the word in general.

In more contrived writing situations, where the parameters and purpose of the student response is narrow, use of specific words in writing may not show whether a student "knows" a word, but rather, which aspects of word knowledge the student might have a firmer grasp of. Correct use of a word in writing might not necessarily reveal knowledge only of that specific word, but rather a better grasp of the patterns and structures of English words in general. For example, it might be rather unlikely for students to encounter the word *unsuspicious*, because of its low frequency in school texts (appearing .0163 times per million words; Zeno et al., 1995). However, if they have a stronger understanding of the morphological patterns of English words, they may be more capable of drawing upon their understanding of distributional morphology to apply an appropriate affix to match their intent. Similarly, children might "invent" non-standard morphological forms to convey their intended meaning (McCutchen & Stull, 2015).

Because of this, it is difficult to assume much about the depth of students' word knowledge when they use a specific word once in their written or spoken language. We can, however, assess the way they use the word in its broader context to gain a better understanding of their ability to use words and language patterns. By tapping into more specific word-production skills, we hope to gain a more nuanced understanding of what students need to be able to successfully use complex words in their academic language production.

Morphological Production

One specific aspect of productive word knowledge we can assess is the ability to manipulate the morpho-syntactic forms of words to fit syntactically within a sentence. For clarity, I refer to this as morphological production, though there is no clear definition of morphological production within the field of morphological awareness research—nor is it necessarily measured as a separate skill. This is partially due to discontinuities in term use in the field of morphological awareness research. Apel (2014) called for increased precision when measuring the construct of morphological awareness, both in defining the skills of interest (whether the object of investigation is morphological *awareness*, morphological *skill*, or morphological *production*, for example) and in creating assessments. Different types of tasks have been used as a general proxy of the broader construct of morphological awareness. Similarly, a range of tasks have been used to measure morphological production. Because this investigation concerns tasks purporting to measure general morphological production, it is necessary to clarify what morphological production is, what it is not, and how it fits into our knowledge about morphological awareness in general.

Morphological Production: Definition and Tasks. Apel (2014) situates morphological production against morphological awareness in terms of consciousness: he describes morphological production as unconscious use of morphemes (typically in spoken language), and morphological awareness as conscious reflection on morphemes. This description may be helpful when describing students' spontaneous use of morphologically complex words in their oral and written expressive language use, and is in fact similar to vocabulary researchers' general conceptualizations of language production. However, in terms of measurement, isolating and assessing different aspects of morphological production presents similar problems to that of isolating and assessing different aspects of morphological awareness: what we call "morphological production" and how we measure it varies across the literature.

Most definitions of morphological awareness refer to both *awareness* of the morphemic structure of words (judgments about semantic and syntactic relationships) and to the ability to *manipulate* that structure for specific use (Carlisle, 2000; Deacon, Kirby, & Casselman-Bell; Kuo & Anderson, 2006; see Apel, 2014 for review). However, as Apel (2014) argues, no researcher-developed tasks measure morphological awareness in a complete way. Morphological awareness—as a construct—is measured across studies using a variety of tasks: production tasks, judgment tasks, blending or segmenting tasks, analogy tasks, and identification of written affixes, to name a few. Though these tasks are all used to measure general morphological awareness, each only taps into a specific aspect (or aspects) of the construct, which prevents researchers from fully understanding the underlying mechanisms of morphological awareness. For example, judgment tasks, like the *comes from* task (Nagy, Berninger, & Abbott, 2006; does *corner* come from *corn*?) show relational morphological knowledge; that is, recognizing when derivations that share the same root morpheme are a part of the same morphological family. Tyler and Nagy's (1989) cloze task asking students to choose the correctly-suffixed word or pseudo-word to complete a sentence shows syntactic morphological knowledge, which is an understanding of the role of suffixes in marking words for syntactic category. Most of these tasks require receptive knowledge of morphology: rather than retrieving new forms from the lexicon, students might choose from a multiple-choice list (Singson, 2000; Nagy, Berninger, & Abbott, 2006), analyze the meaning of provided morphologically-complex words (Bowers & Kirby, 2010; Crosson & McKeown, 2016), or circle a base word within a longer, morphologically-complex word (Bowers & Kirby, 2010).

In the following sections, I will describe the different measures and activities often classified as morphological production in the literature, as well as measures and activities that are not referred to specifically as morphological production, but perhaps should be. For the sake of this review, any task that requires active morphological manipulation of a base word to produce a spoken or written derived form is classified as morphological production.

Production of morphologically-complex words in spontaneous writing**.** What Apel (2014) refers to as morphological production is merely the spontaneous use of words that are morphologically complex. For example, Carlisle (1996) investigated 2nd and 3rd grade students' productive use of morphological inflections and derivations in their spontaneous narrative writing. Green, McCutchen, Schwiebert, Quinlan, Eva-Wood, and Juelis (2003) counted the number of morphologically complex words (inflections and derivations) spontaneously produced by 3rd and 4th grade students in their timed narrative writing in response to a picture. Both studies used their accounts of descriptive word use to describe the developmental trajectory of appropriate productive use of inflectional and derivational forms in spontaneous writing. These studies largely suggest that children produce very few morphologically complex derivations and few academic vocabulary words in their written texts regardless of genre (Green, McCutchen, Quinlan, Eva-Wood, & Juelis, 2009; Berman & Nir-Sagiv, 2007; Olinghouse & Wilson, 2013). While measuring students' spontaneous complex word use can provide information about various aspects of students' word knowledge, it does not provide specific information about students' morphological awareness. That is, use of a morphologically complex word could just as easily be classified as vocabulary knowledge, rather than morphology-specific knowledge. Thus, studies like these do not fit in the classification of morphological production. Additionally, some studies require students to provide a written or spoken definition of a morphologically-complex word (Baumann, Carr Edwards, Bolad, Olejnik, & Kame'enui, 2003). This requires morphological analysis, but not active morphological manipulation of derived forms.

Morphological production: Highly-constrained measures. Some studies administer what Apel (2014) refers to as "fluency tasks" of productive morphological awareness. These tasks provide students with a base word, and prompts them to provide as many morphological relatives to that base word (inflectional or derivational) as possible (Goodwin et al. 2014). Because they require active morphological manipulation and the production of a spoken or written form, these tasks would be considered morphological production tasks. However, these authors used this task not as a measure of morphological production, but rather as a measure representing morphological awareness. Additionally, analogy tasks (sing : singer, ;; magic : ___ (*magician*)) require active manipulation and written or spoken production, but are rarely used as an outcome of investigation.

Carlisle's morphological production task. The most widely-used task of morphological production is Carlisle's (1995, 2000) "Test of Morphological Structure." Carlisle (1995, 2000) developed a measure designed to illuminate the relationship between students' knowledge of word structure and their knowledge of word meaning. Half of the items in this task required decomposition (i.e., generating a base word) (farmer: My uncle lived on a ________.) and half of the items required derivation (i.e., generating a morphological derivation) (create: The child was ________.). Additionally, the words were classified by what she referred to as "linguistic complexity"; that is, half of the items could be classified as requiring a phonological and/or orthographic shift (*wise* → *wisdom*), and half of the items could be classified as requiring no shift in pronunciation or spelling (*farmer* → *farm*). Administering these tasks to third and fifth graders, Carlisle found that for both grades, the derivation tasks were more difficult, and deriving or decomposing words with a phonological and/or orthographic shift was more difficult than transparent word pairs. Thus, we see that production of morphological relatives, even given a syntactic framework to prime responses, is more challenging for students (even in fifth grade still) than identifying root morphemes. Additionally, when using the derivation and decomposition measures to predict students' performance on a morphologically complex word definition task, only derivation was a significant predictor of this writing task, and for fifth graders, the predictors accounted for a significant 43% of the variance in the outcome.

Many researchers have adapted this task for their own use. For example, Apel, Wilson-Fowler, Brimo, and Perrin (2012) use Carlisle's (2000) task as a measure of morphological awareness for second and third graders, citing researchers' suggestions (e.g., Carlisle, 1995; Carlisle & Fleming, 2003) that tasks that require both semantic and syntactic knowledge like the Carlisle (2000) production task may provide a better measure of morphological awareness because they assess both grammatical and meaningful aspects of derivational suffixes. Student success with this morphological production task requires the three areas of knowledge that constitute morphological awareness as outlined by Tyler and Nagy (1989): relational, syntactic, and distributional knowledge of morphology. Because of this, many researchers use the derivation and/or decomposition tasks as a measure of morphological awareness for different purposes (Kieffer & Leseaux, 2008; McCutchen, Green, & Abbott, 2008; McCutchen, Logan, & Biangardi-Orpe, 2009; McCutchen & Logan, 2011; Kraut, 2015). It has been used to predict reading comprehension for fifth grade Spanish-speaking English language learners (Kieffer & Lesaux, 2008) and reading comprehension for typically-achieving fifth- and eighth-grade students (McCutchen, Logan, & Biangardi-Orpe, 2008). In both cases, morphological production was significantly predictive of comprehension after accounting for other relevant skills.

Despite these findings, a robust connection between morphological awareness and reading comprehension is still lacking, and the possible contributions made by morphological awareness to reading comprehension—whether directly or indirectly—are still not understood. This may be, as Apel (2014) argues, partially due to the inconsistencies in measurement across the field. When many different tasks—tasks which vary in the type of skills and knowledge they require for successful completion—are used to measure one broader construct without distinction, it becomes more difficult to tease out the already complicated nature of these relationships. It is likely that it is more difficult, and that it requires more component skills, to produce a morphological derivation given a base word and stem sentence than it is, for example, to pick the correct option from a multiple-choice list.

Because there has been limited empirical investigation into morphological production as a unique skill—or at least a skill requiring slightly different or slightly more processing than more traditional receptive morphological awareness tasks—we don't know what exactly these tasks are measuring. It is unknown whether different tasks can be used interchangeably to measure morphological awareness (Apel, 2014). We do not know, for example, what specific skills might differentiate Carlisle's (2000) morphological production task from Singson's (2000) multiple choice syntactic knowledge task. Both require understanding of morphology and syntax, yet production seems to require just a little bit more. This lack of understanding is due in part to a general trend in research goals: most studies of morphological awareness are concerned with understanding how morphological awareness relates to reading skill, rather than writing skills, and production is generally linked to writing or speaking language. Additionally, productive and receptive language skills are impossible to fully separate, which remains true in the creation of productive and receptive language measures.

Specifically, this morphological production task (and others like it) places many competing demands upon the student: it requires phonological, orthographic, syntactic, and semantic knowledge. Students must recruit phonological and orthographic knowledge to read and identify the word correctly, and then use syntactic and semantic knowledge to identify the correct new suffix. Then, they must use orthographic and phonological knowledge to transcribe the new word correctly (or at least with a phonological approximation). Additionally, this task requires some executive functioning. To produce a new word, students must simultaneously hold the provided base word in mind and manipulate its suffix so that it will fit syntactically within the sentence. Clark and McCutchen (2014) speculated that this might require what Cartwright (2002) refers to as cognitive flexibility, and that this flexibility, or other unidentified processing skills, may provide the link between morphological awareness and reading comprehension. Reading comprehension requires more than just knowledge of word meanings; it requires shifting attention between different word attributes (for developing readers, between phonological and semantic features of words; Bialystok & Niccols, 1989; Cartwright, 2002). Because Carlisle's morphological production task (and others like it) have been found to uniquely predict variance in reading comprehension, there is a question of whether something specific to morphological awareness is predicting comprehension, or whether it has more to do with language-specific cognitive flexibility or other processing artifacts associated with the skills necessary for this shifting-task. Clark and McCutchen (2014) found that even with a parallel task requiring similar process demands (but without the morphological aspects), Carlisle's morphological production task continued to account for unique variance in reading comprehension, concluding that this task captures skill beyond vocabulary and whatever processing skills are necessary for holding words in mind and manipulating them to complete a sentence.

These findings suggest that this task measures skills beyond what other more receptive morphological awareness tasks measure. Similarly, Carlisle's (2000) findings that producing a derivation given a base word and stem sentence was more difficult for both third and fifth graders than identifying the root given a complex derivation and a stem sentence provides evidence that this task requires more than typical receptive measures. This task, which is highly constrained—there is only one (and in some cases two) appropriate word to complete each sentence—provides a good measure with which to explore not only children's morphological awareness, but also their ability to produce syntactically- and semantically-appropriate complex words in context, somewhat simulating the use of words in more complex academic language. When students are compelled to produce a derivation to fit into a sentence, can they?

Demands of Written Production

Other tasks of morpho-syntactic production require the student to produce more than just a single word. When this is the case, other skill related to writing are necessary. For developing writers, writing a complete sentence—especially a sentence containing complex words—can be difficult. In general capacity models for writing, both storage/retrieval and processing demands compete for limited working memory space (Kellogg, 1996; McCutchen, 1996). In Berninger and Amtmann's simple view of writing (2003), transcription (turning language into symbols on a page), text generation (forming ideas into language), and executive functioning (planning, goal-setting, revising, etc.) compete for a finite amount of short-, working-, and long-term memory space needed to write. For children who are learning how to write, lower-level demands can place an especially significant burden on their ability to successfully produce writing. Skills such as spelling and handwriting may significantly hinder a child's ability to attend to higher-level functions (Berninger, Yates, Cartwritght, Rutberg, Remy, & Abbott, 1992; Olive & Kellogg, 2002; Alves, Branco, Castro, & Olive, 2012). Without the automaticity of certain skills, precious cognitive resources must be allotted to simpler motor and lower-level writing skills. Struggling young writers are often slower to retrieve words than their higher-skilled peers (McCutchen, Covill, Hoyne, & Mildes, 1994), and even among adults, fluency, measured by length of writing bursts (which occur during text generation), has been found to be predictive of skill (Chenoweth & Hayes, 2001). Over time, and with mastery of these skills, students are better able to attend to other writing goals. Because writing involves multiple competing demands, fluency and flexibility with production at different levels of language help to free up the space needed to attend to higher-order skills.

Syntactic production. For fourth and fifth grade students, forming a complete sentence can present challenges. It can be especially difficult to adjust to the syntax and register of school writing, given that it differs so much from the register a student might speak at home or with friends. Students this age tend to produce simply-formed sentences, stringing together clauses using common connectors like “and,” “but,” or “because” (Myhill, 2008; Northey, McCutchen, & Sanders, 2016). Weaker writers (and younger writers; writing development is marked by age as well as skill) are more likely to lose control of their sentences by writing poorly-formed sentences (fragments or run-ons) or by failing to express their intended purpose (Myhill, 2008). Higher quality student writing in the expository school genre, as judged by holistic writing rubrics, tends to also have more syntactic variety, such as varying sentence length or greater use of non-finite clauses or adverbials to open a sentence rather than a subject (Myhill, 2008). The stage of writing development when writers begin to use various syntactic devices to fulfill various purposes appears after middle school (Ravid & Tolchinsky, 2002; Berman, 2009). However, though late-elementary and middle-school students tend not to use complex syntactic structures in their spontaneous writing, it does not mean that they are unable to use them. Berman (2009) suggests that part of syntactic development depends on acquiring more advanced vocabulary and word knowledge so that already-available syntactic structures can be adapted to become more complex.

Consider the following example, where Berman (2009) traces how developing writers might adapt familiar word forms (use of the suffix *-ing*, an inflectional form that is mastered early in development) to be used in more syntactically-advanced ways:

> Juvenile use of these forms in isolation is supplemented by the auxiliary *be*…as complements of perception verbs (*see someone running*), later as nominalized forms (e.g., *fighting is not good*), and even later for non-finite subordination in adverbial

> clauses (*Running to meet her, he fell and twisted his ankle*) or relative clauses (*an issue requiring a solution*)" (Berman, 2009, p. 354).

These more advanced examples of word use provide an illustration of how flexible production of different morpho-syntactic structures can be used to achieve different communicative goals (Ravid & Tolchinsky, 2002). Morphological awareness, and the flexibility to manipulate word-level structures to reach sentence-level goals, is needed to maneuver through these complex sentence structures. Little research in this area—specifically, how morphological skill might contribute to control over syntax—limits our understanding of the relationship. However, we do have some ideas of how morphological skill supports control over syntax, which are couched in capacity theories of writing (Berninger & Amtmann, 2003; McCutchen, 1996), and Perfetti's lexical quality hypothesis (2007). At the sentence level, knowledge of lexical morphology (Jarmulowicz & Taran, 2013), with its morpho-syntactic aspects, is likely evidence of exposure to and uptake of a variety of syntactic frames and academic vocabulary, and productive use of this knowledge could help a writer manage syntactic choices by assisting with the fluent change of verbs into nominalizations, or the reverse, via manipulation of suffixes. McCutchen and Stull's (2015) data suggested that children use their morphological skill not only to *retrieve* words they know but also to *construct* novel morphological forms to fit the developing syntax of their sentences (e.g., *solidize,* by analogy with *crystallize*). If students can manipulate words that they already know by altering suffixes, they may be better able to express their intended meaning more precisely. Thus, morphemes may serve as a bridge that relates the word level to the sentence level, with word-level manipulations assisting with sentence-level syntax.

