ABSTRACT

Title of Dissertation: EFFECTS OF WORD AND STUDENT FACTORS ON INSTRUCTIONAL REINFORCEMENT IN A VOCABULARY PROGRAM IMPLEMENTED WITH SECOND GRADE BILINGUALS

Lauren Artzi, Doctor of Philosophy, 2018

Dissertation directed by: Megan Madigan Peercy, Advisor and Chair, Education Department

This study investigates word learning effects from a twenty-minute read aloud vocabulary program developed for second grade English language learners (ELLs). In one condition, words were not instructed (exposed only condition), in another condition words were reinforced six times (reinforcement condition), in a third condition words were reinforced fourteen times (reinforcement plus condition). Employing a within-subjects design where each study participant received each level of treatment, this study examines generalized English vocabulary knowledge, Spanish language proficiency, and type of word (abstract words, concrete words, cognates, and noncognates) as possible influencers on the word learning in each condition. Two hundred and twenty-eight students across twelve classrooms participated in this study.

Findings reveal that the vocabulary instruction in the two instructed conditions compared to the exposed only condition garnered significant strong positive effects (d=.64). There were small additional positive effects for the added reinforcement in the reinforcement plus condition (d=.24). Learning was promoted with higher levels of generalized English vocabulary knowledge, as determined through the Test of Oral Language Development-P:4 scores on the Oral Vocabulary subtest (TOLD). Students scoring in the higher half of the sample with respect to the TOLD performed better on word learning in the exposed and instructed conditions, compared to students who scored in the lower half of the sample with respect to the TOLD (d=.45). However, when comparing the word learning of the students who scored in the higher half of the sample on the TOLD in the reinforcement condition to those students who scored in the lower half of the TOLD in the more intensive reinforcement plus condition, the gap lessens between the two groups (d=.21).

Higher levels of Spanish language proficiency, as determined through the Aprenda 3 scores, facilitated the acquisition of cognates. Although students in this study more readily acquired the cognates in the exposed condition, the difference between cognates and noncognates dissipated in the instructed conditions. This finding indicates that without instruction, students are bootstrapping onto their cognate knowledge, but when instruction intervenes, students learn both the cognates and noncognates.

Trends show that concrete words were more easily learned relative to abstract words. However, there were no significant differences between abstract and concrete words.

EFFECTS OF WORD AND STUDENT FACTORS ON INSTRUCTIONAL REINFORCEMENT IN A VOCABULARY PROGRAM IMPLEMENTED WITH SECOND GRADE BILINGUALS

by

Lauren Artzi

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2018

Advisory Committee:
Professor Megan Madigan Peercy, Chair
Ana Taboada Barber, Deans Representative
Jennifer Danridge Turner, Member
Jade Wexler, Member
Christopher Barr, Special Member

Dedication

To Eitan and Meirav, who are my continual inspiration and make me strive to be better every day of my life.

Acknowledgements

The research project reported on in this dissertation was supported by a grant HD039530 from the Eunice Kennedy Shriver National Center for Child Health and Human Development during its implementation. Thank you to the Center for Applied Linguistics, the institution for which this study was carried out, for their support and facilitation of this work.

I extend my sincere gratitude to the Principal Investigator of the research project reported on in this dissertation, Dr. Diane August and Co-Investigators, Drs. Christopher Barr and Paola Uccelli. Their enduring mentorship has facilitated not only this one endeavor, but my overall professional learning. I never imagined that the project I worked on in my very first research job would ultimately shape my own research interests. I would like to acknowledge Lindsey Massoud at the Center for Applied Linguistics, who collaborated with Dr. August, Dr. Uccelli, Dr. Barr, and myself on this project. Thank you to the staff at the school district as well as the teachers and students who made this study possible.

I am so grateful to my past and present colleagues, whose guidance over the years has been crucial to my professional development as a teacher and researcher. I am fortunate to work for, collaborate with, and learn from such extraordinary educators and scholars.

A heartfelt thank you to my advisor, Dr. Megan Madigan Peercy, and to my dissertation committee for their review, feedback, and guidance of this dissertation; for asking questions that prompted me to consider the best way to conduct this work; and for their continued support of my studies: Drs. Ana Taboada Barber, Jennifer Danridge Turner, Jade Wexler, and Christopher Barr.

I would not have set out on this path to a doctorate had it not been for Drs. Elana Shohamy and Ofra Inbar, my MA advisors at Tel Aviv University, who kindled my desire for research.

Most of all I am thankful for my friends and family, especially my parents and grandparents who instilled in me a love for learning; and, my husband Nir, who pushed me to pursue the doctorate while I worked at the same time, watched the children while I wrote, and patiently supported me throughout this lengthy endeavor.

My friends and family, colleagues and mentors, advisors and teachers have all walked this road with me. For all its twists and turns, we took this journey together. I look forward to what lies ahead.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Background

This study investigates the effects of a core vocabulary program developed for second grade English language learners (ELLs). It looks at instruction using methods that heavily reinforce a word versus reinforcement of the word to a lesser degree and no instruction. It also examines generalized English vocabulary knowledge, Spanish language proficiency, and type of word as possible influencers on the word learning in the program.

For students to have a moderate understanding of the text they are reading, research indicates that readers need to know approximately 90-95% of words on a given page of text (Nation, 2001). It is worth noting that recent college and career readiness (CCR) standards, such as the Common Core State Standards (CCSS), have heightened the level of academic texts and materials. Shifting to more rigorous academic texts in CCR standards likely increases the vocabulary demand necessary for adequate comprehension of academic texts -- as text complexity increases so do rare words, longer more multi-morphemic words, and more polysemous and ambiguous words (Hiebert, 2012; Nation, 2001).

Research shows that vocabulary instruction for ELLs is vital in order to support their vocabulary acquisition and reading comprehension. ELL students may present with more limited vocabularies in English, compared to their non-ELL peers (August & Shanahan, 2006). Since English vocabulary is a critical element in English reading comprehension and text levels have increased with new CCR standards, it is imperative for educators to promote ELL learning and acquisition of English vocabulary in the early elementary grades.

English Language Learners in U.S. Schools

According to a recent National Academies of Sciences, Engineering and Medicine (NASEM) report, approximately ten percent of school-aged children in U.S. schools are ELLs, and most of the U.S. ELL population speak Spanish as their home language (NASEM, 2017). Over half of ELL students are in lower elementary school grades (Batalova & McHugh, 2010).[1]

The U.S. Department of Education (USED) considers an ELL student to be a person "who was not born in the United States or whose native language is a language other than English; or who comes from an environment where a language other than English is dominant; or who is an American Indian or Alaska Native and who comes from an environment where a language other than English has had a significant impact on his or her level of English language proficiency; and who, by reason thereof, has sufficient difficulty speaking, reading, writing, or understanding the English language to deny such individual the opportunity to learn successfully in classrooms where the language of instruction is English or to participate fully in our society" (OESE, 2018, https://eddataexpress.ed.gov/definitions.cfm).

[1] When discussing ELL demographics, it is also worth noting that in U.S. schools, ELLs are a heterogeneous group of learners. ELLs vary in terms of home languages spoken, proficiency in the home language and in English, educational backgrounds of guardians, and socioeconomic status (SES) (August & Shanahan, 2006; Duursma et al., 2008; NASEM, 2017).

For the purposes of this study, ELLs are defined in much the same way as the USED definition. Because the students in this study were identified as ELLs using school district indicators at the time the research took place, they were receiving school-based services for English language development (ELD). In the district where this study took place, the ELD services were delivered through a transitional bilingual program model.

Transitional Bilingual Program Settings

The underlying goals of dual language programs, whether transitional bilingual or a longer-term developmental program, is to develop bilingualism and biliteracy while promoting grade-level academic achievement (NASEM, 2017). Unlike a longer-term developmental dual language program, the goal of transitional bilingual programs is to eventually transition ELL students to English-only instructional settings. Accordingly, transitional bilingual programs are designed to foster the first language in the early grades, which research shows helps to promote English acquisition, and ultimately transition the student to an all English environment by the middle elementary grades (NASEM, 2017).