Moderately-constrained measures of morpho-syntactic production**.** In their meta-analysis of writing instruction techniques, Graham and Perin (2007) reported that sentence combining instruction has good sized effects (*d*=0.50) on the quality of student writing. Nagy and Townsend (2012) suggested that while research supports sentence combining instruction that focuses on the subordination of clauses, it was yet unknown whether sentence combining had been successful in supporting the use of nominalization as a means to condense information. This process requires morphological derivations. They provide an example: *John investigated the problem. This led him to revise his theory.* By deriving a new form of the word *investigated* (in this case, deriving a nominalization), the two sentences can be condensed into one: *John's investigation of the problem led him to revise his theory.* To explore these word-level syntax manipulation skills, McCutchen, Stull, Herrera, Lotas, & Evans (2014) developed a sentence combining task in which sentences could be combined by deriving a different form of content words from the original sentences. This included the possibility of forming nominalizations, but also of changing a noun to an adjective (*ink* → *inky*), or an adjective to an adverb (*direct* → *directly*). Using this task, McCutchen & Stull (2015) found that a more traditional test of receptive morpho-syntactic awareness was uniquely predictive of the number of derivations a student made in the sentence combining task above word reading and receptive vocabulary.

Northey, McCutchen & Sanders (2016) administered the same morphological sentence combining task to a group of 5th and 8th graders. They found that, when controlling for grade, reading comprehension, and writing fluency, the total number of derivations the students made in combining sentences was uniquely predictive of the essay subtest from the *WIAT III Wechsler Individual Achievement Test*, *Third Edition* (Pearson, 2010). Their model predicted that for every standard deviation increase in morphological derivations, the essay score increased by a predicted 0.20 standard deviations. While any explanation of this relationship can only be speculative, the authors hypothesized that, following Berninger and Amtmann's (2003) *simple* *view of writing*, fluency in morpho-syntactic skills might allow for working memory resources to be allotted towards higher-order thinking processes, such as planning, which is important for any essay-writing task.

McCutchen et al. (2014) designed an intervention where elementary classroom teachers integrated instruction of the morphological aspects of science-related words into their fifth grade science curriculum (e.g., *erosion* and *conservation*). Student activities included both reading and writing tasks in analyzing the morphological components of these words and in practicing producing them (sentence completion and writing novel sentences using different word forms). At pre- and posttest, students completed the morphological sentence combining task, where students combined three or four short sentences into one longer sentence, and scores were based on the number of derived words in the final sentences. They found a significant effect of group and of pretest, as well as a significant interaction of group and pretest, suggesting that the treatment had a larger effect for students that scored lower at pretest. An illustrative qualitative analysis of sentence quality indicated that responses containing morphological derivations were more concise and contained more syntactic changes from the original sentences. Thus, while there is evidence that interventions can improve students' use of complex derivational word forms in a moderately constrained task, further analysis of word- and student-level characteristics could provide useful insight into specific effects of instruction.

Though this task is not widely used in the literature, it could be viewed as a moderately-constrained task of morphological production. That is, though students can derive morphological relatives—and the task is designed to elicit these types of word-level manipulations in service of larger syntactic changes—students are not required to change the words to successfully combine sentences. Like morphological production, this task requires students to hold information in their minds and manipulate it, though in this case they must hold larger units of information. They have more flexibility to manipulate the syntax and the morphological structure of the words as they like. Thus, this task provides a measure of morphological production which arguably taps into children's control over word- *and* sentence-level construction, as well as their flexibility to maintain meaning while changing structural aspects of words and syntax. This is an important aspect of academic language proficiency, as students must both build and reconstruct sentences using more complex words to express their ideas. When students can morphologically manipulate words in service of their syntax, do they?

Tests like these—tests used to measure students' linguistic abilities—are created using a specific set of words. Though these tests are created to capture variation between students, they also inevitably capture variation between words. Because some words are more difficult to learn than others, it is also important in this type of research to account for student skills by word characteristics. The next chapter will review the steps that have been taken to better model the complexities of student abilities as related to varying word characteristics.

CHAPTER 3

Modeling the Complexities of Language Processes

In linguistic research, researchers are often interested in testing knowledge of a specific category of words. In the construction of their measures, they carefully select a group of words that are representative of the larger population from which they are drawn. Consider, for example, Tyler & Nagy's (1989) investigation of the development of children's understanding of relational morphology. The authors developed a multiple-choice assessment where students saw either a frequent base word or a less frequent morphological derivation of that base word used in a sentence. Their task was then to pick a synonym that matched the meaning of the target word. If students could identify the high-frequency base within the more complex, less frequent form, they might have an easier time picking its synonym. For example, students saw one of the following sentences: *"I'm in a* celebratory *mood," Mary announced.* or *"I'm in the mood to* celebrate*," Mary announced.* The sentence was followed by the following options: *Mary felt like a) having a party, b) being alone, c) going to sleep, d) having a fight, or e) don't know.* If the student could recognize *celebrate* within the more complex and less frequent *celebratory*, they may have been more likely to choose option *(a) having a party*. This measures relational morphological knowledge, which is the ability to see the relationship between different words within the same morphological family. Not all base and complex word form pairs are created equal, however. Some word pairs have a less obvious semantic relationship between them because of their structural features, so the authors designed their measure to attempt to account for this difficulty.

To do this, the authors compiled a list of word that fit into one of two categories: words with neutral suffixes like *-ness, -er, -ize*, and *-ment*, and words with nonneutral suffixes, like *-ity*,

-*ify*, -*ous*, and -*ive* (Tyler & Nagy, 1989). Neutral suffixes attach to free morphemes, so they do not generally change the spelling or pronunciation of the base word they attach to (*social* → *socialize*). Nonneutral suffixes often attach to free morphemes. When you remove -*ify* from *beautify*, for example, you are left with *beaut*, which is not a word. Nonneutral suffixes can also change the stress (*advantage* → *advantageous*) and the spelling (e.g., *attend* → *attentive*) of the base they are attached to, making the base more difficult to identify. Lastly, the meanings of words with nonneutral suffixes do not often have a clear semantic relationship with their related base words. The meaning of *emergency*, for example, as related to the meaning of *emerge*, is not as clear as the semantic relationship between *art* and *artist*. This makes words with neutral suffixes characteristically different from words from words with nonneutral suffixes. Because nonneutral-suffixed words are different, and, as it turns out, generally more difficult (Tyler & Nagy, 1989) to learn than neutral-suffixed words, we can expect the patterns of student responses to differ by condition (suffix type).

This is true for many other word characteristics: some words are intrinsically harder to learn than other words, so we can expect student response patterns to differ based on word-specific characteristics. Models based on categorized variables like this can be limiting, however. When a measure is developed to test neutral- vs. nonneutral-suffixed words, for example, other factors must be controlled for within the measure. In the process of designing the measure, steps are taken to ensure that the only major difference between the two word conditions is the fixed effect of suffix type itself, so that any differences in student response patterns by condition can be attributed to the suffix type, and not to an unknown factor. Word frequency, for example, can be controlled across conditions. Tyler and Nagy (1989) developed their measure so that half the items were words with neutral suffixes, the other half were words with non-neutral suffixes, and the halves were balanced for word frequency.

While appropriate for certain research questions, these types of study designs have some major drawbacks. First, words are not always so easily categorized. Linguistic researchers are often interested in considering, for example, a word's frequency, bigram frequency (how common every 2-letter sequence within a word is), or phonological Levenshtein distance (how phonologically unique a word is). All of these characteristics have been theorized to play a role in how difficult a word is to learn. These types of variables, which have a wide range and vary considerably, are not suitable for dichotomization. Researchers also have an interest in exploring many possible word-level characteristics simultaneously rather than isolating one specific characteristic for study. The purpose of a study may be to discern which word-level characteristics matter when predicting various outcomes, so many word characteristics need to be simultaneously considered within a statistical model. Beyond that, there is also a theoretical interest in exploring interactions between continuous word-level variables and student-level variables. For example, while a word's frequency may play a role in whether a student can successfully read a multi-syllabic word, it may play a larger role for students with lower reading comprehension. Because words and their characteristics naturally vary, some words being intrinsically more difficult to learn than others, it is often within the interest of linguistic researchers to capture the full range of that variation within their statistical models of word learning.

It is also true that the difficulty of word learning will vary among students. Students' reading and writing knowledge plays a role in how they respond to word-specific tasks. Students who have a more advanced understanding of the orthographic and phonological shifts typically involved in affixing nonneutral suffixes will likely show response patterns more similar to those of the neutral-suffixed words, for example. Because many student-specific characteristics naturally vary among students, it is within theoretical interest to also consider student skills when investigating word-specific responses.

Recent methodological advances have allowed for this word and student variability to be better represented statistically, allowing researchers to harness the natural variation of both words and students to explore word-level skills. Reading researchers have begun exploring how word- and student-level characteristics contribute to word-specific outcomes (like reading morphologically-complex words) using cross-classified mixed models with random effects (Goodwin et al. 2013; Goodwin et al. 2014; Goodwin & Cho, 2016; Dobbs & Kearns, 2016; Kearns, Steacy, Compton, Gilbert, Goodwin, Cho, Lindstrom, & Collins, 2016) as suggested by methodologists (Baayen, Davidson, & Bates, 2008; Quené & van den Bergh, 2008). This allows for the simultaneous exploration of fixed effects describing words (frequency, morphological family size, etc.) and students (decoding skill, comprehension, etc.) on multiple word-level outcomes while treating both words and students as random effects. Treating words and students as random effects allows findings to be extended beyond the specific words and students in the investigation to their broader populations. The outcomes of these models are student responses to specific items rather than aggregated scores, which allows for consideration of the individual word- and student-level factors that contribute to a given student's response to a given word.

This body of research is framed by the lexical quality hypothesis (Perfetti, 2007), which provides a model for how word- and student-level variables interact to contribute to word learning.

Theoretical Reasons for Simultaneously Considering Word- and Student-Level Variables

Lexical processing. The lexical quality hypothesis (Perfetti & Hart, 2002; Perfetti, 2007) provides a model for how knowledge of words is constructed in, stored in, and retrieved from the mental lexicon. The nature of this knowledge building, storing, and retrieval is completely dependent upon the word being learned and the person doing the learning. It has three major components: 1) word knowledge can be broken down into understanding of phonological form, orthographic form, semantics, and morphosyntax; 2) individuals vary in their understanding of these different aspects of word knowledge (thus creating varying levels of mental representation of a lexical item); and 3) differences between individuals' quality of lexical representation have consequences for reading skill (Adlof & Perfetti, 2014). Thus, according to this hypothesis, the quality of a word's representation depends on the characteristics of the word, the knowledge of the student, and their interaction.

The mixed-effects models described previously are well-suited to exploring this interplay of word- and student-level variation for a few reasons. First, as previously explained, the lexical quality hypothesis rests upon the fact that the varying characteristics of words and learners interact to determine word-level knowledge and to play a role in broader reading outcomes. Additionally, it theoretically links the fields of vocabulary research and morphological awareness research by providing a model to explain word learning and storage, which involves both words as parts, and words as wholes.

***Words as parts*.** Knowledge of word parts—in addition to whole word knowledge—supports general language proficiency. Words are stored in the mental lexicon both as whole words and as morphemes (Schreuder & Baayen, 1995). Carlisle (2010), in her review of the field of morphological awareness research, describes how this works as a process of development. Children learn morphemes (bases and affixes) as they learn language. Morphemes and whole words are encoded as phonological units, and with use and exposure, become more well-represented. We see evidence of morphological encoding in children's early experimentation with novel word forms, such as adding the inflectional suffix *-ed* to make *goed* rather than using the past participle *went*. When the suffix *-ed* has its own separate representation (as a result of frequent and varied exposure), it can easily be applied to other verbs that fit in this pattern (Schreuder & Baayen, 1995). Because the lexicon is organized morphologically, morphological knowledge plays a role in more efficient word processing (Kuo & Anderson, 2006). When there are so many words to learn and store, it is likely more effective to store words in patterns, rather than as unique units. Thus, children with more advanced morphological awareness may be better at building and retaining representations of morphologically complex words.

***Words as wholes*.** Central to the lexical quality hypothesis is the idea that knowledge about a word can be partial (Frishkoff, Perfetti, & Collins-Thompson, 2010). I may be familiar with a word in that I've seen it before, or know, for example, that it has a negative connotation. But that does not mean that I could define that word, or spell it, or use it properly in writing or in spoken language. Exposure to a word used in various contexts and in slightly different ways will strengthen its lexical representation. With these exposures, the meaning of the word will be decontextualized (Stahl & Fairbanks, 2006) and the form of the word will be solidified in representation. Higher quality representations, therefore, allow for *flexible* but *precise* understanding of a word (Perfetti, 2007). As an example of the importance of both flexibility and precision in lexical quality, Perfetti (2007) presents the following two sentences: "You need a *record* of the transaction." and "They can't *record* the conversation." (p. 359). A high-quality representation of *record* allows for the precise pronunciation of both uses, as well as a flexible understanding of the word's meaning, given the different contexts it is presented in. Repeated and varying exposures to a word allow for these flexible but precise representations, as a high-quality representation allows for adjustment in interpretation given context and purpose.

Lexical quality and reading comprehension. Language comprehension relies on the efficiency of many simultaneous processes, including decoding, phonological processes, retrieval, memory, and automaticity (Perfetti, 2007). When a word has a high-quality representation in the mental lexicon, it can be retrieved more efficiently, making the way for easier comprehension. Perfetti argues that essential to this efficiency is knowledge about word forms and meanings. Thus, a given word's lexical quality is comprised of meaningful and pragmatic aspects of word knowledge. The meaningful aspects concern the word's meaning and how it is used (aspects of vocabulary knowledge), while the pragmatic aspects concern spelling, pronunciation, and morphological structure (aspects of morphological knowledge). These representations are built over time with multiple and various exposures with the word. Increased exposure to a word improves its representational quality both by decontextualizing it and by solidifying the representation of its form (Stahl & Fairbanks, 2006). Therefore, because word learning is tied to both the word and the student, it is important to account for both the word and the student when considering student word knowledge, as well as the effects of word-level instruction.

Statistical Reasons for Simultaneously Considering Word- and Student-Level Variables

In addition to accounting for some theoretically-selected fixed-effects of the tested word sample—characteristics specific to the words included in the sample—we also have an interest in treating words as a random effect. In linguistic research, tests are often developed with a specific sample of words, and that specific sample is tested across all participating subjects. That is, all students respond to the same set of words. If words are treated as fixed effects rather than as random effects, it is assumed that the tested words constitute the entire population of words the researchers wish to generalize their results to. If findings from the specific subset of words are generalized to the larger population which they represent, it is a statistical error, which Clark (1973) refers to as the "language-as-fixed-effect fallacy." Instead, the tested words must be treated as a random subset of the larger population. We do this to be able to generalize findings to all the possible existing words with these specific characteristics, rather than attributing findings only to the specific words included in the text. The words included in a measure are a systematic source of variation, and so ignoring this variation may lead to high Type I error rates (Baayen, Davidson, & Bates, 2008). Similarly, we wish to treat participants as a random effect. Participants within a study are thought to be a random sample of a larger population to which we wish to generalize our findings.

One common approach to analyzing data like these is to use repeated-measures analyses of variance (RM-ANOVA). Scores across similar items are aggregated at the student level (i.e., each student will have a total score for neutral suffixes and a total score for nonneutral suffixes) and condition (neutral or nonneutral) is included as the within-subjects factor. The random variance within these models is separated into two terms: one variance component attributed to the variance between participants, and one for the within-participant residuals (Quené & van den Bergh, 2008). This type of model allows us to account for both fixed effects (word condition: neutral or nonneutral suffix) and random effects (participants and words).

Methodologists have recently begun to advocate for a more advanced method to address these model structures with mixed-effects models, rather than to using univariate or multivariate approaches to RM-ANOVA. Mixed-effects models are more powerful than RM-ANOVA, and they are more robust to missing data (Quené & van den Bergh, 2008). We can also use a mixture of discrete and continuous variables as model predictors, as well as test binary outcomes. Binary outcomes are of interest for linguistic research, because researchers often want to test item- and student-level effects on whether students respond correctly to a specific word. The model outcome will be whether a student responds correctly to the specific word *gratify*, for example, receiving either a 1 or a 0. Though these types of data (binary outcomes) are not normally distributed, mixed-effects models can easily accommodate Bernouli distribution.

Another advantage to mixed-effects models is that by including both participants and items (words) as random effects, we can also consider fixed effects for both participants and items within the same model. For example, when considering student performance on neutral vs. nonneutral suffixes, we might want to consider student reading comprehension as a predictor, as well as the frequency of the specific word.

Additionally, with linguistic tests like the neutral vs. nonneutral suffix example, researchers have an interest in treating items as crossed random effects. This means that a specific word is treated as the same word across all students, rather than as words nested within students. This is important to do because the error terms for specific items are correlated (Quené & van den Bergh, 2008). That is, the way Student A responds to the word *gratify* will likely be more similar to the way Student B responds to the word *gratify* than the way Student A responds to the word *artist*. This is important because for linguistic items, many of the predictors we include within the model are intrinsically linked to the items. *Gratify* intrinsically has a nonneutral suffix. The fact that item 1, the word *gratify*, is the same word across all participants, is built into crossed random-effects models. Lastly, crossed random-effects provides correct estimates of the standard deviations characterizing the random effects (Baayen et al., 2008).

Cross-Classified Mixed Models with Random Effects Studies

Over the last few years, a small body of research has emerged that uses crossed-classified mixed models with random effects to investigate word-level student response outcomes. This research has explored word- and student-level effects on the reading of complex words (Kearns et al. 2016), or on student attempts to use a complex word in extended writing (Dobbs & Kearns, 2016). To my knowledge, only one study (Goodwin & Cho, 2016) has also considered the effects of an instructional intervention in addition to word- and student-level effects.

The lexical quality hypothesis and lexical production. Of this recent body of work, only one study considers the contribution of word- and student-level variation to language production. This study (Dobbs & Kearns, 2016) provides insight into how the lexical quality hypothesis might account for production of words in extended writing. Specifically, Dobbs and Kearns considered the factors that may contribute to whether students use certain academic words in their opinion essay writing. In an instructional intervention where classroom teachers taught general-use words from the Academic Word List (Coxhead, 1998), sixth, seventh, and eighth grade students learned five new complex words each week. The teaching of new words was embedded within broader curriculum material, which explored different high-interest topics on which students might have an opinion. Each of the new words was presented multiple times in different contexts for students to see and hear, and students were encouraged to use the words to express their opinions. At the end of each week, students were provided with a list of the five words from that week’s curriculum, and were encouraged to use those words in their response to a prompt to write an opinion essay.

Dobbs and Kearns’ main purpose in this study was to explore the word- and student-level characteristics that contributed to the likelihood of students using a specific word in their opinion essays. They hypothesized that willingness to use a more difficult word may be a reflection of students' lexical quality; that is, a student may be more likely to use a word if he or she has a high quality representation of it. As discussed previously, researchers have hypothesized that whether a student is willing or able to produce a word in expressive language requires a certain level of word-specific knowledge. While students could use a word of which they have partial knowledge, Dobbs and Kearns argue that more robust word knowledge may make a word more readily available for a student to retrieve while writing. Efficient retrieval of words and their features is especially important for developing writers, who must juggle many competing demands that have not yet many automatized as they write (McCutchen, Covill, Hoyne, & Mildes, 1994; McCutchen, 1996; Berninger & Amtmann, 2003). High-quality representations are not only dependent on student knowledge, as discussed previously. Some words, depending on various intrinsic characteristics, may be more easy to build representations of. For example, words that are imageable (that is, words that have meanings that can be mapped onto a mental image) are often easier to build representations of because their meanings are often more easily understood. Additionally, words that are encountered more frequently (words with high frequency) may be more likely to have high quality representations, because more frequent encounters allow for the building of more flexible but precise representations. Thus, to explore the interplay of both words and students on item responses, Dobbs and Kearns accounted for various student- and word-specific characteristics.

When accounting for both word- and student-level characteristics, Dobbs and Kearns (2016) found that proficiency on a language arts state achievement test (binary variable), English language proficiency (binary variable), and word frequency (continuous variable) were all uniquely predictive of whether students attempted to use a specific word. Bigram frequency (orthographic uniqueness), morphological family size, imageability (abstractness of a word's meaning) and phonological Levenshtein distance (phonological uniqueness of a word) were not uniquely predictive. The authors provide a few potential explanations for the unique effect of word frequency on word use. First, they hypothesize that high frequency words have more utility. Because the standard frequency index (SFI) provided by Zeno, Ivens, Millard, & Duuvuri (1995) is reflective of how often a word occurs across contexts (i.e., domain-specific words typically have lower SFIs), they argue that words with higher frequencies are used more often and more widely because they are more useful. Secondly, they hypothesize that the quality of students' lexical representations will be better for higher frequency-words because they are more likely to be encountered.

Additionally, Dobbs and Kearns (2016) provide some possible explanations for the student effects that uniquely predicted word use. Proficiency on a state test was a unique predictor of complex word use, as well as limited English proficiency status, which had a unique negative coefficient. They proposed that students who reached proficiency on the language arts state test were likely better readers and writers in general, and therefore that they were probably more skilled at learning new words. They also argue that in general, good writers also have good general language skills, which helps them to juggle the many demands that writing tasks call for. More ease in text generation may allow for more flexibility with word choice and incorporating different ideas in their writing.

With their study, Dobbs and Kearns (2016) provide a theoretical extension of the lexical quality hypothesis—which typically concerns processes related to reading skills—to writing processes. Their study investigated which factors contributed to a student's willingness to use a word in a loosely-constrained task (that is, in a task where students have freedom to construct text and use specific words as they see fit). Though the authors did assume that students would only use a word in their essays if they felt comfortable using them in a syntactically and semantically appropriate way, they did not consider how well the word was used. They also did not explore student-level factors that are more well-known to contribute to word-level and writing outcomes, such as reading comprehension. Additionally, *willingness* to use complex words is different from *ability* to use complex words. There are many other possible explanations for why students may not have used the complex words, as the authors suggest: greater task persistence, inhibitory control, or ability to follow directions. Despite these limitations, this study provides a unique contribution to our understanding of students' lexical production in their spontaneous writing.

The Current Study

This study builds upon Dobbs and Kearns' (2016) findings regarding students' lexical production in loosely-constrained writing tasks. In their study, Dobbs and Kearns did not assess how well a word is used in student writing. How well a word is used in context can be measured in many ways, one of which is to assess whether a derived word fits syntactically within the sentence the student produces. This study, assessing students' ability to morphologically manipulate a word to fit into an existing sentence or to revise existing sentences, provides insight into one measure of quality of language production (i.e., how well a word is used). By placing more constraint upon students by compelling them to use specific words, this study creates the opportunity to assess students' *ability* to control morpho-syntactic production in addition to their *willingness*. Based on Dobbs and Kearn's (2016) findings, this study will also examine word-level student responses to morphological tasks as predicted by base word standard frequency, prior morphological knowledge at pretest, and students' reading comprehension level. Moreover, though the Dobbs and Kearns findings were part of a larger intervention study, they did not specifically test intervention effects of lexical production against a control group; thus the present study will also contribute to the literature by testing the effect of classroom morphological instruction on students' word-level responses to morphological tasks.

The purpose of the current study is to investigate the role of relevant word- and student-level characteristics when assessing the effects of a 12-week instructional program on students' productive use of morphologically complex words in writing tasks. This addresses a few major limitations of the morphological awareness and vocabulary literature: 1) little is known about the underlying skills required for morpho-syntactic production, 2) few instructional interventions teaching word-level skills concern morpho-syntactic production outcomes, and 3) the studies that do consider morpho-syntactic production are often not designed to consider the range of variation of word- and student-level characteristics simultaneously.

By constructing models of the proximal production outcomes of our morphologically focused vocabulary instruction, this study attempts to provide a preliminary exploration of what accounts for students' control over production morpho-syntactic forms, as well as the possibilities provided by an embedded instructional intervention.

Research Questions

1) What are the *direct effects* of classroom morphological instruction, student pretest word-level information knowledge, student pretest comprehension, and prompt word frequency level on students' syntactically and semantically appropriate production of morphologically complex words, for two different levels of text constraint (more and less constrained tasks)?
2) What are the *unique effects* of each of the factors above for each of the two levels of text constraint?

Method

Participants

Teachers. Two cohorts of fourth and fifth grade teachers were recruited from a large metropolitan area in the Pacific Northwest. A total of twenty-two classrooms participated in the study: 12 in the first cohort (after two dropped before instruction began due to scheduling conflicts) and 10 in the second. Class sizes ranged from 19 to 34 students per class, with a mean class size of 25.5. For each cohort, teachers were randomly assigned to either a morphological classroom instruction condition (treatment), or to a business-as-usual (control) condition, in which teachers implemented their typical literacy instruction. Teacher demographic characteristics are described in Table 1.

Table 1. Teacher characteristics

	Treatment		Control	
	N	*%*	*N*	*%*
Male	3	27%	3	25%
Education				
Bachelor's	10	100%	12	100%
Master's	8	82%	8	67%
School Type				
Public	4	40%	6	50%
Private	6	70%	6	50%
Years of Experience				
1-5	3	27%	3	25%
6-10	3	27%	1	8%
11-15	3	27%	2	17%
>15	2	18%	6	50%

Note. Treatment n = 10, Control n = 12. There were no significant between-group differences on any characteristics.

Students. Of the 561 fourth and fifth grade students in participating classrooms, 499 students in 22 classrooms participated in all phases of the study (autumn pretest and spring posttest). Student demographic characteristics are provided in Table 2.

Table 2. Student characteristics

	Treatment		Control	
	N	*%*	*N*	*%*
Male	113	48%	131	49%
Minority	52	22%	66	25%
Hispanic Heritage	20	9%	23	9%
ELL	3	1%	3	1%
SPED	16	7%	18	7%
4th Grade	124	53%	122	46%

Note : Treatment *n* = 233; Control *n* = 266. ELL and SPED are reported by official school designation. There were no significant between-group differences on any characteristics.

Classroom Instruction Intervention

The intervention consisted of 12 weeks (20-40 minutes per day, 4 days per week) of instruction implemented by the classroom teacher. The learning objectives for the intervention were varied, focused on a broader understanding of how words work rather than on a specific skill for growth (spelling, word reading, vocabulary, etc.), as is often seen in interventions teaching morphology (see Goodwin & Ahn, 2013, for a review). We were interested in not only supporting student growth in receptive and productive knowledge for specific vocabulary, but more importantly in promoting a broader understanding of the phonological, orthographic, semantic, and syntactic patterns of the English language. Therefore, classroom teachers taught specific word families (typically a base word and its common related derived forms; described below) within the context of teaching broader patterns of language structure and function. Teachers implemented instruction with the following learning objectives: 1) break down words into morphemes, 2) learn the meanings of common prefixes and suffixes, 3) learn ways that morphology impacts spelling, 4) gain a general awareness that words can be related in predictable ways, 5) use morphological analysis to infer meanings of unfamiliar words, and 6) use morphologically complex vocabulary to fit a sentence- or text-level writing purpose. These learning objectives range from gaining insight into the morphological structure of words to applying that insight when encountering and producing complex words in text.

Instructed word families. A total of 40 word families were chosen for instructional focus. Families were typically introduced as a base word and two to four related derivations (although additional family members often came up in discussions and in student work). Instructional units were words, rather than morphemes, to account for words with bound morphemes (e.g., *produce*). Although explicit attention to Greek- and Latinate-bound morphemes was not emphasized within the curriculum itself, teachers were free to draw attention to these word features if they felt it would help their instruction. Word families introduced related words whose semantic relationship was more obvious (Nagy & Anderson, 1984). The word family for *sympathy*, for example, included *sympathizer*, *sympathetic*, and *sympathetically*, but not *empathy*, *pathology*, or *pathetic*. SFIs of base words ranged from 31.8 to 62.4, with a median of 54.7. Derivations within the family varied more: some words were more content-specific (*inoperable*; *SFI*=13.2), others were incredibly common (*direction*; *SFI*=61.1). Words that were chosen for instruction had relevance to instructional topics (explained below), a range in semantic familiarity where the meanings of base word may be more familiar (*adventure*) or less familiar (*prohibit*) to students, and a mixture of phonologically and orthographically opaque and transparent word pairings, to allow for exploration of how spellings and pronunciations may change within morphological families. Instructional activities covered a wide range of prefixes and suffixes, focusing on their phonological, orthographic, semantic, and grammatical properties.

Curriculum format and activities. Throughout the intervention, students participated in lessons within a 4-day cycle (see Table 3 for a description of daily lesson plans). Each new week, a new group of word families was introduced, and students repeatedly encountered and practiced speaking and writing different forms of the word families by reading passages, responding to comprehension questions, practicing sentence-level writing, and responding to prompts with extended writing. Embedded within the lessons were morphology-specific tasks, including morphological analysis (identifying morphemes in morphologically complex words), synthesis (building words from parts), and oral and written production of derivations in context (for a review of typical tasks and content in morphological interventions, see Bowers, Kirby & Deacon, 2010). Additionally, though the curriculum activities were built with a morphological lens, teachers supported and emphasized morphological insight throughout the implementation of the curriculum—especially through feedback to student responses and questions.

Table 3. Typical Weekly Lesson Plan

Day 1: Engage with word families	**Day 2**: Encounter words in text	**Day 3**: Produce language	**Day 4**: Write longer text
• Introduce new word families and theme of the week	• Practice close readings with academic text	• Practice using words in short writing activities	• Use different morphological forms of vocabulary words in longer text production
• Write words and learn meanings	• Interpret words in context	• Practice using different morphological forms in correct syntax	• Learn and practice the traits of different written school genres
• Make connections between vocabulary words and student experiences	• Consider multiple meanings and usages of words	• Practice using complex words in place of more common words/phrases	• Practice using other complex words in text production, including transition words
• Learn parts of speech and affixes	• Use vocabulary words in responses to text-based questions	• Produce more dense, complex sentences	• Review theme of the week
• Bring attention to word roots		• Practice revising sentences into more exact language	

Morphological awareness instruction was embedded into broader reading and writing activities. The curriculum was organized by instruction on the three school genres emphasized in the CCSS: the first four weeks focused on narrative texts, the second four weeks on informative texts, and the last four weeks on opinion texts (National Governors Association and Council of Chief School Officers, 2010). Teachers introduced word families within the context of three different themes as well: neighborhoods, space exploration, and Lewis and Clark. For example, students were introduced to the word family *function* (malfunction, functional, etc.) on a space exploration week, and to the word family *reside* (residential, residence, etc.) on a neighborhoods week (see Table 4 for word families taught across themes and genres). Each week of instruction, teachers introduced four or five new word families. Weeks four, eight, and twelve were used for reviewing vocabulary, encouraging playful use of words and morphemes, and highlighting morphological insights and strategies.

Table 4. Instruction Overview

Lesson Theme	**Neighborhoods**	**Space Exploration**	**Lewis & Clark**	**Review**
Narrative	**Week 1** commonly directly distinctive socialize	**Week 2** differentiate operational suspiciously adventurer	**Week 3** descriptive documentation exploratory vividly	**Week 4**
Informative	**Week 5** collaboration defiantly prohibitive sympathize	**Week 6** adaptation imitation experimental malfunction	**Week 7** valuable cleverly emotional translation receptive	**Week 8**
Opinion	**Week 9** division management residential sensible various	**Week 10** technological financial cooperate continuously advantageous	**Week 11** unimaginable secrecy intentionally memorable determination	**Week 12**

Measures

Norm-referenced measures were used to assess students' vocabulary, reading, and comprehension skills at pretest only in autumn of the academic year. Researcher-developed identical measure forms to assess two types of morphological knowledge (more and less constrained) were administered both at pretest and posttest, with posttesting occurring in spring of the academic year.

Comprehension. The Passage Comprehension subtest of the *Woodcock Johnson III Tests of Achievement* (Woodcock, McGrew, & Mather, 2001) is an oral cloze task, asking students to orally provide one word to appropriately complete a short passage. The test manual (McGrew, Schrank, & Woodcock, 2007) reports split-half reliability of .89 for age 10 and .83 for age 11.

Morphological production (MP). As a highly constrained measure of written production, this morphological production task (see Carlisle, 2000; McCutchen et al., 2009) invites students to change a base (prompt) word so that it fits both syntactically and meaningfully in a short sentence. The prompt words provided for each task (referred to as "base" words) were simple words to which multiple affixes might be added. However, these base words were not necessarily single-morpheme root words. For example, the provided base word for the derivation *emotional* is *emotion* (rather than *emote*), because *emotion* has a higher-frequency and is more likely to be familiar to students than *emote*.

Students were given examples and instructed to complete the sentence by changing the underlined word (e.g., <u>memory</u>: The party was _____.) Points were given for every syntactically- and meaningfully-correct morphological derivation: *memorable, remembered*, and *memorial* were given points; *memorized*, *memorific*, and *a memory/memories* were not. Phonological approximation misspellings were awarded points; morphological inventions (e.g., *memorific*), were not.