In this study district's transitional bilingual program, a Spanish-English bilingual teacher used Spanish to instruct students in the content areas (math, science, and social studies). In addition to that instruction in Spanish, the ELL students in the program where this study was carried out were instructed in English for a daily ninety-minute English language development (ELD) block. That ELD block focused on developing English literacy skills. The intervention in this study took place in that ninety-minute ELD block.

Vocabulary's Effect on Reading Comprehension

Although ELLs are not a monolithic population, research tends to find that ELL students show difficulty comprehending grade-level English academic text, with those struggles persisting as students advance through school grades. Accordingly, some ELL students may have challenges acquiring academic content due to difficulties comprehending the grade-level texts used in those content areas (August & Shanahan, 2006).

Research also finds that ELL students may present with more limited vocabularies in English, compared to their non-ELL peers (August & Shanahan, 2006; Mancilla-Martinez & Lesaux, 2010; NASEM, 2017). As one component of oral language proficiency, limited English vocabulary can lead to difficulties with reading comprehension by the middle elementary grades (Nakamoto, Lindsey, & Manis, 2007; Nation, 2001).

Longitudinal investigations reveal that students, including ELLs, with limited English vocabularies in first grade tend to perform significantly below their peers in reading comprehension by third grade (Nakomoto, Lindsey, & Manis, 2007; Mancilla-Martinez & Lesaux, 2010). This research sheds light on the fact that though there are other elements in the reading process, vocabulary is one of the key components that needs to be promoted and strengthened in ELL students in order to benefit reading comprehension down the road. Furthermore, with more limited vocabularies in the early elementary grades, acquiring the meanings of new words from text and classroom discourse can be hindered (see Stanovich, 1986 for a review of the Mathew Effect explained in Chapter 2). Teachers can support ELL learning and reading comprehension by strengthening students' vocabulary, oral language proficiency, and use of reading comprehension strategies.

According to the research, the disparity between ELLs and their non-ELL peers with respect to the text-level literacy skills, such as vocabulary, oral language proficiency, and reading comprehension, does not tend to exist to the same degree in relation to word-level, code-based skills, such as decoding and word reading (August & Shanahan, 2006). While ELL students are apt to decode, or read the actual words on a page of text, as well as their non-ELL grade-level peers, the research generally demonstrates that they may have difficulties understanding what they are reading and understanding what the vocabulary words mean in a given text.

Limited English vocabulary in early elementary school can obstruct later reading comprehension in the middle and later elementary grades, when students acquire content information from grade-level text. Therefore, it is crucial for teachers to foster ELL students' vocabularies in the lower elementary grades (Lesaux, Kieffer, Faller, & Kelley, 2010; Mancilla-Martinez & Lesaux, 2011).

Consideration for Vocabulary Instruction

As part of a comprehensive approach to vocabulary instruction for all students, including ELLs, researchers recommend explicit and robust instruction of key school-relevant words and continued reinforcement of those meanings, instruction and reinforcement of word learning strategies to help student readers infer word meanings from text, immersion in language-rich environments, and promotion of learners' word consciousness (Baker et al., 2014; Gersten et al., 2007; Graves, August, & Mancilla-

Martinez, 2013). This evidence-based four-part approach to vocabulary instruction is described in Graves (2006) and Graves, August, and Mancilla-Martinez (2013).

Although the four-pronged vocabulary instructional approach is often found in the research to benefit all students, there are important instructional considerations when executing the four components with ELL students. These considerations, described in more detail in Chapter 2, relate to the need for teachers to account for and leverage the attributes of an ELL student's bilingualism as well as attend to the language acquisition needs of a student acquiring English.

While explicit robust vocabulary instruction with reinforcement garners significant moderate to high effects in past research studies (see Baker et al., 2014), the time needed to execute that level of instruction is an issue with robust instruction. Therefore, only a limited number of words can be targeted through such intensive instruction (Baker et al., 2014). To that end, researchers recommend that educators be highly judicious in selecting the words targeted for more robust vocabulary instruction.

To keep up with grade-level academic text, it is estimated that students schooled in the U.S. will need to acquire approximately 3,000 words per year (Graves, 2006; Nagy & Herman, 1984). Prior experimental research studies examining the impact of vocabulary interventions have targeted 8-15 words per (often weekly) unit, which works out to be roughly 400 words per year (Baker et al., 2014). As a result, 400 words per year represent approximately 10% of the total number of words school children need to acquire to keep pace with academic text.

That instructional time needed for robust instruction should be focused specifically on the words that are high-utility for the students (what Uccelli et al., 2015 refer to as “school relevant words”) and likely to not be acquired by the student without the additional instructional support robust vocabulary instruction provides. School relevant words are those words that are important for academic learning and text reading and are words that likely would not be acquired by the student without the additional instructional support (see Uccelli et al., 2015)

Researchers suggest that educators choose words that are high frequency in academic text, vital to the content area, and have properties, or attributes, that make them more cognitively arduous to learn; thus, likely require the more intensive instruction to be adequately learned by the student (e.g., Nagy & Hiebert, 2011). Exploring word characteristics related to vocabulary learning and acquisition can inform judicious word selection for vocabulary instruction.

Role of Word Properties in Vocabulary Instruction

Word characteristics are often parsed into two categories: syntactic and semantic (Nation, 2001). Syntactic properties are related to the word form itself, such as lexical categories or spelling patterns. Semantic word attributes are related to the word’s meaning, such as the degree to which a word is concrete or perceived through the senses. For example, the degree to which a word is concrete, perceived through the senses, or abstract, conveyed only through language, may make the word more or less arduous to learn for all students-- ELLs and non-ELLs (Nagy & Hiebert, 2011). If research finds that abstract words are more difficult to learn, then pedagogically speaking more instructional time should be spent on those more difficult and abstract words.

Another semantic attribute important to explore in specific relation to ELL student learning is cognate status, or the degree to which a word is shared across two languages. Research shows that ELLs acquire and learn cognates more quickly than non-ELLs (Genesee, 2006, NASEM, 2017). However, the research also shows that ELLs may not always draw on their cognate awareness knowledge when reading academic texts (Carlo et al., 2004). First language proficiency may factor into the degree students are able to bootstrap onto their cognate knowledge because a reader needs the cognitive lexical representation of the word from their first language for that cognate transfer to occur (August & Shanahan, 2006; Graves et al., 2013). Therefore, investigating student factors, such as first and second language proficiency, in addition to word-level factors can shed light on the nature of vocabulary learning in more targeted ways.

Role of Student Factors in Instruction

Person-level factors, such as bilingualism and first and second language proficiency, may also impact the ease of vocabulary learning and acquisition of some words. Prior research shows that cognates are often easier to acquire in a second language when there is cognitive representation of that word in a learner's first language (Nation, 2001). Other research shows that generalized English vocabulary knowledge, a component of oral language proficiency, promotes or hinders the process of acquiring new word meanings from text and discourse without additional instructional support (Stanovich, Nathan, & Vala-Rossi, 1986).

Reinforcement's Role in Instruction

McKeown, Beck, Omanson, and Pople (1985) raise the point that two components are important to explore with respect to vocabulary instruction: the type of instruction and the frequency the words are encountered, or reinforced, in the curriculum. Prior research indicates that if may take upwards of ten encounters with a word for the word to be sufficiently acquired by the student. However, that figure may also depend on student-level factors (Horst, 2013).

To that end, research is mixed on the level of reinforcement needed for a word to be adequately learned by school-aged learners. McKeown et al. (1985) found stronger outcomes for a sample of fourth grade students who encountered words twelve times in instruction compared to four times. Elley (1989) found effects for words that were more heavily encountered through one additional repeated reading of a children's book. However, in an ELL study with third and fourth graders in a summer school program, August, Artzi, and Barr (2016) did not find significant gains for the words learned in a condition of instruction that included additional reinforcement.