Of the 40 total items, 20 items contained instructed word families, and 20 items contained non-instructed word families. Word frequency was controlled across word type (instructed vs. non-instructed). Additionally, instructed and non-instructed words were balanced by the orthographic and phonological relationship between the stem word and the target derivation. That is, for half of the items, the shift between the stem and target words was orthographically and phonologically transparent, meaning that there were no spelling or pronunciation changes in the stem in the form of the derivation (*agree* → *agreement*). The other half of the items included both an orthographic and phonological shift from the stem word to the target word (*mystery*→*mysterious*). Students were tested both at pre- and posttest. The estimated internal consistency (Cronbach's alpha) was 0.92.

Morphological sentence combining. As a moderately constrained measure of written production, this sentence combining task, adapted from McCutchen and colleagues (2014), was designed to measure students' flexible and precise use of word forms and sentence structures. Students were instructed to combine three short kernel sentences into one longer sentence without using the word "and". These kernel sentences were constructed such that students could combine them effectively without requiring clausal subordination (which is typical of most sentence combining tasks). Instead, the structure of the kernel sentences invited students to make syntactic manipulations at the word level. The following is one item from the task:

The boy described his plan.

He often imagines.

The way it was described was vivid.

Each set of kernel sentences included two content words whose grammatical categories could easily be changed to reduce an independent clause to a word. Students could, for example, change the verb *imagine* to the adjective *imaginative*, and the adjective *vivid* in to the adverb *vividly*: The *imaginative* boy *vividly* described his plan. They could reduce an independent clause to a noun or adverb phrase, thereby increasing the density of another independent clause: The *boy who often imagines* described his plan *in a vivid way*. Students could also, in some situations, appropriately infer causal or ordinal relationships between clauses, and create a final sentence by using subordinating conjunctions to connect independent and dependent clauses: The way the boy described his plan was vivid *because he often imagines*. Each of these responses successfully combines the sentences in a way that maintains the general meaning of the original sentences and is grammatical.

Our interest in administering this task was in observing students' flexible and precise use of complex morphological forms, but also in their flexible and precise manipulation of syntax in sentence-level writing. Thus, this task was scored in two ways: as the number of derivations used during sentence combining (which is the score used for these analyses), and as the number of derivations during sentence composition (which focuses more on quality of the sentence generated).

The sentence combining derivations score reflects the number of morphological derivations the students make that fit both syntactically and semantically within their sentences. The morphological derivations score is not a measure of writing quality; it is, rather, meant to be used as a measure of students' flexible but precise word-level manipulations in service of sentence-level syntactic revision. Points were given for every syntactically- and semantically-appropriate morphological derivation used within student responses. As a reminder, a derivation is identified by a change in grammatical category (e.g., *imagines* → *imaginative* or *imaginary*), which usually requires changing the suffix to create a new word. Inflectional changes, which signal tense changes (if the base is a verb; e.g., *imagines* → *imagined*) or plurality changes (if the base is a noun; e.g., full of *terror* → full of *terrors*), did not receive points. Items were designed to include two base words (unidentified to the students) that could potentially be changed to combine the sentences. Of the 16 total target words, 8 were instructed base words, and 8 were non-instructed base words. Importantly, students were not required to make derivations to be successful on this task, so reported total scores should not necessarily be interpreted as a proportion of correct responses. Instead, they can be thought of as a descriptive account of the strategies students took up to revise the existing syntax.

The sentence composition score, alternatively, is considered a measure of sentence combining quality. The score reflects three major criteria: whether all of the essential information was included, whether the sentence was complete (not a run-on or fragment), and whether the sentence more or less maintained the meaning of the original sentences. This score is thought to reflect the quality of sentence combining regardless of morphological manipulations. Zero-order correlations from pretest morphological sentence combining and sentence composition reflect a significant positive relationship between scores, ($r = 0.64$, $p < .001$). Thus, while the morphological derivations score does not represent sentence writing quality, the two traits are well correlated.

Base word frequency was balanced between instructed and non-instructed words. The estimated internal consistency (Cronbach's alpha) at pretest was 0.79, and inter-rater reliability (Pearson's *r*) was estimated at 0.99.

Base Word Characteristics for Morphological Tasks

The findings of the recent body of work using mixed models to explore student responses to different word-level tasks suggest that both word frequency and word transparency play a role in determining students' ability to read complex morphological derivations (Goodwin et al. 2013; Goodwin et al. 2014; Kearns et al., 2016) or willingness to use complex academic words in their essay writing (Dobbs & Kearns, 2016). Specifically, Goodwin et al. (2013) studied both word- and student-level characteristics as they relate to morphologically complex words. They found that a student's ability to read the derived form was predicted by their ability to also read the base form of the word (e.g., reading *logically* and *logic*), but a significant root-reading by phonological and/or orthographic transparency interaction implied that root word reading was only predictive for phonologically and orthographically transparent word pairs. Kearns et al. (2016), when taking into account both word- and student-level characteristics predicting morphologically complex word reading, found that root word familiarity (whether the student was familiar with the spoken form of the word or not), root word recognition (word reading), word frequency, and phonological and/or orthographic transparency of the complex word predicted successful reading of morphologically complex words for both typically-achieving students and students with reading difficulties. Lastly, Dobbs and Kearns (2016) found that word frequency uniquely predicted whether students attempted to use a complex academic word in their essay writing, while no other word-level predictors were significant. Because word frequency and transparency were relevant word-level characteristics across these studies, they will be considered here.

Base word frequency. The frequency of the provided base (prompt) word is reported as the SFI (Standard Frequency Index) from *The Educator's Word Frequency Guide* (Zeno, Ivens, Millard, & Duvvuri, 1995). The frequencies are calculated from a corpus of school printed texts for grades K-12. SFI is a logarithmic transformation of the U statistic, which represents frequency per million words, weighted by frequency across contexts. SFIs were standardized (z-scores) within each task.

Word Shifts. The MP task is comprised of 20 base words requiring a transparent shift to the target and 20 base words requiring an opaque shift to the target. A pair of related words have a transparent shift when neither the spelling nor the pronunciation of the base word changes in its derivation (e.g., *paint* → *painter*), whereas a pair of related words have an opaque shift when either the spelling or the pronunciation (or both) of the base word changes in its derivation (e.g., *wise* → *wisdom*). Word shifts were effect coded: 1 for transparent shift, and -1 for opaque shift.

For Sentence Combining Derivations, because we expected students to respond with a variety of derivations related to a single prompted word, we did not consider whether shifts between prompt and response words might be transparent or opaque when selecting words to include in the measure. Therefore, this task was not built with the intent of controlling for shift type.

Data Analysis Plan

Consistent with recommendations by Baayen, Davidson, and Bates (2008), I adopted a cross-classified multilevel modeling (MLM) analysis approach that takes into account the dependencies in item responses due both to subjects (students) and items (target words); in addition, we also accounted for the partial nesting of the data as students were nested within classrooms and we wished to test classroom morphological instruction (treatment) on item responses. In summary, item responses (Level 1) were nested within both words as well as students (words and students fully crossed at Level 2), but only students were specified as nested within classrooms (Level 3). Because the item-level responses were binary, the outcomes were modeled with a log-linear link function, similar to logistic regression models (i.e., all coefficients in log-odds, or logits).

All models were estimated using the **lmer** function of the **lme4** package in *R* 3.3.3. For each outcome, I estimated three models. First, the intercept-only model (M0) was estimated to obtain an estimate of the intraclass correlations (degree of dependencies in the data). Second, a *direct* effects (M1) model was estimated *for each predictor* in order to understand which predictors had any relationship to the outcome, controlling for the hierarchical data structure.

Finally, a *unique* effects model (M2) was estimated to determine the non-overlapping contributions of each predictor to the outcome (i.e., all predictors entered simultaneously within the same model), controlling for the hierarchical data structure. For ease of interpretation, all categorical predictors were effect coded (i.e., classroom treatment condition = 1 and control = -1) and all continuous variables were standardized (*z*-scores; specifically, we standardized student comprehension scores across grades, and the word-level standard frequency index across target words).

Coefficient *t*-tests at the word-level were based on *df* = number of words – number of coefficients tested at that level; similarly, coefficient *t*-tests at the student-level were based on *df* = number of students – number of coefficients tested at that level.

Preliminary analyses indicated that there were no significant differences between Cohorts 1 and 2 on student demographics, nor on student pretest scores, and as such, cohort was not considered in the outcomes models. Additional analyses showed that the treatment group was lower than the control group on Sentence Combining Derivations at pretest (mean difference of 0.91, favoring control group). No other pretest differences by treatment condition were present; however, respective pretests were used in all outcomes models to control for these pretest differences (despite random assignment) and to maximize statistical power.

RESULTS

Student-Level Descriptive Statistics

Disaggregated descriptive statistics for both outcome measures (Morphological Production and Sentence Combining Derivations) are provided in Tables 5 and 6. Results suggest that both groups made gains from pre- to posttest on both outcomes and for both instructed and non-instructed words. Raw score means for pretest Comprehension show fourth grade students across treatment condition performed similarly (treatment: *M* = 30.82, *SD* = 3.59; control: *M* = 30.92, *SD* = 3.37). Fifth grade students across classrooms also performed similarly (treatment: *M* = 31.83, *SD* = 3.21; control: *M* = 32.36, *SD* = 3.71). Table 7 reports zero-order correlations for the student-level variables at pre- and posttest, disaggregated by treatment group. All predictors were positively correlated with the outcomes as well as with each other. Relations among the various student-level predictors ranged in magnitude from .33 to .80.

Table 5. Student-level descriptives for Morphological Production

		Instructed Words		Non-Instructed Words	
		Treatment	Control	Treatment	Control
		M *(SD)*	*M* *(SD)*	*M* *(SD)*	*M* *(SD)*
4th Grade	Pre	10.79 (4.25)	11.01 (3.89)	10.73 (3.95)	10.71 (3.94)
	Post	17.07 (2.82)	13.34 (3.66)	13.99 (3.04)	12.57 (3.65)
	Gains	6.28	2.33	3.26	1.86
5th Grade	Pre	12.72 (3.89)	13.10 (4.46)	12.28 (3.68)	12.57 (4.27)
	Post	17.85 (2.06)	14.53 (4.18)	14.94 (2.90)	13.85 (4.17)
	Gains	5.13	1.43	2.66	1.28

Note. 4th grade Treatment *n* = 124; 4th grade Control *n* = 122; 5th grade Treatment *n* = 109; 5th grade Control *n* = 144. Raw score means and standard deviations reported. 20 Instructed words and 20 Non-Instructed words total.

Table 6. Student-level descriptives for Sentence Combining Derivations

		Instructed Words		Non-Instructed Words	
		Treatment	Control	Treatment	Control
		M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
4th Grade	Pre	1.34 (1.44)	1.52 (1.42)	1.72 (2.05)	1.72 (2.01)
	Post	3.99 (2.16)	2.52 (1.70)	4.22 (2.58)	3.02 (2.45)
	Gains	2.65	1.00	2.50	1.30
5th Grade	Pre	1.65 (1.44)	2.01 (1.62)	1.98 (2.01)	2.53 (2.20)
	Post	4.36 (2.28)	3.24 (1.95)	4.54 (2.24)	3.88 (2.49)
	Gains	2.71	1.23	2.56	1.35

Note. 4th grade Treatment *n* = 124; 4th grade Control *n* = 122; 5th grade Treatment *n* = 109; 5th grade Control *n* = 144. Raw score means and standard deviations reported. 8 Instructed words and 8 Non-Instructed words total.

Table 7. Zero-order correlations between student-level measures

	1.	2.	3.	4.	5.	6.	7.
Pretest							
1. Comprehension	--	0.65	0.55	0.64	0.66	0.52	0.54
2. MP	0.65	--	0.58	0.79	0.80	0.56	0.51
3. SC	0.45	0.49	--	0.55	0.56	0.66	0.69
Posttest							
4. MP Instructed Words	0.48	0.64	0.33	--	0.83	0.56	0.53
5. MP Non-Instructed Words	0.63	0.76	0.50	0.73	--	0.56	0.56
6. SCD Instructed Words	0.52	0.56	0.57	0.51	0.61	--	0.68
7. SCD Non-Instructed Words	0.51	0.56	0.60	0.46	0.61	0.72	--

Note . *N* = 499; Treatment *n* = 233, Control *n* = 266. Treatment classrooms below the diagonal. MP = Morphological Production; SCD = Sentence Combining Derivations.
All correlations are significant at p <.001.

Item-Level (Word) Descriptive Statistics

Word characteristics across both outcome measures are reported in Table 8. Table 9 presents item difficulty (based on Classical Test Theory; means by item) scores for Morphological Production; Sentence Combining Derivations item difficulty can be found in Table 10. For Morphological Production at pretest, the mean difficulty of instructed words and non-instructed words was 0.59 and 0.58, respectively. For Sentence Combining Derivations at pretest, the mean difficulty of instructed and non-instructed words was 0.21 and 0.25, respectively. Averaged item means disaggregated by treatment condition suggest that at posttest, items with instructed words on both outcomes were consistently less difficult for students in treatment classrooms, while items with non-instructed words appeared to only be slightly less difficult, on average. For Morphological Production, item means disaggregated by shift type (phonologically and orthographically transparent or opaque), instruction type, and treatment condition, suggest that while opaque words were more difficult for all students than transparent words, instructed words were less difficult for students in treatment conditions regardless of shift type.

Table 8. Word characteristics

Base Word	Part of Speech	Instruct Type	Base Word SFI	U	Morph. Family Size	In MP	In SCD
accept	VB	N-I	57.1	51.00	15	1	
achieve	VB	N-I	54.5	28.00	13	1	
adapt	VB	I	49.9	9.00	21	1	
advantage	NN|VB	I	56.7	47.00	8	1	
adventure	NN	I	53.3	21.00	13		1
agree	VB	N-I	57.6	57.00	17	1	
assign	VB	N-I	44.5	2.00	12	1	
circle	NN|VB	N-I	58.2	65.00	23	1	
clear	JJ|VB|RB	N-I	62.6	180.00	27	1	
clever	JJ	I	52.6	18.00	5	1	1
collaborate	VB	I	31.8	0.15	7		1
construct	VB	N-I	49.1	8.00	24	1	
courage	NN	N-I	53.4	21.00	17		1
create	VB|NN	N-I	56.5	45.00	14	1	1
defy	VB	I	42.1	1.00	7		1
depend	VB	N-I	57.5	56.00	17	1	
describe	VB	I	59.1	81.00	11	1	
direct	JJ|VB|NN|RB	I	58.9	78.00	28	1	
emotion	NN	I	50.3	10.00	7	1	
exclude	VB	N-I	41.9	1.00	15	1	
experiment	NN	I	56.6	46.00	10	1	1
explode	VB|NN	N-I	47.1	5.00	12	1	
explore	VB|NN	I	53.9	24.00	13	1	
finance	NN|VB	I	49.6	9.00	9	1	
function	NN|VB	I	57.8	59.00	20	1	
history	NN	N-I	62.4	174.00	18		1
imagine	VB	I	58.9	78.00	16		1
imitate	VB	I	48.0	6.00	9	1	
impress	VB	N-I	45.2	3.00	19		1
intend	VB	I	49.4	8.00	18	1	
invent	VB	N-I	49.5	8.00	10	1	
investigate	VB	N-I	50.9	12.00	11		1
manage	VB	I	53.2	21.00	15	1	
memory	NN	I	56.9	48.00	21	1	
miracle	NN	N-I	49.9	9.00	5	1	
mystery	NN	N-I	54.8	30.00	5	1	
nutrient	NN	N-I	45.3	3.00	14		1
observe	VB|NN	N-I	55.6	36.00	19	1	
operate	VB	I	55.2	32.00	32	1	
pride	NN|VB	N-I	54.7	29.00	9		1

Table 8. Continued

Base Word	Part of Speech	Instruct Type	Base Word SFI	U	Morph. Family Size	In MP	In SCD
produce	VB	N-I	62.0	158.00	32	1	
reflect	VB	N-I	53.7	23.00	14	1	
secret	JJ|NN	I	58.1	64.00	7	1	
sense	NN|VB	I	61.6	145.00	62		1
severe	JJ	N-I	54.0	25.00	9	1	
sincere	JJ	N-I	46.5	4.00	6	1	
suspect	NN	I	51.3	13.00	13	1	
sympathy	NN	I	49.5	8.00	11	1	
terror	NN	N-I	52.3	17.00	20		1
translate	VB	I	45.8	3.00	12	1	
visual	JJ|NN	N-I	51.6	14.00	21	1	
vivid	JJ	I	49.9	9.00	3	1	1

Note. VB = verb; NN = noun; JJ = adjective, RB = adverb. I = Instructed word; N-I = Non-instructed word. SFI = Standard Frequency Index, from Zeno et al. (1995). U = weighted instances per million, from Zeno et al. (1995).

Table 9. Item difficulty: Morphological Production

Instructed Words	SFI	Pretest All *n* = 499 *M* (*SD*)	Posttest Treatment *n* = 233 *M* (*SD*)	Posttest Control *n* = 266 *M* (*SD*)	Non-Instructed Words	SFI	Pretest All *n* = 499 *M* (*SD*)	Posttest Treatment *n* = 233 *M* (*SD*)	Posttest Control *n* = 266 *M* (*SD*)
Opaque Instructed					Opaque Non-Instructed				
advantage	56.7	0.03 (.16)	0.73 (.44)	0.04 (.19)	circle	58.2	0.26 (.44)	0.34 (.47)	0.37 (.48)
describe	59.1	0.59 (.49)	0.87 (.33)	0.69 (.46)	clear	62.6	0.20 (.40)	0.45 (.50)	0.35 (.48)
finance	49.6	0.30 (.46)	0.75 (.43)	0.44 (.50)	create	56.5	0.84 (.36)	0.95 (.22)	0.88 (.33)
imitate	48.0	0.56 (.50)	0.86 (.34)	0.71 (.45)	exclude	41.9	0.44 (.50)	0.64 (.48)	0.55 (.50)
intend	49.4	0.39 (.49)	0.82 (.39)	0.48 (.50)	explode	47.1	0.88 (.33)	0.97 (.18)	0.91 (.29)
memory	56.9	0.69 (.46)	0.97 (.18)	0.75 (.44)	miracle	49.9	0.36 (.48)	0.47 (.50)	0.45 (.50)
operate	55.2	0.86 (.35)	0.97 (.17)	0.92 (.27)	mystery	54.8	0.75 (.43)	0.82 (.39)	0.80 (.40)
suspect	51.3	0.47 (.50)	0.71 (.45)	0.58 (.49)	observe	55.6	0.78 (.41)	0.92 (.27)	0.86 (.34)
sympathy	49.5	0.67 (.47)	0.91 (.29)	0.78 (.42)	produce	62.0	0.25 (.44)	0.49 (.50)	0.36 (.48)
translate	45.8	0.48 (.50)	0.75 (.43)	0.61 (.49)	severe	54.0	0.05 (.22)	0.10 (.30)	0.12 (.32)
Opaque Totals	52.2	0.50 (.47)	0.83 (.35)	0.60 (.42)	Opaque Totals	54.3	0.48 (.40)	0.61 (.38)	0.57 (.41)
Transparent Instructed					Transparent Non-Instructed				
adapt	49.9	0.33 (.47)	0.78 (.42)	0.55 (.50)	accept	57.1	0.47 (.50)	0.61 (.49)	0.48 (.50)
clever	52.6	0.89 (.31)	0.99 (.11)	0.92 (.27)	achieve	54.5	0.90 (.30)	0.96 (.20)	0.92 (.27)
direct	58.9	0.69 (.46)	0.90 (.30)	0.84 (.37)	agree	57.6	0.78 (.41)	0.86 (.35)	0.89 (.31)
emotion	50.3	0.73 (.44)	0.91 (.29)	0.79 (.41)	assign	44.5	0.89 (.31)	0.97 (.18)	0.93 (.25)
experiment	56.6	0.54 (.50)	0.77 (.42)	0.67 (.47)	construct	49.1	0.39 (.49)	0.64 (.48)	0.52 (.50)
explore	53.9	0.96 (.20)	0.99 (.09)	0.95 (.21)	depend	57.5	0.51 (.50)	0.69 (.46)	0.67 (.47)
function	57.8	0.54 (.50)	0.82 (.38)	0.68 (.47)	invent	49.5	0.83 (.38)	0.92 (.27)	0.85 (.36)
manage	53.2	0.76 (.42)	0.95 (.21)	0.85 (.36)	reflect	53.7	0.50 (.50)	0.77 (.42)	0.59 (.49)
secret	58.1	0.75 (.44)	0.88 (.32)	0.84 (.37)	sincere	46.5	0.74 (.44)	0.89 (.31)	0.86 (.35)
vivid	49.9	0.66 (.47)	0.95 (.23)	0.82 (.38)	visual	51.6	0.73 (.44)	0.88 (.33)	0.81 (.39)
Transparent Totals	54.1	0.68 (.46)	0.89 (.28)	0.79 (.38)	Transparent Totals	52.2	0.67 (.42)	0.82 (.35)	0.75 (.39)
All Words Totals	53.2	0.59 (.46)	0.86 (.31)	0.68 (.40)	All Words Totals	53.3	0.57 (.41)	0.70 (.37)	0.65 (.40)

Note . Item difficulty based on Classical Test Theory (means by item). 20 Instructed Words (10 Opaque, 10 Transparent); 20 Non-Instructed Words (10 Opaque, 10 Transparent).

Table 10. Item difficulty: Sentence Combining Derivations

		Pretest	Posttest				Pretest	Posttest	
		All	Treatment	Control			All	Treatment	Control
		n = 499	*n* = 233	*n* = 266			*n* = 499	*n* = 233	*n* = 266
	SFI	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)		SFI	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Instructed Words					Non-Instructed Words				
adventure	53.3	0.36 (.48)	0.73 (.45)	0.64 (.48)	courage	53.4	0.37 (.48)	0.68 (.47)	0.59 (.49)
clever	52.6	0.30 (.46)	0.56 (.50)	0.43 (.50)	create	56.5	0.30 (.46)	0.64 (.48)	0.56 (.50)
collaborate	31.8	0.08 (.27)	0.46 (.50)	0.25 (.43)	history	62.4	0.30 (.46)	0.65 (.48)	0.53 (.50)
defy	42.1	0.03 (.18)	0.24 (.43)	0.07 (.26)	impress	45.2	0.15 (.36)	0.34 (.47)	0.25 (.43)
experiment	56.6	0.18 (.38)	0.55 (.50)	0.32 (.47)	investigate	50.9	0.15 (.36)	0.45 (.50)	0.29 (.46)
imagine	58.9	0.24 (.43)	0.58 (.50)	0.43 (.50)	nutrient	45.3	0.22 (.42)	0.55 (.50)	0.43 (.50)
sense	61.6	0.10 (.30)	0.50 (.50)	0.22 (.42)	pride	54.7	0.24 (.43)	0.50 (.50)	0.41 (.49)
vivid	49.9	0.36 (.48)	0.54 (.50)	0.56 (.50)	terror	52.3	0.29 (.45)	0.56 (.50)	0.43 (.50)
Total	50.9	0.21 (.37)	0.52 (.48)	0.37 (.44)		52.6	0.25 (.43)	0.55 (.49)	0.44 (.48)

Note. Item difficulty based on Classical Test theory (means by item). 8 Instructed words, 8 Non-Instructed words.

Partially Nested Cross-Classified Model Results

Morphological Production. Model results for the morphological production task are provided in Tables 11 and 12 for instructed and non-instructed words, respectively. Recall that Model 1 fitted the unconditional model—in other words, estimating the mean log-odds (logits) of the probability of correctly responding to the item, controlling for item, student within classroom, and classroom random effects. For instructed words, the intercept estimate was 1.94 logits (which translates to a mean predicted probability of correct response of 87% across the entire sample). There was also substantial variability due to items (s^2 = 1.54), students within classrooms (s^2 = 1.28), and classrooms (s^2 = 0.66). For non-instructed words, the intercept estimate was 1.32 logits (which translates to a 79% probability of correct response across the entire sample). Again, there was substantial variability due to items (s^2 = 2.79), students within classrooms (s^2 = 1.44), and classrooms (s^2 = 0.09).

Next, *direct* effects models (Model 2) were conducted to estimate the direct relationship between each of the fixed predictors and the likelihood of correctly deriving a word. This set of models enabled me to understand the relationship between each of the predictors and the outcome while controlling for the random effects/hierarchical structure of the data (unlike the zero-order correlations, which are likely to over-estimate relationships), but yet without controlling for all the overlapping covariation among the fixed effects predictors. For instructed words, the results showed that item-level shift type, student-level comprehension and pretest morphological production, and classroom treatment condition were each positively predictive of a correct response. In fact, the only predictor that was not directly predictive of correct response was item-level SFI. For non-instructed words, the results were similar, except that item-level shift type was not significantly predictive of correct response for this set of items.

Last, a *unique* effects model (Model 3) was fit to the data, with all predictors entered simultaneously together to determine non-overlapping relationships between each predictor and the likelihood of a correct response. For instructed words, results showed the same pattern as found for our direct effects model (Model 2): item-level shift, student-level pretest comprehension and morphological production, and classroom-level treatment condition were each uniquely and positively predictive of the probability of correct response. When translated from logits to predicted probabilities, these estimates indicate that, holding all else constant, students with relatively higher levels of pretest comprehension (+1 *SD*) had a 90% predicted probability of a correct response compared to students with average comprehension (88%) and those with relatively lower levels of comprehension (85%). In a similar vein, students with relatively higher levels of pretest morphological skill (+1 *SD*) had a 94% predicted probability of a correct response compared to average (88%) and lower (-1 *SD*) pretest knowledge (77%). Last but not least, the predicted probability of a correct response for students in treatment classrooms was 94% compared to those in control classrooms (77%). Again, similar to Model 2, for non-instructed words, with all of the fixed effects entered simultaneously into the model, student-level comprehension and morphological production pretests as well as classroom-level treatment all remained significant. When translated from logits to predicted probabilities, these estimates indicate that, holding all else constant, students with relatively higher levels of pretest comprehension (+1 *SD*) had a 84% predicted probability of a correct response compared to students with average comprehension (79%) and those with relatively lower levels of comprehension (73%) (see Figure 1). In a similar vein, students with relatively higher levels of pretest morphological skill (+1 *SD*) had a 90% predicted probability of a correct response compared to average (79%) and lower (-1 *SD*) pretest knowledge (62%) (see Figure 2). Last but not least, the predicted probability of a correct response for students in treatment classrooms was 83% compared to those in control classrooms (74%) (see Figure 3).

Table 11. Morphological production: Instructed words

	Model 1 (Intercept Only)			Model 2 (Direct Main Effects)			Model 3 (Unique Main Effects)		
Fixed Effect	*Coeff*	*SE*	*t*	*Coeff*	*SE*	*t*	*Coeff*	*SE*	*t*
Intercept (Mean)	1.94	0.33	5.88 *	--	--	--	1.98	0.26	7.62 *
Item-Level									
SFI				0.24	0.34	0.71	0.08	0.32	0.25
Shift Type				0.54	0.25	2.16 *	0.53	0.26	2.04 *
Student-Level									
Comprehension				0.75	0.05	15.00 *	0.25	0.05	5.00 *
Pretest				0.97	0.04	24.25 *	0.79	0.05	15.80 *
Classroom-Level									
Treatment				0.75	0.09	8.33 *	0.78	0.06	13.00 *
Random Effect	*Var*			*Var*			*Var*		
Items	1.54			--	--	--	1.23		
Students/Classrooms	1.28			--	--	--	0.34		
Classrooms	0.66			--	--	--	0.04		
Fit Indices									
Deviance (-2LL)	8095.33			--	--	--	7675.95		
AIC	8103.33			--	--	--	7693.95		
BIC	8132.17			--	--	--	7758.83		

Note. *N* =499 students; 233 Treatment, 266 Control. 20 items.
$^{*}p < .05$

Table 12. Morphological production: Non-instructed words

	Model 1 (Intercept Only)			Model 2 (Direct Main Effects)			Model 3 (Unique Main Effects)		
Fixed Effect	*Coeff*	*SE*	*t*	*Coeff*	*SE*	*t*	*Coeff*	*SE*	*t*
Intercept (Mean)	1.32	0.38	3.47 *	--	--	--	1.34	0.33	4.06 *
Item-Level									
SFI				0.52	0.31	1.68	-0.43	0.30	-1.43
Shift Type				0.63	0.35	1.80	0.53	0.34	1.56
Student-Level									
Comprehension				0.83	0.05	16.60 *	0.34	0.05	6.80 *
Pretest				1.07	0.04	26.75 *	0.84	0.05	16.80 *
Classroom-Level									
Treatment				0.23	0.08	2.88 *	0.28	0.04	7.00 *
Random Effect	*Var*			*Var*			*Var*		
Items	2.79			--	--	--	2.19		
Students/Classrooms	1.44			--	--	--	0.28		
Classrooms	0.09			--	--	--	0.01		
Fit Indices									
Deviance (-2LL)	8720.56			--	--	--	8207.26		
AIC	8728.56			--	--	--	8225.26		
BIC	8757.39			--	--	--	8290.14		

Note. *N* =499 students; 233 Treatment, 266 Control. 20 items.
$^{*}p < .05$

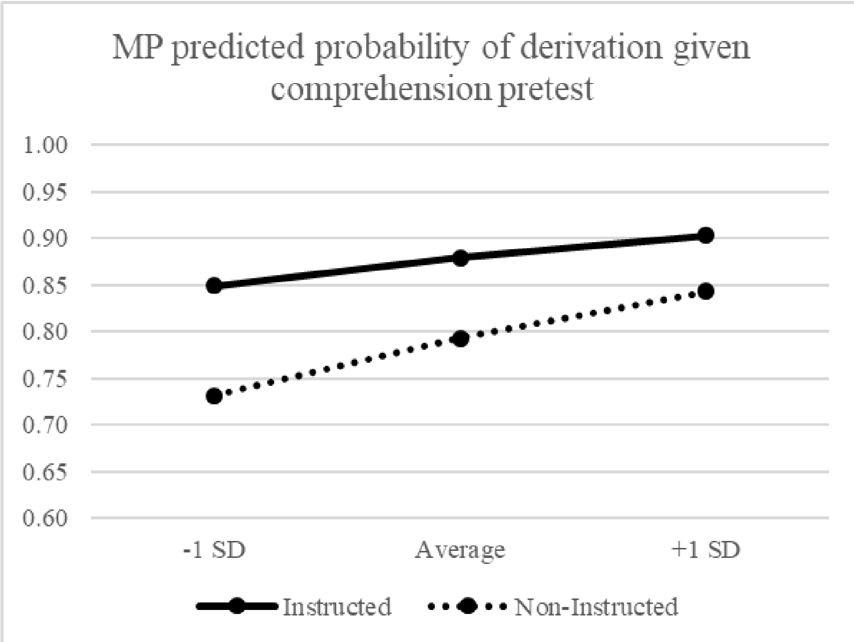

Figure 1. Probability of a student who scored at average on MP at pretest in an average classroom responding with a correct response on posttest MP for an average frequency word.

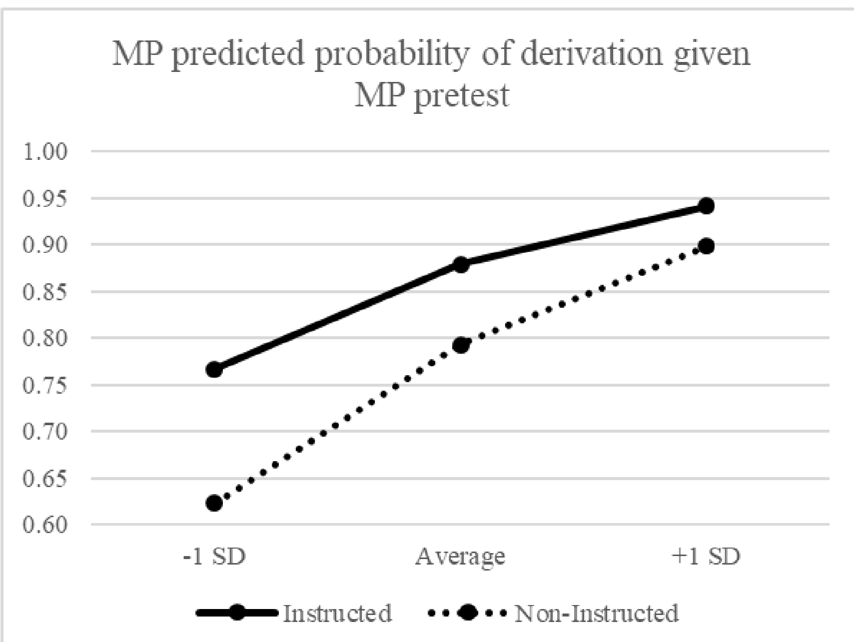

Figure 2. Probability of a student who scored at average on comprehension at pretest in an average classroom responding with a correct response on posttest MP for an average frequency word.