Accordingly, this work examines the level of instructional reinforcement (encounter with the word), type of word, and students' generalized English vocabulary knowledge and Spanish language proficiency in a core vocabulary program for second graders. For word-type, this work focuses on two semantic word properties: cognate status and level of abstractness (also referred to as conceptual complexity). This study examines two student-level factors: generalized English vocabulary knowledge and Spanish language proficiency.

Problem Statement

To synthesize the research for this dissertation, I examined the studies from the National Literacy Panel (NLP, August & Shanahan, 2006), two research syntheses from the NLP authors published after the NLP (August & Shanahan, 2009; August, Marcardle, & Shanhan, 2014), two extant Institute of Educational Sciences (IES) practice guides on literacy instruction for ELLs (Baker et al., 2014; Gersten et al., 2007), and another comprehensive synthesis conducted by the Center for Research on Education, Diversity and Excellence Center (Genesee, Lindholm-Leary, Saunders, & Christian, 2006). The syntheses referenced above represent resources currently in the field that systematically reviewed research studies related to ELL literacy instruction from the years 1980-2016. As an additional measure, I conducted a keyword search using the JSTOR and PSYCHinfo research portals at the University of Maryland library using phrases of *English learner, EL, ELL, vocabulary instruction and ELL.*

In the research body uncovered by my review, I was unable to identify any studies that examined the interaction between the level of vocabulary reinforcement, defined as the degree words were encountered in a curriculum, and word-level attributes of abstractness level and cognate status. I found no studies that also considered the student-level factors of English vocabulary knowledge and Spanish language proficiency in analyses of those studies.

Although there are four prior studies that have investigated level of reinforcement (August et al., 2016; Elley, 1989; McKeown, et al., 1985; Penno, Wilkinson, and Moore, 2002), these studies were not carried out with second grade ELLs. Three studies (Elley, 1989; McKeown et al., 1985; Penno et al., 2002) did not disaggregate findings by ELL status and more importantly did not examine the interaction between level of vocabulary reinforcement, word-type, and student factors of English vocabulary and Spanish language proficiency. The fourth study (August et al. 2016) was carried out with a sample of ELL students in third and fourth grade and implemented reinforcement as one additional encounter through a game. However, authors did not find effects related to that one additional reinforcement encounter in the game. Therefore, it is worth investigating instructional reinforcement implemented with an ELL sample at second grade in a more systematic way.

This investigation also adds to the body of ELL vocabulary instructional research. Across all the ELL vocabulary instructional studies I reviewed, the majority of the studies explored word learning in either very young learners, such as kindergarteners or first graders, or older learners such as upper elementary and middle school. Fewer studies examined word learning in the formative year of second grade.

By the time students reach the middle elementary grades they use academic texts to assimilate content and academic information to a greater degree. Therefore, second grade is a vital year for literacy development in elementary school. Second grade is also the year prior to the implementation of state achievement assessments in reading. Prior longitudinal research found that ELLs who demonstrate more limited vocabularies in the first grade initially kept pace with their peers in terms of comprehension skills; however, they fell significantly behind their peers in reading comprehension by third grade (Mancilla-Martinez & Lesaux, 2010; Nakomoto et al., 2007). Therefore, second grade is a critical year to investigate strategic vocabulary instruction.

In this investigation, I evaluate the relationship between type of word (cognates, noncognates, abstract words, and concrete words), level of reinforcement, defined as encounters with the target vocabulary through three instructional conditions (reinforcement plus, reinforcement, and incidental exposure), and the student-level factors of English generalized vocabulary knowledge and Spanish language proficiency. Accordingly, this study addresses three major research gaps: (1) the lack of studies that investigate vocabulary acquisition and learning in second grade ELLs, (2) the lack of studies that investigate outcomes of vocabulary instruction in relation to word-level attributes and student-level factors, and (3) paucity in the ELL research related to level of word reinforcement. The following research questions guide this work:

1. What is the effect of the three conditions (exposure, reinforcement, and reinforcement plus) on word learning?
2. Does type of word (abstract cognates, concrete cognates, abstract noncognates, concrete noncognates) influence students' acquisition and learning in each condition?
3. Does ELL word acquisition and learning across the three conditions vary as a function of generalized English vocabulary knowledge?
4. Does ELL word learning across the three conditions vary as a function of Spanish language proficiency?

The Study

Purpose

The study was conducted as part of the Acquisition of Vocabulary in English (AVE) project, a large research project aimed at investigating vocabulary instruction and word learning for lexicalized (words that carry meaning) and delexicalized (function) words in second grade ELL students. The AVE project was part of a broader research program called Vocabulary Instruction and Assessment for Spanish Speakers (VIAS), funded by the National Institutes of Child and Human Development (NICHD), an arm of the National Institutes of Health and the study itself was conducted through the Center for Applied Linguistics.

The AVE project took place over three years with the final year of the project designed as the culminating between-subjects investigation. In the final year the study was implemented with a sample of students across twelve classrooms who received the AVE program as the intervention versus a business-as-usual comparable sample of students.

As part of that culminating study in Year 3, there was an incorporation of a within-subjects subset of the larger principal study (referred to as substudy) allowing for the investigation of varying levels of word reinforcement (the conditions: reinforcement, reinforcement plus, and incidental exposure) connected to word-type (concrete cognates, abstract cognates, concrete non-cognates, and abstract non-cognates) of the lexicalized words. Additional information collected also allows for the exploration of student-level factors on the AVE program learning. The substudy reported on in this dissertation is referred to as AVE 3B (3 denotes the third year of the project and "B" indicates the subset study). As such, the research investigators and institution agreed to my use of the student assessment dataset to investigate the within-subjects substudy, AVE 3B, more closely.

Student measures for the AVE 3B investigation are described in Chapter 3 and include:

- Test of Oral Language Development IV;

- Aprenda 3 Spanish language proficiency test;
- A researcher-developed, curriculum-aligned measure that assessed intervention words, called the Acquisition of Vocabulary in English Vocabulary Assessment (AVEVA).

Fidelity to the AVE3B program was examined through the implementation of a fidelity protocol. That instrument measured not only instructional adherence to the AVE curriculum but quality of implementation as well.

Study Participants

Two hundred and twenty-eight students across twelve classrooms participated in the AVE3B substudy (reported on in this dissertation). All participants were second grade Spanish-speaking ELLs attending schools in a district in the Southwestern United States. In the study schools, over 75% of the students received free and reduced lunch. Students participating in the AVE3B study were receiving instruction through a transitional bilingual program in the Southwestern region of the United States. Instruction for these students took place predominately in Spanish with one ninety-minute instructional ELD block. The AVE vocabulary program was delivered daily for 20 weeks during the 90-minute ELD block. It required implementation by the classroom teachers for four out of five days each week.

My Role

I worked on the project as the lead research associate from its inception to completion—September 2009 to May 2013. In that role, I managed the curriculum-and assessment development for the project. As part of the research team that included Dr.

August and co-principal investigators Dr. Christopher Barr, Dr. Paola Uccelli, and Lindsey Massoud, I contributed to the research design across the project waves, implementing elements of the design and helping to operationalize the design through the curriculum development. With the investigators oversight, I led the curriculum development, co-developed the researcher-developed measures of vocabulary, and helped conceptualize the within-subjects word-matching scheme for the AVE3B substudy.

Rationale for the Use of the AVE Data

Using the AVE3B data allows me to examine a sample of second grade ELL students within the context of an experimental intervention study that utilized evidence-based practices for vocabulary instruction. With a large sample of Spanish-speaking ELL students in a transitional bilingual program, the AVE3B dataset is a robust data source that allows me to examine all four of the research questions noted in this chapter. Because Spanish is the primary home language of many U.S. school children who are ELLs, the sample of student data in the AVE3B dataset reflects the overall demographics of many schools across the United States.