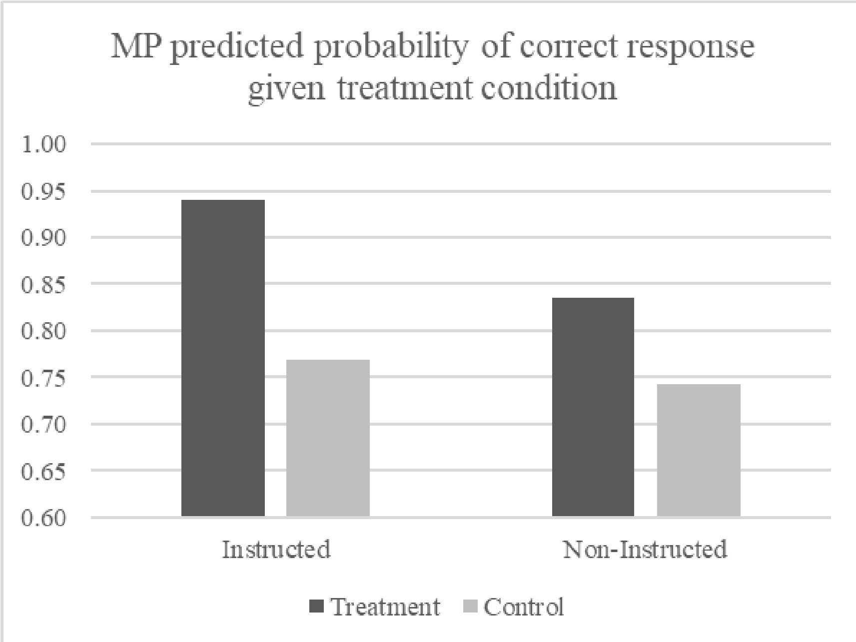

Figure 6. Probability of a student who scored at average on MP and comprehension at pretest responding with a correct response on posttest MP for an average frequency word, given treatment condition.

Sentence Combining Derivations. Model results for the sentence combining task are provided in Tables 13 and 14 for instructed and non-instructed words, respectively. Recall that Model 1 fitted the unconditional model—in other words, estimating the mean log-odds (logits) of the probability of correctly responding to the item, controlling for item, student within classroom, and classroom random effects. For instructed words, the intercept estimate was not significant (which translates to a mean predicted probability of an appropriate derivation of 50% across the entire sample). There was also substantial variability due to items ($s^2 = 0.76$), students within classrooms ($s^2 = 1.17$), and classrooms ($s^2 = 0.38$). For non-instructed words, the intercept estimate was not significant (which translates to a 50% probability of an appropriate derivation across the entire sample). Again, there was substantial variability due to items ($s^2 = 0.46$), students within classrooms ($s^2 = 2.46$), and classrooms ($s^2 = 0.30$).

Next, *direct* effects models (Model 2) were conducted to estimate the direct relationship between each of the fixed predictors and the likelihood of correctly deriving a word. For instructed words, the results showed that student-level comprehension and pretest sentence combining derivations, and classroom treatment condition were each positively predictive of an appropriate morphological derivation. Item-level word frequency was not uniquely predictive. For non-instructed words, the results were similar, except that item-level word frequency was uniquely and positively predictive of an appropriate derivation.

Lastly, a *unique* effects model (Model 3) was fit to the data, with all predictors entered simultaneously together to determine non-overlapping relationships between each predictor and the likelihood of a correct response. For instructed words, results showed the same pattern as found for our direct effects model (Model 2): student-level pretest comprehension and sentence combining derivations, and classroom-level treatment condition were each uniquely and positively predictive of the probability of correct response. When translated from logits to predicted probabilities, these estimates indicate that, holding all else constant, students with relatively higher levels of pretest comprehension (+1 *SD*) had a 67% predicted probability of a making an appropriate derivation compared to students with average comprehension (56%) and those with relatively lower levels of comprehension (46%) (see Figure 4). In a similar vein, students with relatively higher levels of pretest morphological skill (specifically, making morphological derivations in service of combining sentences at pretest) (+1 *SD*) had a 71% predicted probability of a correct response compared to average (56%) and lower (-1 *SD*) pretest knowledge (40%) (see Figure 5). Last but not least, the predicted probability of an appropriate derivation for students in treatment classrooms was 69% compared to those in control classrooms (43%) (see Figure 6). Again, similar to Model 2, for non-instructed words, with all of the fixed effects entered simultaneously into the model, student-level comprehension and sentence combining derivations pretests as well as classroom-level treatment all remained significant. Additionally, item-level word frequency was uniquely and positively predictive of the outcome. When translated from logits to predicted probabilities, these estimates indicate that, holding all else constant, students with relatively higher levels of pretest comprehension (+1 *SD*) had a 68% predicted probability of an appropriate derivation compared to students with average comprehension (55%) and those with relatively lower levels of comprehension (42%). In a similar vein, students with relatively higher levels of pretest morphological skill (+1 *SD*) had a 76% predicted probability of an appropriate derivation compared to average (55%) and lower (-1 *SD*) pretest knowledge (44%). Last but not least, the predicted probability of an appropriate derivation for students in treatment classrooms was 67% compared to those in control classrooms (44%).

Table 13. Sentence combining derivations: Instructed words

	Model 1 (Intercept Only)			Model 2 (Direct Main Effects)				Model 3 (Unique Main Effects)			
Fixed Effect	*Coeff*	*SE*	*t*	*Coeff*	*SE*	*t*		*Coeff*	*SE*	*t*	
Intercept (Mean)	0.36	0.34	1.06	--	--	--		0.26	0.39	0.67	
Item-Level											
SFI				0.23	0.19	1.21		0.23	0.18	1.28	
Student-Level											
Comprehension				0.76	0.06	12.67	*	0.43	0.06	7.17	*
Pretest				0.84	0.06	14.00	*	0.65	0.06	10.83	*
Classroom-Level											
Treatment				0.44	0.11	4.00	*	0.53	0.07	7.57	*
Random Effect	*Var*			*Var*				*Var*			
Items	0.76			--	--	--		0.62			
Students/Classrooms	1.17			--	--	--		0.37			
Classrooms	0.38			--	--	--		0.07			
Fit Indices											
Deviance (-2LL)	4702.64			--	--	--		4406.97			
AIC	4710.64			--	--	--		4422.97			
BIC	4735.81			--	--	--		4473.30			

Note. *N*=499 students; 233 Treatment, 266 Control. 20 items.

[*] $p < .05$.

Table 14. Sentence combining derivations: Non-Instructed Words

	Model 1 (Intercept Only)			Model 2 (Direct Main Effects)				Model 3 (Unique Main Effects)			
Fixed Effect	*Coeff*	*SE*	*t*	*Coeff*	*SE*	*t*		*Coeff*	*SE*	*t*	
Intercept (Mean)	0.10	0.28	0.36	--	--	--		0.22	0.19	1.16	
Item-Level											
SFI				0.55	0.21	2.62	*	0.56	0.20	2.80	*
Student-Level											
Comprehension				1.02	0.08	12.75	*	0.53	0.07	7.57	*
Pretest				1.17	0.07	16.71	*	0.95	0.08	11.88	*
Classroom-Level											
Treatment				0.34	0.12	2.83	*	0.47	0.06	7.83	*
Random Effect	*Var*			*Var*				*Var*			
Items	0.46			--	--	--		0.23			
Students/Classrooms	2.46			--	--	--		0.96			
Classrooms	0.30			--	--	--		0.00			
Fit Indices											
Deviance (-2LL)	4615.31			--	--	--		4284.77			
AIC	4623.31			--	--	--		4300.79			
BIC	4648.47			--	--	--		4351.12			

Note. N=499 students; 233 Treatment, 266 Control. 20 items.
* $p < .05$.

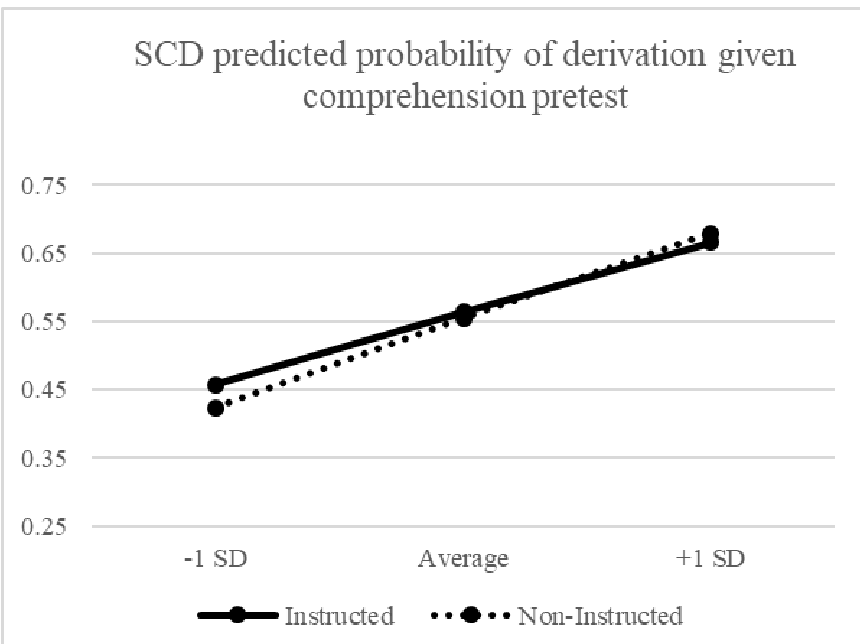

Figure 4. Probability of a student who scored at average on SCD at pretest in an average classroom responding with an appropriate derivational change on posttest SCD for an average frequency word.

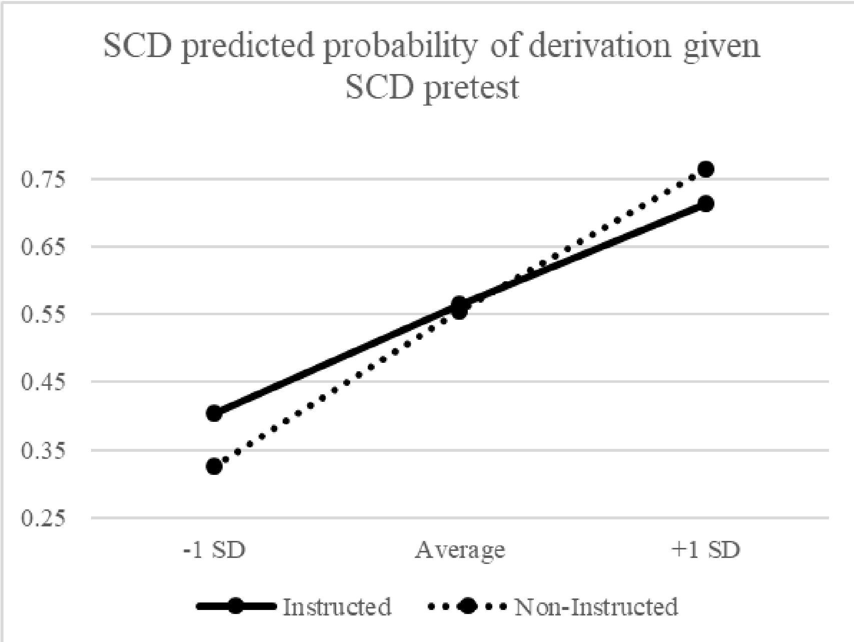

Figure 5. Probability of a student who scored at average on comprehension at pretest in an average classroom responding with an appropriate derivational change on posttest SCD for an average frequency word.

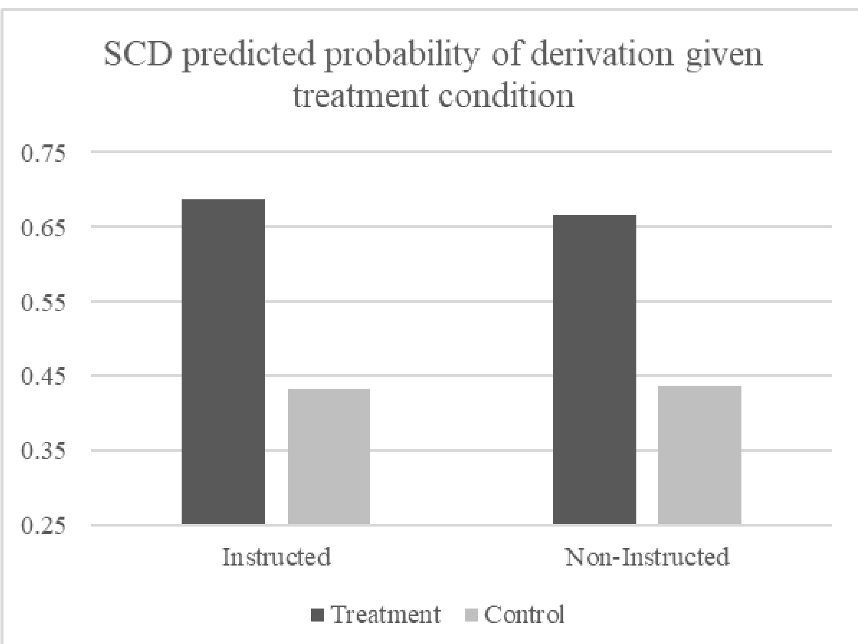

Figure 6. Probability of a student who scored at average on SCD and comprehension at pretest responding with an appropriate derivational change on posttest SCD for an average frequency word, given treatment condition.

DISCUSSION

Findings

This study provides strong evidence that embedded vocabulary instruction with a morphological focus can improve students' productive control of morphologically complex words in constrained sentence-writing production tasks. In general, results from the morphological production tasks provided evidence to show that students *can* manipulate base words to fit syntactically within an existing sentence. Similarly, results from the sentence combining derivations task provided evidence to show that students are *willing* manipulate base words to revise existing sentences. Whether they can—and whether they are willing to—morphologically manipulate base words was considered for words that were explicitly instructed and practiced within the intervention curriculum classrooms, as well as for words that were not instructed or practiced, to determine whether the morphological principles learned in through instruction would transfer to non-instructed words.

Treatment effects on instructed vs. non-instructed words. Overall, the model results show that effect of the morphological instruction treatment on the probability of a correct derivation for morphological production were larger for instructed words than for non-instructed words, which is to be expected (see again Figures 1-3). For the sentence combining task, students appeared to be equally likely to change instructed and non-instructed words (see again Figures 4-6). In general, students in treatment classrooms gained morpho-syntactic control over instructed words, and transferred knowledge of morphological structure to non-instructed words at a higher rate than students in control classrooms. The difference in treatment effects between instructed and non-instructed words was also larger for the morphological production task (treatment-control difference in predicted probabilities was 17% and 9% for each set of words, respectively), compared to the difference between treatment and control classrooms on instructed and non-instructed words for the sentence combining derivations task (treatment-control differences were 25% and 23% for the two sets of words, respectively).

Word-level SFI. The standard frequency index (SFI) of a word was not a unique predictor of whether students would correctly derive a complex instructed or non-instructed word to complete a sentence (morphological production) or of whether students would derive an instructed word in service of combining sentences (sentence combining derivations for instructed words). However, SFI was a unique predictor of whether students would derive a non-instructed word in service of combining sentences (sentence combining derivations for non-instructed words). Though these models reflect the response patterns of all students, rather than just students in treatment classrooms, these results suggest a few different things about the impact of instruction on students' productive control of complex morphologically forms.

First, consider morphological production. Recall again that this task provides a measure of whether students *can* change the words correctly, when prompted by a base and a stem sentence. SFI did not play a role in determining whether students (in either treatment or control classrooms) would derive instructed words. This result may likely be carried by students in treatment classrooms: because the words were taught and practiced in treatment classrooms, the frequency of words mattered less. However, SFI also did not uniquely predict a correct response for non-instructed words. It may be that students in both treatment and control classrooms were able to apply morphological principles to words that were not specifically instructed (with treatment classroom students responding correctly at a higher rate) regardless of their relative frequency, showing evidence of relational, syntactic, and distributional morphological knowledge. Additionally, though base word SFI was accounted for in these models, there was still substantial variation due to item-level random effects, suggesting that other word-specific factors not considered in these models might be at play more so than the relative SFI of the provided base words.

Sentence combining derivations, on the other hand, allows us to measure whether students *will* change a word in service of condensing syntax. Results suggest that students in both treatment and control classrooms did take up the strategy of deriving morphological relatives in order to condense syntax. However, students in treatment classrooms were more likely to derive words to combine sentences than students in control classrooms, both for instructed and non-instructed words. SFI did not play a role in determining whether students (in either treatment or control classrooms) would derive instructed words. Similarly to Morphological Production, this may be carried by students in treatment classrooms: because the words were taught and practiced in treatment classrooms, the frequency of words mattered less. Interestingly, base word frequency did uniquely predict derivations on non-instructed words, where words with higher frequency (relative to the frequency of other words in the task) were more likely to be changed by students in both treatment and control classrooms. These findings are consistent with Dobbs and Kearns' (2016) results, which showed that students were more likely to attempt to use higher frequency words in a loosely constrained task (that is, a writing task where students have freedom to construct text and use specific words as they see fit). In a task where students have more options for responding strategies, students may be less willing to attempt to change a word if it is less familiar to them.