Moreover, exploring data with respect to ELL students in the early elementary grades, specifically second grade, is valuable because research indicates that the early elementary years are critical years in terms of vocabulary development and reading comprehension (August & Shanahan, 2006). Since second grade is a period when students begin using academic text to assimilate content knowledge to a greater degree and there is a paucity of vocabulary-oriented studies with second grade ELLs, second grade is a year that is of particular interest to me.

From an English language development perspective, second grade is a significant year to study. In a transitional bilingual program structure, second grade is often the year before instruction fully transitions to English. Therefore, understanding the mechanisms to promote vocabulary in the pre-transition year of second grade is imperative.

Investigation

Within the ELD block of the district transitional bilingual program, the AVE3B intervention took place in core instruction, meaning that the AVE3B intervention was delivered to all students, not only students receiving intervention instruction support as part of a multi-tiered systems of supports (MTSS) model (see Center on RTI, 2011, for explanation of tiered levels of instruction within an MTSS system).[2] The AVE3B intervention consisted of nine thematic units that were aligned to state English language arts (ELA) content standards. The units were delivered over four days a week for twenty weeks, for approximately 20 minutes per day. A tenth two-week unit was designed to review the program words from the prior weeks.

Words targeted in the AVE3B study were split into three conditions: reinforcement (six word encounters), reinforcement plus (fourteen word encounters), and incidental exposure (no instruction). Informed by the work of educational research such as Nation (2001) and Nagy and Hiebert (2011), and psycholinguistic work, such as Schwanenflugel (1991), the words were theoretically equated using a number of variables

---

[2] In Multi-tiered Systems of Supports (MTSS) instruction is delivered through tiered levels of supports (Center for RTI, 2010). In the MTSS model core instruction should sufficiently serve 85% of students. Databased decision making, which is using screening and progress monitoring data, is used to determine students who would benefit from receiving additional instructional support through Tier II intervention (also referred to as supplemental intervention). In the MTSS model, Tier II intervention should meet the needs of 10% of those students receiving Tier II support. About 5% of the students may still struggle. For those students not responding to Tier II intervention, instructional intensification is necessary. This is referred to as Tier III intensive intervention support.

such as part of speech (POS), word frequency indicators from the Educator's Word Frequency Guide (Zeno et al., 1995), and word knownness indicators from the *Living Word Vocabulary* (LWV) (Dale and O'Rouke, 1981)[3].

Given the unique within-subjects design of the study, equating the words through the variables of POS, word frequency, and word knownness remained important to ensure that all three conditions were equal in theoretical vocabulary difficulty at pretest (wave 1). The words were also stratified by cognate status and abstractness status bringing about the four types of words under investigation: concrete cognates, concrete non-cognates, abstract cognates, and abstract non-cognates. Stratifying the words by cognate status and abstractness (conceptual complexity) status allows for the investigation of learning within condition (exposure only, reinforcement, reinforcement plus) by those four types of words (see Appendix B for the words by type of word and instructional condition) through a within-subjects design methodology that is concurrent in nature.

A priori theoretical difficulty of each condition was also empirically investigated and there were no detectable statistical differences between the three conditions (exposure, reinforcement, and reinforcement plus) at pretest. To assess post-intervention vocabulary and language outcomes, students were given a curriculum-aligned vocabulary measure as well as the Test of Oral Language Development IV (Hammill & Newcomer, 2008), a metric of English oral language proficiency. Student scores on the Aprenda 3 (Pearson, 2005), an assessment of Spanish proficiency, were also obtained, allowing for student-factors to be examined as possible influencers on the word learning in each

[3] Word knownness is one's demonstrated ability to understand a word meaning. Garnered from assessment outcomes across the United States, the LWV provides data on the percentage correct of students who knew the words at grades 4, 6, 8, 10, and 12 (see Dale & O'Rouke, 1981 for more information on the assessment procedures of the LWV).

condition by the four types of words (concrete cognates, abstract cognates, concrete noncognates, abstract noncognates).

Significance

This study explores the effectiveness of various levels of instructional reinforcement for words in a core read aloud program for second grade ELLs that utilized evidence-based practices for vocabulary instruction. This study investigates the impact that type of word, stratified by dimensions of cognate status and abstractness status, has on vocabulary learning acquisition. It also considers student-level factors, such as generalized English vocabulary and Spanish language proficiency, as influencers on the word learning.

Although there is experimental research focused on the development of vocabulary in ELLs (e.g., Gersten et al., 2007; Baker et al., 2014; Silverman et al., 2017), and other work that compares vocabulary instructional conditions that differ on levels of reinforcement with respect to instruction (i.e., August et al., 2016; Beck, Omanson, & McKeown, 1982; Elley, 1989; Penno et al., 2002), there is no work that empirically explores the interactions between reinforcement conditions, word-level properties, and student-level factors on an ELL sample at second grade. Because there is often limited time for instruction, it is important for educators and researchers to understand whether words of various types and properties interact with the level of instructional vocabulary reinforcement, or encounters, in the curriculum. It is also important for educators and researchers to understand if that interaction varies by generalized English vocabulary knowledge and Spanish language proficiency of the students.

Chapter 2: Review of Literature and Theoretical Framework

Introduction

There is a strong empirical relationship in the research between English vocabulary knowledge and English reading comprehension (August & Shanahan, 2006; Nakamoto, Lindsey, & Manis, 2008; NASEM, 2017). However, research tends to find that ELLs may display more limited vocabularies in English, which likely causes difficulties for reading comprehension (August & Shanahan, 2006; Mancilla-Martinez & Lesaux, 2010; Nakamoto, Lindsey, & Manis, 2007). National assessments illustrate this trend. On the 2013 National Assessment of Educational Progress (NAEP), for example, fourth and eighth grade scores in reading comprehension were significantly associated with scores in vocabulary and ELLs made up a large portion of the lowest-performing percentiles (NCES, 2013).

It is projected that a reader needs to access the meanings for 90-95% of the words on a page to sufficiently comprehend that text (Lauffer, 1998; Nation, 2001). Because English is a vocabulary-rich language, characterized by a dense and ranging lexicon, researchers estimate that school children in the U.S. need to acquire approximately 3,000 words per year to keep pace with academic grade-level text (Graves, 2006; Nation, 2001).

Robust explicit vocabulary instruction, also called by researchers extended vocabulary instruction (e.g., McKeown et al., 1982; Coyne et al., 2007), is a vocabulary instructional method that promotes students' deep processing of a word. With robust instruction students are given multiple opportunities to encounter the word, see images that illustrate the word, engage in demonstrations or the use of multimedia with the word, hear examples and nonexamples of the word, and use the word in discussions, reading, writing, and listening activities. In the research extended (robust) vocabulary instruction is shown to promote word knowledge as well as reading skills (August & Shanahan, 2006; Baker et al., 2014; Beck et al., 2007; Gersten et al., 2007). However, robust vocabulary instruction requires ample instructional time, an important limitation of that instructional approach. In an experimental instructional study, Coyne et al. (2007) found, for example, that it took as many as five minutes to initially instruct a vocabulary word through direct and robust means versus one minute to provide an in situ brief explanation of the target word. Other research finds that it takes multiple encounters with a word for it to be sufficiently acquired by the student (McKeown, Beck, Omanson, & Pople, 1985).

As such robust vocabulary instruction, the kind that is multimodal and accompanied by deep processing activities and multiple encounters with the word (see Baker et al., 2014; Gersten et al., 2007), is time-intensive. Consequently, a limited number of words can be targeted through robust vocabulary instruction with high levels of reinforcement or encounters with the word-- a main shortcoming of that vocabulary instructional method.

Research is mixed regarding the degree words need to be reinforced, or encountered, for them to be sufficiently learned by students. McKeown et al. (1985) studied varying levels of reinforcement, defined by the number of encounters with a given word, to determine if words need to be reencountered to a high degree by elementary-aged learners. McKeown et al. found the words that were more heavily encountered, up to fifteen times in the lesson, garnered the greatest gains; however, the study participants also learned the moderately encountered words to a significantly greater degree compared to the words that received no instruction.