Word-level shift type. Recall that the item-level predictor of shift type was only applicable to Morphological Production. The shift type between base and target words on the morphological production task was a significant direct and unique predictor of instructed words, such that students were more likely to respond correctly for word pairs that are phonologically and orthographically transparent rather than opaque (transparent shift: *adapt* → *adaptable*; opaque shift: *circle* → *circular*). Interestingly, this pattern did not translate to non-instructed words. In general, word pairs that are phonologically and orthographically transparent are easier for students to spell, read, and generate (Goodwin, 2016), so we might expect students to respond correctly at a higher rate for transparent word pairs. These trends were present among these data from the item means difficulty analysis (see Table 9): for all students and across all words, word pairs with a transparent shift (where adding an affix does not change the pronunciation or spelling of the base word) were less difficult to correctly derive than opaque words. Disaggregated item means suggest that students in treatment classrooms responded correctly at a similar rate across opaque and transparent instructed words, providing evidence that specific-word instruction may overcome the difficulties of opaque words. However, when all students were considered in the models, results indicated that shift type was still a unique predictor of correct response, where students were more likely to respond correctly if the shift is transparent. It is less clear why shift type was not a significant predictor of correct response for non-instructed words, though means suggest transparent non-instructed words were less difficult than opaque non-instructed words for all students,

Student pretest knowledge. As expected, students' initial skill levels—specifically, reading comprehension and morphological skill—were each positively and uniquely predictive of the probability of a correct response for both tasks and each word set within tasks. To better understand the effect of pretest morphological skill on Morphological Production, consider Figure 2. Holding all other predictors constant (including treatment condition), predicted probabilities suggest that students who scored one standard deviation below the mean on morphological skill pretest were considerably less likely to respond correctly on non-instructed words (instructed-non-instructed difference in predicted probabilities was 15%), while the probability of correct response between instructed and non-instructed words for students who scored one standard deviation above the mean on morphological skill at pretest was more similar (instructed-non-instructed difference in predicted probabilities was 4%). Though these predicted probabilities apply to students in the average classroom (neither treatment nor control), the difference cannot be attributed to characteristics of instructed vs. non-instructed words alone, because each type was equally difficult at pretest. Instead, this difference may be carried by the high correct response rate on instructed words by the students in treatment classrooms, for whom specific-word instruction strongly improved posttest correct responses. It appears that though students with lower pretest morphological knowledge improved on instructed words, they were less likely to transfer their knowledge to non-instructed words. Comprehension pretest score followed a similar pattern between predicted probabilities of instructed and non-instructed words (instructed-non-instructed difference in predicted probabilities was 12% for students scoring -1 SD and 6% for students scoring +1 SD on comprehension at pretest).

For Sentence Combining Derivations, consider Figures 4 and 5. Predicted probabilities of SCD responses followed similar patterns whether considering comprehension pretest or morphological skill (SCD) pretest. Specifically, holding all other variables constant, students performing one standard deviation below the mean at pretest had a lower predicted probability of deriving both instructed and non-instructed words at posttest than students performing one standard deviation above the mean at pretest. In contrast to Morphological Production, students were more-or-less equally likely to derive either instructed or non-instructed words on Sentence Combining Derivations, regardless of their pretest performance. This may be due to the moderately-constrained nature of the task in general, or to the nature of the words and items themselves. For example, it appears that at posttest, students in both treatment and control classrooms were more likely to derive non-instructed words than instructed words (see Table 10). Because students were not compelled to change words in order to be successful on this task, we could expect to see more variation in response strategies given both the content of the kernel sentences and the characteristics of the words involved. Some sentences, depending on their content and structure, may have lent themselves better to morphological changes than others. When they had more freedom to manipulate syntax as they saw fit, students within pretest skill levels changed instructed and non-instructed words at an equal rate, regardless of whether the word was practiced. Additionally, although word frequency was balanced across instruction type, it is possible too that the non-instructed words were more easily manipulated or familiar to the students, a characteristic not able to be captured by word-level frequency alone.

These results show that pretest knowledge does influence the outcomes we might expect from this type of word-level instruction. Students with higher pretest knowledge were more likely to respond correctly and derive words across both outcome measures at a more-or-less equal rate among instructed and non-instructed words, regardless of treatment condition. Though these models cannot specifically provide evidence of differential gains over time, it does provide evidence that students scoring low at pretest were more likely to correctly derive instructed words than non-instructed words for the highly-constrained Morphological Production task, suggesting that direct instruction on specific words can help low-performing students make gains in control over complex words in sentences for tasks with scaffolding.

Present Study Findings in Context

Though not discussed in the present study, the main intervention study associated with the present study (McCutchen, Herrera, Northey, Clark, & Huey, 2017) found a similar pattern of results for the previously discussed sentence combining *quality* score, which was used as a measure of overall sentence quality without accounting for morphological derivations. Specifically, students in treatment classrooms scored significantly higher on the sentence combining quality metric than students in control classrooms for sentences containing instructed and non-instructed words, suggesting that the instruction not only increased the likelihood of students making word-level manipulations in service of sentence-level syntax, but it also improved their control over the revision of sentence-level syntax to write high-quality sentences.

Our investigation of the broader effects of our intervention on a range of reading and writing measures reveal results consistent with previous findings in the literature (McCutchen et al., 2017): our results revealed smaller or non-significant effects on measures not specifically related to morphological skill (typical spelling, vocabulary, and comprehension tasks) and moderate to large effects on measures directly requiring morphological skill (see meta-analyses by Goodwin and Ahn (2013) for morphology and Elleman, Lindo, Morphy, and Compton (2009) for vocabulary). However, McCutchen et al. (2017) also added uniquely to our understanding of vocabulary and morphology interventions by testing the effects of intervention on writing tasks: specifically, on morphological production, sentence combining, and essay writing. Word-level instructional interventions overwhelmingly focus on reading outcomes, focusing on growth in definitional vocabulary, comprehension, word reading, and various measures of receptive morphological awareness. Few studies (if any) have considered intervention effects on students' ability to use specific words productively, even though students are expected to use the words they use in their written and oral responses across many school activities.

Consider the following set of sentences from the sentence combining task: *The person searched the house. The house has history. The person investigates crimes.* Here is the response from a student in a treatment classroom at pretest: *The person investigates crimes like a history house*.; and the same student's response at posttest: *The crime investigater* [sic] *searched the historic house*. Though the word families for *history* and *investigate* were not specifically taught in the intervention, students were able to use their existing knowledge of the word families and their morphological structure to not only create more condensed, grammatical sentences, but also to form responses that perhaps more closely resembled the meaning of the original sentences. A different student in a treatment classroom provides an example of growth between pretest and posttest in making syntax even denser and more clear. In responding to the following item: *The boy described his plan. He often imagines. The way it was described was vivid*., one student wrote at pretest: *The boy with an imagination described his plan in a vivid way*.; and at posttest: *The amaginitive* [sic] *boy decribed* [sic] *his plan vividly*. Though students in treatment and control classrooms improved on both the number of derivations made and the quality of the sentences written over the course of the intervention, the evidence from this study suggests that students receiving direct instruction on word-level structural attributes and their syntactic functions showed larger gains in specific word-level production and control over complex syntax.

Instructional Implications

These findings provide some initial insights into best approaches for word-level instruction. In considering word characteristics that may influence learning, these data suggest that direct instruction on specific words and their families can directly improve students' morpho-syntactic control on instructed words. Transfer of morphological knowledge for fourth and fifth grade students may be more likely to be applied to morphological families whose phonological and orthographic relationships are more transparent, and whose words are more familiar—especially for students with lower general literacy skills. It appears as though direct instruction on specific words can help students overcome the difficulties of opaque and lower-frequency words, suggesting that careful selection of words for instructional focus is important. Additionally, these data suggest that direct instruction on the structural features of words in general can have an effect on children's flexible manipulation of word- and sentence-level syntax to produce more complex and precise sentences. Thus, word-level instruction can be usefully applied to sentence-level writing instruction.

Limitations

At the onset of designing this study, I had planned to examine additional word-level characteristics, as well as interactions between word-, student-, and classroom-level effects. I had also planned to include both instructed and non-instructed words within the same model so that I could directly compare performance on each type. However, when these factors were included in the models, the models were unable to converge, suggesting collinearity between factors. This may be due to a number of the characteristics that made these models complex, including the binary outcomes, the small number of items on each measure (40 on Morphological Production and 16 on Sentence Combining), and the combination of cross-classification and partial nesting of items and students within classrooms. This led me to pursue other methods to examine the relevant relationships. Nevertheless, the findings here show that instruction had a unique and positive effect on students' morpho-syntactic production in tasks with more or less constraint and for instructed and non-instructed words.

Future Research

This study provides evidence that instruction can improve students' appropriate morpho-syntactic production in constrained writing tasks for a range of instructed and non-instructed words. Still, research is needed to better understand students' word-level production more broadly. Understanding students' lexical and morpho-syntactic production is valuable for understanding practical outcomes—the skill of using words in spoken and written language production—but it is also an untapped area that could provide insight into literacy skill in general. If we can better understand the conditions which are necessary for a student to choose to use a certain word or word form, or for a student to use a word correctly, we might better understand the nature of word knowledge, for example. Studies that examine the convergent and divergent validity of morphological production measures could help us to better understand the possible differences between productive and receptive morphological skill. The development of measures that can accurately assess word-level production in ways that are perhaps less contrived would aid in our understanding of this skill, as well as to provide a more situated and purposeful writing activity for students to engage in. Other areas of word-level production should also be considered. Beyond morpho-syntactic production, for example, students' use of words might be assessed for quality of usage (including semantic and collocational appropriateness).

Word-level instructional interventions should also expand their scope to look beyond reading outcomes to include different levels of writing skill. With the recent increased emphasis on writing standards through the Common Core, all levels of the writing process need to be better understood (production at the word-, sentence-, and text-level of writing), so that effective instructional approaches can be developed. Research in learning and teaching writing has largely focused on higher order writing processes, such as planning and organizing (Connors, 2000), though there is a still a need for more word- and sentence-specific writing instruction. Sentence combining instruction has been found to improve the writing quality of both typically achieving students and for students with writing difficulties (Saddler & Graham, 2005; Saddler, Behforooz, & Asaro, 2008), providing an opportunity for word- and sentence-specific instruction. Promising results from an instructional intervention embedding meaningful grammar instruction into more general writing curriculum suggested that focus on word- and sentence-level structural features can improve essay writing quality—especially for classrooms whose teachers have heightened metalinguistic awareness (Jones, Myhill, & Bailey, 2013; Myhill, Jones, & Watson, 2013). Word-level instructional interventions like those discussed in this paper have great potential for supporting students' productive use of words in their sentence and essay writing. And yet, few have investigated those types of outcomes, even though one of the goals of specific-word instruction is to encourage word ownership through spoken and written use.

Conclusions

Word-level instruction—especially instruction that supports the flexible use of word forms to maneuver more complex syntax—has the potential to impact student sentence-level production in a direct way. This study presents evidence that embedded, morphologically focused vocabulary instruction can help students better control the type of complex morphological forms and syntactic structures in which they are used, for words ranging in frequency.

References

Allen, L. K., Snow, E. L., & McNamara, D. S. (2016). The narrative waltz: The role of flexibility in writing proficiency. *Journal of Educational Psychology, 108*(7), 911-924.

Alves, R. A., Branco, M., Castro, S. L. & Olive, T. (2012). Effects of handwriting skill, output modes, and gender on fourth graders' pauses, language bursts, fluency, and quality. In V. W. Berninger (Ed.), *Past, present and future contributions of cognitive writing research to cognitive psychology*, (pp. 389-402). New York, NY: Psychology Press.

Apel, K., Brimo, D., Diehm, E., & Apel, L. (2013). Morphological awareness intervention with kindergartners and first- and second-grade students from low socioeconomic status homes: A feasibility study. *Language, Speech, and Hearing Services in Schools, 44*, 161-173.

Apel, K., Wilson-Fowler, E. B., Brimo, D., & Perrin, N. A. (2012). Metalinguistic contributions to reading and spelling in second and third grade students. *Reading and Writing, 25*, 1283-1305.

Apel, K. (2014). A comprehensive definition of morphological awareness. *Topical Language Disorders*, 34(3), 197-209.

Baayen, R. H., Davidson, D.J., & Bates, D.M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390-412.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., et al. (2007). The English Lexicon Project. *Behavior Research Methods, 39*, 445–459.

Baumann, J. F., Carr Edwards, E., Boland, E., 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*(2), 447-494.

Baumann, J. F., Carr Edwards, E., 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., & Ash, G. E. (2003). Research on vocabulary instruction: Voltaire redux. In J. Flood, D. Lapp, J. R. Squire, & J. M. Jensen (Eds.), *Handbook of Research on Teaching the English Language Arts*, 2nd Ed. (pp. 752-785). Mahwah, N.J.: Lawrence Erlbaum Associates.

Beck, I. L., & McKeown, M. G. (1985). Teaching vocabulary: Making the instruction fit the goal. *Educational Perspectives, 23*, 11-15.

Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction*. New York, NY: Guilford.

Berman, R. (2009). Developing linguistic knowledge and language use across adolescence. In E. Hoff & M. Shatz (Eds.), *Blackwell Handbook of Language Development* (pp. 347-367). Oxford: Blackwell Press.

Berman, R. A., & Nir-Sagiv, B. (2007). Comparing narrative and expository text construction across adolescence: A developmental paradox. *Discourse Processes, 43*(2), 79-120.

Berninger, V. W., Abbott, R. D., Nagy, W., & Carlisle, J. (2010). Growth in phonological, orthographic, and morphological awareness in grades 1 to 6. *Journal of Psycholinguistic Research, 39*, 141-163.

Berninger, V. W., & Amtmann, D. (2003). Preventing written expression disabilities through early and continuing assessment and intervention for handwriting and/or spelling problems: Research into practice. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.) *Handbook of learning disabilities* (pp. 345-363). New York: Guilford Press.

Berninger, V. W., Fuller, Fl. & Whittaker, D. (1996). A process model of writing development across the lifespan. *Education Psychology Review, 8*(3), 193-217.

Berninger, V. W., Nagy, W., & Beers, S. (2011). Child writers' construction and reconstruction of single sentences and construction of multi-sentence texts: Contributions of syntax and transcription to translation. *Reading and Writing, 24*, 151-182.

Berninger, V. W., Yates, C., Cartwright, A., Rutberg, J., Remy, E., & Abbott, R. D. (1992). Lower-level developmental skills in beginning writing. *Reading and Writing, 4*, 257-280.

Bowers, P. N., & Kirby, J. R. (2010). Effects of morphological instruction on vocabulary acquisition. *Reading and Writing: An Interdisciplinary Journal, 23,* 515-537.

Carlisle, J. (1995). Morphological awareness and early reading achievement. In L. Feldman (Ed.) *Morphological aspects of language processing* (pp. 189-210). Hillsdale, N.J.: Lawrence Erlbaum Associates.

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

Carlisle, J. F. (2010). Effects of instruction in morphological awareness on literacy achievement: An integrative review. *Reading Research Quarterly*, *45*(4), 464-487.

Carlisle, J. F. & Fleming, J. (2003). Lexical processing of morphologically complex words in the elementary years. *Scientific Studies of Reading*, *7*(3), 239-253.

Carlisle, J. F., & Katz, L. A. (2006). Effects of word and morpheme familiarity on reading of derived words. *Reading and Writing: An Interdisciplinary Journal, 19*, 669-693.

Carlisle, J. F., Kelcey, B., & Berebitsky, D. (2013). Teachers' support of students' vocabulary learning during literacy instruction in high poverty elementary schools. *American Educational Research Journal, 50*(6), 1360-1391.

Carlisle, J. R., & Stone, C. A. (2005). Exploring the role of morphemes in word reading. *Reading Research Quarterly, 40*(4), 428-449.

Chenoweth, N. A. & Hayes, J. R. (2001). Fluency in writing: Generating text in L1 and L2. *Written Communication*, *18*(1), 80-98.

Cho, S. J., & Goodwin, A. P. (2016). Modeling learning in doubly multilevel binary longitudinal data using generalized linear mixed models: An application to measuring and explaining word learning. *Psychometrika*. Advance online publication.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12*, 335-359.

Clark, T. L., & McCutchen, D. (2014, April). The contribution of morphological awareness to reading beyond decoding, vocabulary, and processing artifacts. Paper presented at the meeting of the American Educational Research Association. Philadelphia, PA.

Connors, R. J. (2000). The erasure of the sentence. *College Composition and Communication,* 52(1), 96-126.

Coxhead, A. (1998). *An academic word list*. Wellington, NZ: Victoria University of Wellington.

Crosson, A. C., & McKeown, M. G. (2016). Middle school learners' use of Latin roots to infer the meaning of unfamiliar words. *Cognition and Instruction*, *34*(2), 148-171.

Cunningham, A., & Stanovich, K. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33*(6), 934-945.

Dale, E. (1965). Vocabulary measurement: Techniques and major findings. *Elementary English, 42*, 82-88.

Deacon, S. H., & Dhooge, S. (2010). Developmental stability in changes in the impact of root consistency on children's spelling. *Reading and Writing: An Interdisciplinary Journal, 23*, 1055-1069.