Word encounter studies also take the form of repeated reading explorations. For example, Elley (1989) studied three reading conditions. In one group students listened to a story one time without additional instruction related to the words, a second group listened to the story once and the teacher provided in situ parenthetical explanations for the words during the read aloud. A third group of students in Elley's study listened to the story three times and in that condition the teacher also provided the brief word explanations during the read aloud session. The average percent gain in vocabulary for students in the first two conditions, where students heard the story once, were comparable-- 19% for the first condition and a 20% gain for students in the second condition. Students who heard the story three times, however, learned a significantly greater amount of words compared to the first two conditions--a 33% gain in word learning from pretest to posttest.

Penno, Wilkinson, and Moore (2002) studied vocabulary learning and acquisition from shared book reading with 5- to 8-year-old children in elementary school using similar methods. Comparable to the findings that Elley reported, the students in the Penno et al. study who heard a story read three times instead of one time showed greater acquisition of the target vocabulary words and deeper knowledge of the words. In that study, researchers found a linear effect for each additional rereading of the story--the greatest mean gain occurred with the students in the condition that heard the story read three times with in situ parenthetical explanations by the teacher for vocabulary words in the story. Both of these studies (Elley, 1989; Penno et al., 2002) reveal that the frequency the word was encountered by the student participants in the form of rereading positively impacted outcomes and that the brief explanations provided by the teacher increased the magnitude of learning. These findings corroborate the findings by McKeown et al. (1985) --the greater a word is encountered, the more likely it is to be learned by students in early elementary.

In a study carried out with a sample of ELL third and fourth grade students, August, Artzi and Barr (2016) designed a summer school intervention for third and fourth grade ELLs in a summer school program. They investigated robust vocabulary instructional methods against leaner embedded instruction, where in-situ parenthetical explanations were given during a shared reading event by the teacher, for domain-specific and general academic words. As an additional component to the study, the research team also explored whether additional reinforcement of the vocabulary garnered better outcomes; however, no significant differences for the additional reinforcement condition were found.

In that study (August et al., 2016), reinforcement was conceptualized as one additional end-of-unit game, and the authors hypothesized that the reinforcement did not add additional value to the student outcomes because either (1) the program's vocabulary was adequately learned without the additional reinforcement or (2) the reinforcement activities were themselves not effective. Because students in that study did not score at the ceiling with respect to the vocabulary outcome measure, the authors believe that the activities themselves were less effective or perhaps the way in which the reinforcement condition was conceptualized, as one additional encounter through a student-directed game activity, was not effective (see August et al., 2016).

Although August et al. (2016) did not find effects for reinforcement, the findings in McKeown et al. (1985), Elley (1989), and Penno et al. (2002) illustrate that more vigorous reinforcement yields greater effects for word learning. Those studies also illustrate greater effects for words that are encountered with less reinforcement compared to no instruction. In the context of implementing strategic vocabulary instruction it is useful to determine the kinds of words that need the more intensive instructional reinforcement, the types of words that are sufficiently served through fewer encounters, and words that students are likely to acquire without any additional instructional assistance.

Raising the point that vocabulary is a heterogeneous construct, Nation (2001) and Nagy and Hiebert (2011) explain that different word-types have different learning burdens and therefore might require different intensities of instruction. For example, concrete words, such as *apple*, that can more easily be mentally imaged and perceived through the senses, may be less laborious to acquire and learn compared to abstract words that require language to convey the true sense of the word, such as the word *efficient.* Carlo et al. (2004) found that cognates, words that are orthographically and phonologically similar across two languages and present with similar meanings, tend to be easier to acquire by someone who speaks both of those languages.

In another study, August, Artzi, Barr, and Francis (2018) examined the word properties of cognate status and abstractness status with respect to the level of vocabulary instructional intensity of embedded instruction (in situ brief explanations provided for the words) versus more robust and extended vocabulary instruction (multimodal deep-processing instructional methods). That study was carried out with a sample of second grade ELLs in a transitional bilingual program in the Southwest using the AVE instructional program. The researchers found a significant mean difference for all comparisons at posttest with an 18% correct for concrete cognates (*d* = 1.2), 25% correct for concrete noncognates (*d* = 1.6), 18% correct for abstract cognates (*d* = 1.2), and 18% for abstract noncognates (*d* = 1.3). Further, the research team found no significant differences between concrete cognates for the embedded brief explanations for the words (embedded instruction) compared to the more robust (extended) instruction. This finding indicates that the investigated word-types, stratified by dimensions of cognates status and abstractness status, might be differentially learned by ELL students in second grade. For example, study findings indicate that concrete cognates might be a word-type that is sufficiently acquired by ELLs through leaner vocabulary instructional methods, such as embedded brief explanations, compared to other word-types that need the additional multimodal, and deep-processing instruction that characterizes robust instruction, such as abstract noncognates. Therefore, investigating the influence of the four types of words (concrete cognates, abstract cognates, concrete noncognates, abstract noncogntes) with respect to ELL word learning is worthwhile.

However, word-level differences may not be the only influencers of vocabulary learning and acquisition. Nation (2001) specifies that the learning burden of words likely interacts with person-level factors and may shift based on one’s background characteristics, such as language background and proficiency. ELLs bring the attribute of their bilingualism with them to their learning opportunities and that asset might be critical in vocabulary acquisition (see Moll, Amanti, Neff, & González, 1992 and 2005, for a discussion related to leveraging learning attributes, or funds of knowledge).

With respect to cognate learning in bilingual students, Ramírez, et al. (2013) and Genesee et al. (2006) found evidence for cross-language transfer with Spanish-English bilingual learners; although, word-level factors concerning the degree the cognates were orthographically and phonologically related across both languages mediated the magnitude the transfer occurred. In those studies, student-level factors also mediated the cross-language transfer relationship, including the proficiency level of the first language, Spanish, and second language, English. These findings demonstrate that in order to bootstrap onto their cognate knowledge, student participants in those studies needed a sufficient level of first language proficiency and the cognitive representation of the word in their first language.[4] Therefore, the type of word (i.e., abstract and concrete words and cognates versus noncognates) and levels of instructional reinforcement can be more thoroughly investigated when considering person-level factors, such as language background and first language proficiency, in the research analytics.

The study reported on in this dissertation extends the work of August et al. (2016), August et al. (2018), McKeown et al. (1985), Elley (1989), and Penno et al. (2002) through studying reinforcement levels in relation to type of word (defined as degree of abstractness and cognate status). In addition to looking at type of word and pre-to-posttest (wave 1 to wave 2) growth related to level of reinforcement, this work investigates the way in which person-level factors such as language proficiency impact the outcomes by instructional condition and type of word.

---

[4] Genesee et al. (2006) also found negative transfer effects for false cognates. False cognates are words that show orthographic and phonological similarities across both languages, but have different meaning (i.e., *pie* in English versus *pie* in Spanish, which means *foot*).

In this chapter, I review the evidence from existing research that supports each component in this study: (1) the key role vocabulary plays in reading comprehension and academic success, (2) the way in which I am theoretically situating academic vocabulary knowledge, and (3) current vocabulary instructional research framed by a balanced literacy approach. The following review explores these topics in three sections: (1) vocabulary knowledge and the lexical levels, (2) components of strategic vocabulary instruction, and (3) theoretical foundations.

I used a rapid evidence assessment (REA) approach to synthesize the research base. An REA begins with a systematic review of extant reports, meta-analyses, and other significant syntheses, with specified criteria for inclusion. REAs may include important literature from practice and policy sources as well as the research sources. For inclusion in the REA study database, I utilized the following criteria: (1) studies are focused on students in elementary school or middle school, included an element of vocabulary instruction, and reported at least one vocabulary-based outcome; (2) when reporting on practices for ELLs, the studies disaggregated study data by ELL status, although the criteria for ELL status may differ across the studies; and (3) the studies were published in peer-reviewed articles or conference papers, policy briefs, practice guides, research syntheses, chapters, or books, following professional evidence criteria standards and REA specifications.

The existing syntheses I used for this Chapter include studies from the National Literacy Panel (August & Shanahan, 2006), studies from two more recent syntheses from the NLP authors (August & Shanahan, 2009; and August, McCardle, & Shanahan, 2014), a Center for Research on Educational Diversity and Excellence (CREDE Center) synthesis (Genesee et al., 2006) and two Institute of Education Sciences (IES) practice guides (Gersten et al., 2007; Baker et al., 2014). These six syntheses are widely referenced in the field and capture experimental research from 1980 to 2013. As an additional measure, I keyed the search phrases *English language learners*, *English learners, second language learners, and vocabulary instruction* into the PSYCHinfo and JSTOR portals at the University of Maryland library site.

Often, REA inclusion criteria allow for the addition of promising practices as well as evidence-based practices in a study pool. To that end I used the Collaboration for Effective Educator Development and Reform (CEEDAR) Center Evidence Standards (see, http://ceedar.education.ufl.edu/wp-content/uploads/2014/08/Evidence-Based-Practices-guide.pdf) because those standards consider a range in research evidence standards, and allows for the inclusion of promising practices. Accordingly, for this review I included studies in submission and papers presented at national conferences as references to promising practices in vocabulary instruction and research.

Literature Review

Vocabulary is multifaceted construct. Expressive vocabulary knowledge, or the active use of vocabulary, is distinct from receptive vocabulary knowledge where someone listens to or recognizes a word (Nation, 2001). Meara (2002) points to larger lexical units as important components in vocabulary knowledge. Zeroing in on the word-level only, where this specific investigation is situated, there is still significant heterogeneity in the pool. There are, for example, domain-specific and general academic words (Nation, 2001; Nagy & Hiebert, 2011).

Domain-specific words, or technical words (i.e., *photosynthesis, hypotenuse, legislation*), are those that are mainly found in one subject area and are characterized by low cross-subject dispersion (Nation, 2001). Compared to words that cross-cut subject areas, domain-specific words tend to be lower frequency in school text; yet the domain-specific words are also often conceptually key to content-area learning.

General academic words (i.e., *identify*, *confidence, excellent*), where this investigation lies, cross-cut subject areas and are characterized by high cross-subject dispersion. This pool of words constitutes significantly greater coverage in academic text compared to domain-specific words. For example, close to 95% of words in academic text come from the most frequent 5,000 general academic words (Nation, 2001). While domain-specific words are important for content acquisition, general academic words are vital for comprehension of academic texts used in school (Baker et al., 2014).

Vocabulary Knowledge and the Lexical Levels

Literacy skills can be classified in terms of the lexical levels (e.g., Bowers, Kirby, & Deacon, 2010; Goodwin & Ahn, 2010; Kirby et al., 2012). Sublexical literacy skills are code-based skills that include phonological processing, phonological awareness, decoding, and encoding. These are the skills associated with units that do not carry meaning but are key with respect to ensure fluent reading. On the other hand, lexical skills such as vocabulary knowledge and the supralexical skill of reading comprehension involve processing meaning-related language units (Bowers et al., 2010).

The NLP (August & Shanahan, 2006) classifies literacy skills in terms of word- and text-level skills. Sublexical skills fall under the NLP's word level skill category, and lexical and supra-lexical skills are at the NLP's classification for text-level category.

Figure 1 visually depicts the lexical levels crossed with the NLP classifications of word and text skills.

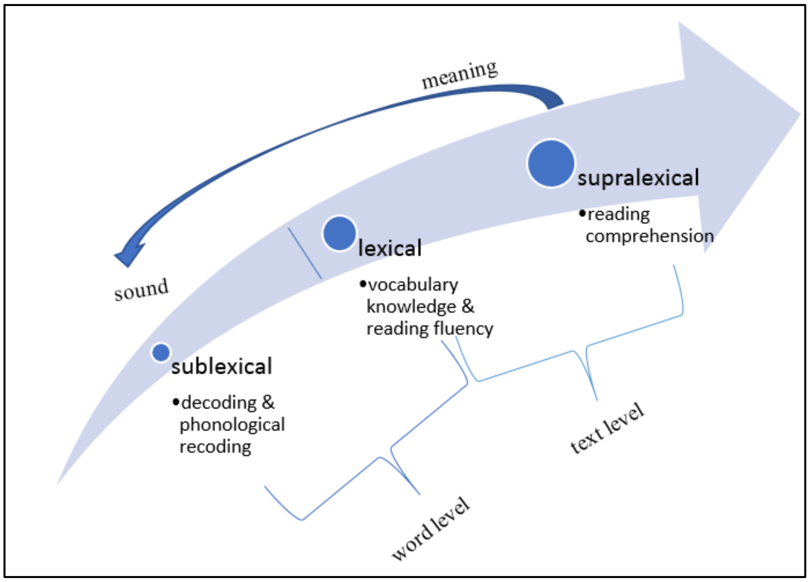


Figure 1. The Lexical levels parsed by the NLP classifications

The following sections detail research related to vocabulary at each of the lexical levels.

Vocabulary's Role on Sub-and-Lexical Skills

Vocabulary knowledge and oral language proficiency aids in student development of the sublexical skills because decoding a word where a reader knows the meaning of that word strengthens the word's orthographic depiction in that reader's cognition (Perfetti, McKeown, & Kucan, 2010). A reader's phonological awareness and word reading are two other sublexical skills that are also facilitated by higher levels of vocabulary knowledge because knowing the meaning of a word helps support readers in recognizing the word's phonological components and make better sense of the sound-letter relationships of that word (Lonigan, Burgess, & Anthony, 2000; Walley, Metsala, & Garlock, 2003).

Other research finds that vocabulary knowledge impacts acquisition at the lexical level. Lower or higher vocabulary levels influence new word acquisition from reading at the lexical level that escalates over time (Nicholson & Whyte, 1992). Stanovich (1986) refers to the lexical impact of more limited vocabulary as the Mathew Effect. In the Mathew Effect, Stanovich theorizes that students with more limited vocabulary tend to read less and in turn pick up less vocabulary through reading. Consequently, there is a compounding effect where a gap in vocabulary levels widens between struggling readers and their more reading proficient peers. As such, the vocabulary and reading comprehension gap for students with more limited vocabularies expands as children advance through school.

Prior studies that have studied word acquisition and learning of ELLs and non-ELLs at the lexical level indicate the possible presence of a Matthew Effect. Cain, Oakhill, and Bryant (2004), for example, found that monolingual children presenting with vocabulary limitations, as determined by student-level outcomes, portrayed difficulty inferring word meanings from context using the written support of that reading passage. In a longitudinal investigation conducted with ELL children in the early elementary grades, Mancilla-Martinez and Lesaux (2011) found that ELLs with limited vocabularies in the lower elementary grades, continued to present with restricted vocabularies in third grade. The authors surmise that the disparity in vocabulary levels persisted in part because discrepancies in vocabulary knowledge at kindergarten was perhaps too great for those students to overcome by third grade.

Vocabulary's Role in Supralexical Skills

Increased levels of vocabulary knowledge are associated with academic success and discourse. Research shows a link between vocabulary knowledge and content-area acquisition as measured through academic GPA (Baumann, Kame'enui, & Ash, 2003; Carrier, 2013 Saville-Troike, 1991) and vocabulary knowledge and academic discourse skills (Uccelli et al., 2015).

There is also a robust relationship between receptive and productive vocabulary knowledge and overall reading comprehension for non-ELLs (Geva & Genesee, 2006; Pearson, Hiebert, & Kamil, 2007; Pressley et al., 2000; Oulette & Beers, 2010; Swart, et al., 2017) as well as ELLs (Kieffer & Lesaux, 2012; Proctor et al., 2005; Nakamoto, Lindsey, & Manis, 2007). Research indicates that the vocabulary-reading comprehension relationship strengthens as the child moves through school (Oulette & Beers, 2010).