Deacon, S. H., Francis, K., & Tong, X. (2017). The relationship of morphological analysis and morphological decoding to reading comprehension. *Journal of Research in Reading, 40*(1), 1-16.

Deacon, S. H., Kieffer, M. J., & Laroche, A. (2014). The relation between morphological awareness and reading comprehension: Evidence from mediation and longitudinal models. *Scientific Studies of Reading, 18*, 432-451.

Deacon, S. H., Kirby, J. R., & Casselman-Bell, M. (2009). How robust is the contribution of morphological awareness to general spelling outcomes? *Reading Psychology, 30*(4), 301-318.

Devonshire, V., & Fluck, M. (2010). Spelling development: Fine-tuning strategy-use and capitalising on the connections between words. *Learning and Instruction, 20*, 361-371.

Dobbs, C. L., & Kearns, D. (2016). Using new vocabulary in writing: Exploring how word and learner characteristics relate to the likelihood that writers use newly taught vocabulary. *Reading and Writing: An Interdisciplinary Journal, 29*, 1817-1843.

Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. Journal of Research on Educational Effectiveness, 2, 1–44.

Flinspach, S. L., Scott, J. A., & Vevea, J. L. (2010). Rare words in students' writing as a measure of vocabulary. In R. T. Jiménez, V. J. Risko, M. K. Hundley, & D. W. Rowe (Eds), *59th Yearbook of the National Reading Conference* (pp. 187-200). Oak Creek, WI: National Reading Conference, Inc.

Gilbert, J. K., Goodwin, A. P., Compton, D. L., & Kearns, D. M. (2014). Multisyllabic word reading as a moderator of morphological awareness and reading comprehension. *Journal of Learning Disabilities, 49*(1), 34-43.

Goodwin, A. P., & Ahn, S. (2013). A meta-analysis of morphological interventions in English: Effects on literacy outcomes for school-age children. *Scientific Studies of Reading, 17*:4, 257-285.

Goodwin, A. P. & Cho, S. J. (2016). Unraveling vocabulary leaning: Reading and item-level predictors of vocabulary learning within comprehension instruction for fifth and sixth graders. *Scientific Studies of Reading, 20*(6), 490-514.

Goodwin, A. P., Gilbert, J. K., & Cho, S. J. (2013). Morphological contributions to adolescent word reading: An item response approach. *Reading Research Quarterly, 48*(1), 39-60.

Goodwin, A. P., Gilbert, J. K., Cho, S., & Kearns, D. M. (2014). Probing lexical representations: Simultaneous modeling of word and reader contributions to multidimensional lexical representations. *Journal of Educational Psychology, 106* (2), 448-468.

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99*(3), 445-476.

Green, L. B., McCutchen, D., Schwiebert, C., Quinlan, T., Eva-Wood, A., & Juelis, J. (2003). Morphological development in children’s writing. *Journal of Educational Psychology* , 95, 752-761.

Hart, B., & Risley, T. (2003). The early catastrophe: The 30 million word gap by age 3. *American Educator, 27*(1), 4-9.

Jarmulowicz, L. & Taran, V. L. (2013). Lexical morphology: Structure, process, and development . *Topics in Language Disorders, 33*(1), 57-72.

Jones, S., Myhill, D. & Bailey, T. (2012). Grammar for writing? An investigation of the effects of contextualised grammar teaching on students' writing. *Reading and Writing, 26*, 1241-1263.

Kamberelis, G. (1999). Genre development and learning: "Children writing stories, science reports, and poems". *Research in the Teaching of English, 33*(4), 403-460.

Kearns, D. M., Steacy, L. M., Compton, D. L., Gilbert, J. K., Goodwin, A., Cho, E., et al. (2016). Modeling polymorphemic word recognition: Exploring differences among children with early- emerging and late-emerging word reading difficulty. *Journal of Learning Disabilities, 49*(4), 368-394.

Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 57-71). Mahwah, NJ: Lawrence Erlbaum.

Kieffer, M. J., Biancarosa, G., & Mancilla-Martinez, J. (2013). Roles of morphological awareness in the reading comprehension of Spanish-speaking language minority learners: Exploring partial mediation by vocabulary and reading fluency. *Applied Psycholinguistics, 34*, 697-725.

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.

Kirby, J., Deacon, S., Bowers, P., Izenberg, L., Wade-Woolley, L., & Parrila, R. (2012). Children's morphological awareness and reading ability. *Reading and Writing, 25*, 389-410.

Kirk, C. & Gillon, G. T. (2009). Integrated morphological awareness intervention as a tool for improving literacy. *Language, Speech, and Hearing Sciences in Schools, 40*, 341-351.

Krashen, S. (2012). Academic jibberish. *RELC Journal*, 43(2), 283-285.

Kraut, R. (2015). The relationship between morphological awareness and morphological decomposition among English language learners. *Reading and Writing: An Interdisciplinary Journal, 28*, 873-890.

Kuo, L. J. & Anderson, R. C. (2006). Morphological awareness and learning to read: A cross-language perspective. *Educational Psychologist*, *31*(3). 161-180.

Leseaux, N. K., Kieffer, M. J., Faller, S. E., & Kelley, J. G. (2010). The effectiveness and ease of implementation of an academic vocabulary intervention for linguistically diverse students in urban middle schools. *Reading Research Quarterly, 45*(2), 196-228.

MacGinitie, W. H., MacGinitie, R. K., Maria, K., Dreyer, L. G., & Hughes, K. E. (2007). *Gates-MacGinitie Reading Tests* (4^{th} ed). Rolling Meadows, IL: The Riverside Publishing Company.

McCutchen, D. (1996). A capacity theory of writing: Working memory in composition. *Educational Psychology Review, 8*(3), 299-325.

McCutchen, D., Covill, A., Hoyne, S. H., & Mildes, K. (1994). Individual differences in writing skill: Implications of translating fluency. *Journal of Educational Psychology, 86*, 256–266.

McCutchen, D., Green, L., & Abbott, R. D. (2008). Children's morphological knowledge: Links to literacy. *Reading Psychology, 29*(4), 289-314.

McCutchen, D., Herrera, B. L., Northey, M., Clark, T., & Huey, E. (2017, April). Effects of embedded morphologically-focused vocabulary instruction on word knowledge and productive use in fourth and fifth grade. Paper presented at the meeting of the American Educational Research Association. San Antonio, TX.

McCutchen, D., Logan, B., & Biangardi-Orpe, U. (2009). Making meaning: Children's sensitivity to morphological information during word reading. *Reading Research Quarterly, 44*(4), 360-376.

McCutchen, D., & Logan, B. (2011). Inside incidental word learning: Children’s strategic use of morphological information to infer word meanings. *Reading Research Quarterly*, *46*(4), 334-349.

McCutchen, D., Stull, S., Herrera, B. L, Lotas, S., & Evans, S. (2014). Putting words to work: Effects of morphological instruction on children’s writing. *Journal of Learning Disabilities, 47*(1), 86-97.

McCutchen, D., & Stull, S. (2015). Morphological awareness and children's writing: Accuracy, error, and invention. *Reading and Writing: An Interdisciplinary Journal, 28*, 271-289.

McKeown, M. G., Beck, I. L., Omanson, R. C., & Pople, M. T. (1985). Some effects of the nature and frequency of vocabulary instruction on the knowledge and use of words. *Reading Research Quarterly, 20*(5), 522-535.

McKeown, M. G., Beck, I. L., & Sandora, C. (2012). Direct and rich vocabulary instruction needs to start early. In E. J. Kame'enui & J. F. Baumann (Eds.), *Vocabulary Instruction: Research to Practice* (2nd ed., pp. 17-33). New York: Guilford Press.

Meara, P. (1997). Towards a new approach to modelling vocabulary acquisition. In Schmitt, N., & McCarthy, M. (Eds.), *Vocabulary: Description, acquisition, and pedagogy* (pp. 109-121). Cambridge: Cambridge University Press.

Menyuk, P. (1988). *Language development: Knowledge and use*. Glenview, IL: Scott, Foresman.

Miller, T. F., Gage-Serio, O., & Scott, J. A. (2010). Word consciousness in practice: Illustrations from a fourth-grade teacher's classroom. In R. T. Jiménez, V. J. Risko, M. K. Hundley, & D. W. Rowe (Eds), *59th Yearbook of the National Reading Conference* (pp. 171-186). Oak Creek, WI: National Reading Conference, Inc.

Myhill, D. (2008). Towards a linguistic model of sentence development in writing. *Language and Education*, *22*, 271-288.

Myhill, D. (2012). The ordeal of deliberate choice: Metalinguistic development in secondary writers. In V. W. Berninger (Ed.), *Past, Present and Future Contributions of Cognitive Writing Research to Cognitive Psychology* (pp. 247-273). New York, NY: Psychology Press.

Myhill, D., Jones, S., & Watson, A. (2013). Grammar matters: How teachers’ grammatical knowledge impacts on the teaching of writing. *Teaching and Teacher Education, 36*, 77-91.

Nagy, W. E. & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly, 19*(3), 304-330.

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, 96*(1), 134-147.

Nagy, W. E., Carlisle, J. F., & Goodwin, A. P. (2014). Morphological knowledge and literacy acquisition. *Journal of Learning Disabilities, 47*(1), 3-12.

Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly, 47*(1), pp. 91-108.

National Governors Association and Council of Chief School Officers. (2010). *Common core state standards.* Retrieved from http://www.corestandards.org/.

Northey, M., McCutchen, D., & Sanders, E. (2016). Contributions of morphological skill to children's essay writing. *Reading and Writing: An Interdisciplinary Journal, 29*(1), 47-68.

Olinghouse, N. G., & Wilson, J. (2013). The relationship between vocabulary and writing quality in three genres. *Reading and Writing, 26*, 45-65.

Olive, T. & Kellogg, R. (2002). Concurrent activation of high- and low-level production processes in written composition. *Memory and Cognition, 20*(4), 594-600.

Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 174-200). Cambridge: Cambridge University Press.

Pearson (2010). *Wechsler Individual Achievement Test, Third Edition*. Bloomington, MN: PsychCorp.

Pearson, P. D., Hiebert, E. H., & Kamil, M. L. (2012). Vocabulary assessment: Making do with what we have while we create the tools we need. In E. J. Kame'enui & J. F. Baumann (Eds.), *Vocabulary Instruction: Research to Practice* (2nd ed., pp. 231-255). New York: Guilford Press.

Perfetti, C. A. & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189-213). Amsterdam, the Netherlands: John Benjamins.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357-383.

Plester, B. Wood, C. & Joshi, P. (2009). Exploring the relationship between children's knowledge of text message abbreviations and school literacy outcomes. *British Journal of Developmental Psychology, 27*, 145-161.

Quené, H., & van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language, 59*, 413-425.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd Ed.). Thousand Oaks, CA: Sage Publications.

Raudenbush, S. W., Bryk, A., & Congdon, R. (2010). *HLM7 Hierarchical Linear and Nonlinear Modeling*. Lincolnwood, IL: Scientific Software International, Inc.

Ravid, D., & Tolchinsky, L. (2002). Developing linguistic literacy: A comprehensive model. *Journal of Child Language, 29*, 419-448.

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.

Saddler, B., Behforooz, B. & Asaro, K. (2008). The effects of sentence-combining instruction on the writing of fourth-grade students with writing difficulties. *The Journal of Special Education, 42*(2), 79-90.

Saddler, B. & Graham, S. (2005). The effects of peer-assisted sentence-combining instruction on the writing performance of more and less skilled young writers. *Journal of Educational Psychology, 97*(1), 43-54.

Schleppegrell, M. (2012). Academic language in teaching and learning. *The Elementary School Journal*, *112*(3), pp. 409-418.
Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. New York, NY: Palgrave Macmillan.
Schreuder, R., & Baayen, R. H. (1997). Modeling morphological processing. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 131-154). Hillsdale, NJ: Erlbaum.
Scott, J. A., Jamieson-Noel, D., & Asselin, M. (2003). Vocabulary instruction throughout the day in twenty-three Canadian upper-elementary classrooms. *The Elementary School Journal, 103*(3), 269-286.
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*, 219-252.
Snow, C. E. & Uccelli, P. (2009). The challenge of academic language. In D. R. Olson & N. Torrance (Eds.), *The Cambridge handbook of literacy* (pp. 112-133). Cambridge: Cambridge University Press.
Spencer, M., Muse, A., Wagner, R. K., Foorman, B., Petscher, Y., Schatschneider, C., Tighe, E. L., & Bishop, M. D. (2015). Examining the underlying dimensions of morphological awareness and vocabulary knowledge. *Reading and Writing: An Interdisciplinary Journal, 28*, 959-988.
Stahl, S. A. (2003). How words are learned incrementally. *American Educator*, *27*(1), 18-19.
Stahl, S. A., & Fairbanks, M. M. (2006). The effects of vocabulary instruction: A model-based meta-analysis. In K. A. Dougherty Stahl & M. C. McKenna (Eds.), *Reading Research at Work* (pp. 226-261). New York, NY: Guilford Press.
Stolz, J. A., & Feldman, L. B. (1995). The role of orthographic and semantic transparency of the base morpheme in morphological processing. In L. Feldman (Ed) *Morphological aspects of language processing* (pp. 189-210). Hillsdale, N.J.: Lawrence Erlbaum Associates.
Vadasy, P. F., Sanders, E. A., & Herrera, B. L. (2014). Efficacy of rich vocabulary instruction in fourth- and fifth-grade classrooms. *Journal of Research on Educational Effectiveness, 8*(3), 325-365.
Woodcock, R., McGrew, K., & Mather, N. (2001). *Woodcock-Johnson III tests of achievement*. Itasca, IL: Riverside.
Wright, T. S., & Peltier, M. R., (2016). Supports for vocabulary instruction in early language and literacy methods textbooks. *Reading & Writing Quarterly, 32*, 527-549.
Zeno, S. M., Ivens, D. H., Millard, R. T., & Duuvuri, R. (1995). *The educator's word frequency guide.* New York, NY: Touchstone Applied Science Associates.

Appendix

Morphological Production Task

1 explore: He wanted to be an _________________________.

2 accept: The food tasted _________________________.

3 memory: The party was _________________________.

4 mystery: The house seemed _________________________.

5 clever: The girl is known for her _________________________.

6 agree: The two groups were in _________________________.

7 translate: He could only read a _________________________.

8 miracle: Their survival was _________________________.

9 experiment: The medicine was _________________________.

10 construct: The meeting was very _________________________.

11 imitate: The painting was an _________________________.

12 circle: He wanted the rumor to _________________________.

13 function: The room was very _________________________.

14 sincere: He spoke very _________________________.

15 advantage: The change was _________________________.

16 explode: They heard an _________________________.

17 direct: The company hired a _________________________.

18 visual: Some things are hard to _________________________.

19 describe: Her instructions were very _________________________.

20 clear: She asked the teacher to _________________________.

21 emotion: The news made her _________________________.

22 invent: The woman was a famous _________________________.

23 suspect: The open door seemed _________________________.

24 create: She presented a new _________________________.

25 adapt: Animals that survive must be _________________________.

26 achieve: This was her greatest _________________________.

27 sympathy: The teacher was _________________________.

28 observe: He wrote down his _________________________.

29 manage: The amount of work was quite _________________________.

30 reflect: She was feeling very _________________________.

31 operate: They prepared for the _________________________.

32 exclude: The brand of clothing was very _________________________.

33 vivid: She remembered her dream _________________________.

34 assign: She finished the _________________________.

35 intend: They had clear _________________________.

36 produce: The day had been _________________________.

37 secret: The group met _________________________.

38 depend: The car is _________________________.

39 finance: Her reasons for moving were all _________________________.

40 severe: They didn't know the problem's _________________________.

Note. Students were given 15 minutes to complete this task. Words that finished the sentence in a syntactically and meaningfully sound way received a point. Non-standard forms (inventions, regularized forms, etc.) did not receive a point for this score. Phonological misspellings (*produckshun*) were counted as correct.

Sentence Combining Task

1	The boy describes his plan. He often imagines. The way it was described was vivid.
2	The bird was saved by a girl. The girl had lots of courage. The bird felt terror.
3	The child climbed the forbidden tree. The child often goes on adventures. He climbed the tree to defy.
4	The chef prepared the soup. She likes to create. The soup had lots of nutrients.
5	The meeting was for a project. The project is an experiment. The meeting was to collaborate.
6	The person searched the house. The house has history. The person investigates crimes.
7	The doctor solved the problem. The way it was solved was clever. The doctor had good sense.
8	The parents smiled at their son. The parents were full of pride. Many are impressed by their son.

Note. Students were given 16 minutes to complete this task. Derived words that fit syntactically and meaningfully sound way received a point. Non-standard forms (inventions, regularized forms, etc.) did not receive a point for this score. Phonological misspellings (produckshun) were counted as correct.