Nakamoto, Lindsey, and Manis (2008) and Mancilla-Martinez and Lesaux (2011) used growth modelling to study vocabulary and reading associations of ELL elementary-aged learners, using expressive and receptive measures of vocabulary. Although first grade ELLs with more limited vocabularies kept pace with grade-level peers in terms of reading and listening comprehension in those studies, the students significantly fell behind their grade-level peers in reading comprehension by third grade. This longitudinal work points to the fact that ELLs who portray more limited English vocabularies in the early elementary years, such as kindergarten and first grade, may likely struggle with comprehending English texts more significantly in third and fourth grade, those middle elementary school years where students begin to transition from learning how to read text to using academic text to assimilate content knowledge.

Other work demonstrates vocabulary knowledge to be a more salient predictor of reading comprehension for second language learners (Shanahan & Beck, 2006; Verhoeven, van Leeuwe, & Vermeer, 2011; Verhoeven & Perfetti, 2011). While background knowledge predicted more variance in second language learners, compared to monolinguals, Verhoeven et al. (2011) and Verhoeven and Perfetti (2011) found that vocabulary knowledge predicted significantly more variance in reading comprehension of second language learners, compared to monolinguals.

Exploring the impact of first language factors on English vocabulary, Proctor, August, Carlo and Snow (2005) examined constituents of Spanish language proficiency related to English literacy skills. In addition to the main effect for Spanish vocabulary knowledge, authors also found a significant interaction between Spanish vocabulary and English fluency–the students who were more fluent readers in English showed greater Spanish vocabulary knowledge compared to the students who were less fluent in English reading. This finding reveals a transfer effect from the first language (Spanish) vocabulary knowledge to the second language (English) reading ability. It also indicates the facilitative effect of the first language in aiding acquisition of English at the supralexical level. Nagy, Garcıá, Durgunoglu, and Hancin-Bhatt (1993) studied the literacy skills of Spanish-English bilingual children in later elementary grades and found a significant association between students' ability for identifying Spanish/English cognates, words that are shared across both languages, and their English reading comprehension.

As such, strategic instruction of key school-relevant vocabulary and instruction that makes the best use of time and leverages cognate awareness is vital for ELLs. The next section describes the research-based vocabulary instructional components found effective for ELLs in primary grades.

Vocabulary Instruction for ELLs

In the most recent Institutes of Education Sciences (IES) practice guide for supporting literacy of ELLs, Baker et al. (2014) recommend the use of explicit instruction of vocabulary augmented with word learning strategy instruction and indirect word exposure through a language-rich environment. With elements of a balanced literacy approach, where teacher directed methods are implemented in tandem with indirect methods, student-directed methods, and exposure to a language-rich environment, a strategic vocabulary instructional approach should entail a judicious selection of high-utility academically relevant words for explicit instruction balanced by immersion in a language-rich environment, promotion of word awareness, and the use of strategy instruction to help students become more effective independent word learners as part of indirect learning. These four prongs to vocabulary instruction are outlined in Graves (2006) and in a later resource centered on vocabulary instruction for ELLs, Graves, August, & Mancilla-Martinez (2013) (see Figure 2 for a visual depiction of the four-pronged approach to vocabulary instruction).

These four prongs to vocabulary instruction align well to the CCSS language strand standards four through six. ELL and non-ELL studies across elementary and middle school have drawn on some, if not all, of the components of the four vocabulary instructional prongs shown in Figure 2 and the majority of these studies have garnered significant positive effects (see Tables 1 and 2; Graves, 2006; Graves et al., 2013).

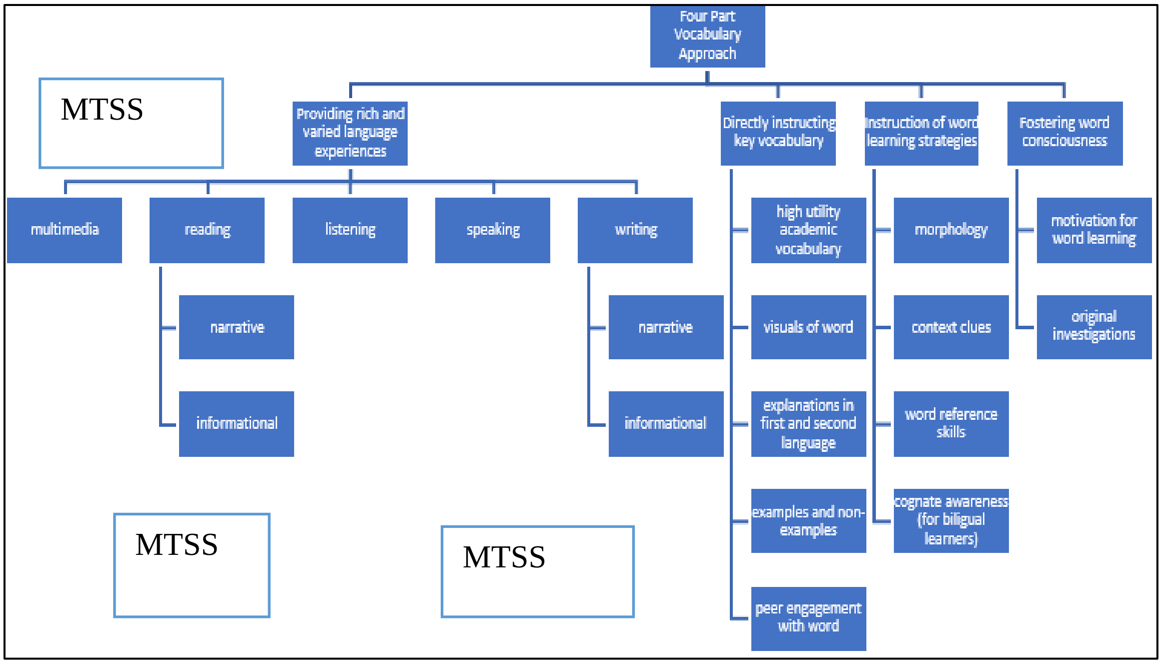


Figure 2. Four theoretical prongs of vocabulary instruction, outlined by Graves (2006) and Graves et al. (2013)

Situating the four prongs to vocabulary instruction, shown in Figure 2, within multi-tiered systems of supports (MTSS) illustrates the nature of vocabulary instruction in today's instructional landscape, where there is prevalence for tiered instructional frameworks. MTSS is a systems-level, data-driven, instructional approach to address the needs of all learners. Many schools, local education agencies, and state agencies are utilizing MTSS frameworks to support learning for all students. This tiered approach to instruction is increasingly applied to general education because it is a unified data-driven system that identifies students at risk for learning difficulties in reading and math and directly addresses their learning and behavioral needs (see National Center on Response to Intervention, 2010 for a review of MTSS elements).

As an instructional framework, MTSS provides a means for which students who need more support to successfully meet grade-level standards with respect to core instruction in reading and math receive increasingly intense intervention. Those students' learning needs are addressed through high-quality evidence-based core instruction (often referred to as Tier I instruction) and the possible use of research-based intervention (often referred to as Tier II instruction) for struggling learners. The intervention support can also be intensified and individualized for students who are not responding to Tier II instruction, with Tier III instruction often referred to as more intensive instruction.

Within a healthy MTSS framework, at least 75-80% of all students will demonstrate appropriate academic growth when provided high-quality core instruction at the Tier I level, an additional 15-20% of students will benefit from some intervention supports at the Tier II level, and 5-10% of students will require more intense level of intervention at the Tier III level, personalized to their individual academic and behavioral needs. Therefore, the four theoretical prongs to vocabulary instruction depicted in Figure 2 should be evident across all MTSS levels, or tiers, with data and formative assessment driving the decision-making pertaining to vocabulary instruction.

With respect to the vocabulary instruction studies uncovered from the REA literature search, Table 1[5] shows those studies cross walked with the categories of the Graves' four prongs to vocabulary instruction as well as the inclusion of instructional practices as part of an MTSS framework in intervention settings.

---

[5] Following REA parameters, findings to the studies in Tables 1 and 2 are reported by program and not citation.

Table 1. Four-Pronged Approach to Vocabulary Dot Plot of REA ELL Studies

Study (Year)	Explicit Instruction	Exposure/ rich language	Reinforce-ment (encounters)	Word Learning Strategies	Multi-media	Tier II or III Vocabulary instruction
August et al. (2014)	●			●		
August et al. (2009)	●			●		
August et al. (2016)	●	●	●	●		
August et al. (2018)	●			●		
Biemiller & Boote (2006)	●	●				

Study (Year)	Explicit Instruction	Exposure/ rich language	Reinforce-ment (encounters)	Word Learning Strategies	Multi-media	Tier II or III Vocabulary instruction
Burns (2011)	●	●				●
Carlo et al. (2004)	●			●		
Calderon et al. (2004)	●	●			●	
Crevecour et al. (2014)	●	●				
Chambers et al. (2006) **Chambers et al.**					●	
Gunn et al. (2000)	●	●				
Gunn et al. (2005)		●				●
Fillipini (2007) Fillipini et al. (2012)	●					
Howard & Arteogotia (2013, 2014)				●		
Kieffer & Lesaux (2012)				●		

Study (Year)	Explicit Instruction	Exposure/ rich language	Reinforce-ment (encounters)	Word Learning Strategies	Multi-media	Tier II or III Vocabulary instruction
Lesaux et al. (2014)				●		
Lesaux et al. (2010)				●		
Nelson et al. (2011)		●				●
Paéz et al. (2013, 2014)	●	●		●		
Peercy et al. (2015) Silverman et al. (2017)	●			●	●	
Proctor et al. (2009)	●	●		●	●	
Ross & Begeny (2011)	●					
Silverman & Hines (2009)	●				●	
Taboada & Rutherford (2011)	●			●		

Study (Year)	Explicit Instruction	Exposure/ rich language	Reinforce-ment (encounters)	Word Learning Strategies	Multi-media	Tier II or III Vocabulary instruction
Townsend & Collins (2009)	●					
Uchikoshi (2005)	●	●			●	
Uccelli & Rosenthall (2013)	●	●				
Vadasy et al. (2013)		●				●
Vadasy et al. (2015)		●				●
Vaughn et al. (2009)	●					●
Vaughn et al. (2006)	●					●
Wijukamer et al. (2016)					●	

As Table 1 depicts, the bulk of the studies I uncovered investigated robust instruction, followed by instruction with word-learning strategies. Few studies looked at leaner instruction methods, such as embedded instruction, and no studies investigated the level of instructional vocabulary reinforcement. Four studies drew on multi-media platforms, either a universal design environment or video platforms. Six studies took place in either Tier II or Tier III intervention instruction found in MTSS frameworks, with the rest of the studies taking place in core instruction. Most of the studies investigating word learning (morphology, context clues, dictionary skills, and cognate awareness) executed the word learning strategies together.

Table 2 provides an overview of each of the REA studies highlighted in Table 1 and includes the grades as well as the outcomes and findings.

Table 2. Overview of REA ELL Studies 2000-2016

Study by Program	Research Design	Findings	Grade	Vocabulary Area	Source	Study Focus Area
August et al. (2014)	-Between- subjects, business as usual -40 minutes per day, 5 days per week, 18 weeks -1,309 students, 353 were ELLs	-Group differences between treatment and control for curriculum-based vocabulary (b= 2.32, p = .05) -No effect for science (b = –0.03, p = .97) or for standardized assessment vocabulary (b = –0.01, p = .98)	7th	English vocabulary and science knowledge	Peer-reviewed journal	Study focused on science knowledge and academic vocabulary development
August et al. (2009)	-Between- subjects, business as usual -40 minutes per day, 5 days per week, 18 weeks -562 ELLs	-Group differences on a researcher developed measure (d=.69, Word Association (d=.27) -Group differences on Academic Word Meanings in Context (d=.22)	6th	English vocabulary and science knowledge	Peer-reviewed journal, IES Practice Guide	Study focused on science knowledge and academic vocabulary development
August et al. (2016) **August et al. (2013)**	-Within- subjects, extended versus embedded vocabulary instruction -20 min per day, 5 days per week, 5 weeks in a summer school program for struggling learners -509 ELLs	-Group differences on the researcher-developed curriculum-based measure: robust instruction (g = 1.7); leaner instruction (g= .57)	3rd and 4th	Domain specific and general academic lexemes	Peer-reviewed journal, Policy brief, peer reviewed book chapter	Summer school ELA block intervention aligned to the FOSS science block

Study by Program	Research Design	Findings	Grade	Vocabulary Area	Source	Study Focus Area
August et al. (2018)	-Within- subjects design, extended versus embedded vocabulary instruction -20 min per day, 5 days per week, 8 weeks -186 ELLs in a transitional bilingual program	-Significant gains of 37% correct (t = 35.5, p < .0001, d = 2.3) found for words taught through robust instruction -Significant gain of 20% correct (t = 19.0, p < .0001, d = 1.3) for words taught through leaner instruction; and a nonsignificant gain of 2% correct (t = 1.78, p = .07, d = 0.1) for words learned through exposure	2nd	General academic lexemes	Peer reviewed journal	-Three methods of vocabulary instruction were explored for second grade Spanish-speaking ELLs--robust instruction, brief explanations, and exposure -Instruction within shared interactive reading -Study also explored interactions between instructional condition and word-type–(abstract cognates, abstract non-cognates, concrete cognates, and concrete non-cognates)

Study by Program	Research Design	Findings	Grade	Vocabulary Area	Source	Study Focus Area
Burns (2011)	-Between-subjects design, intervention condition versus business as usual -30-minute lesson, 5 days a week, 12 weeks -78 ELL students	-Group differences on a vocabulary measure that measured depth of word knowledge (*d*=.20) and standardized grade-level word meaning test (*d*= -.13) -Standardized reading comprehension measure garnered no effects (*d*=02).	1st	General academic vocabulary, scientifically based reading research	Doctoral dissertation, IES practice guide	Study explored small group intervention for struggling readers that focused on the components of scientifically based reading research, conducted as part of an MTSS framework

Study by Program	Research Design	Findings	Grade	Vocabulary Area	Source	Study Focus Area
Biemiller & Boote (2006)	-Between- subjects design, data not disaggregated by ELL status -15 minutes per day -400 students, 50% ELLs	-Authors found a gain of 28% in the brief explanation instruction condition -13% gain for the business as usual condition	3rd	General academic vocabulary in an ELA block	IES practice guide Peer-reviewed journal	Study investigated centered on embedded vocabulary instruction where the instruction was delivered at the core level of instruction included in situ explanations of word meanings encountered in shared text

Study by Program	Research Design and Sample	Findings	Grade	Vocabulary Area	Source	Study Focus Area
Carlo et al. (2004)	-Between-subjects design, Intervention versus business as usual -30-45 minutes per day, 15 weeks -254 students in 5th grade, 62% of the sample ELLs, data disaggregated	Gains on researcher developed word association task (*d*=.34) between the instruction and no-instruction condition	6th	General academic vocabulary in an ELA block	Peer-reviewed journal, IES practice guide, NLP	Intervention included way to infer word meanings from context, definitions, and association tasks, synonym/antonym tasks, semantic feature and morphological analysis
Chambers et al. (2006) **Chambers et al. (2004)**	-Cluster randomized design -90-minute reading lessons -450 first graders, 29% of ELL students in the treatment condition, data is disaggregated	-Group differences indicated on the following: word Identification subtest (*d*=0.15), the Word Attack subtest (*d*=0.32) -No statistical difference in results of subpopulations compared to overall sample	1st	General academic vocabulary in an ELA block	Peer-reviewed journal, IES practice guide	-Study investigated different multimedia support on reinforcing vocabulary within the Success for All -Program as a follow up study to Chambers et al. (2004)

Study by Program	Research Design and Sample	Findings	Grade	Vocabulary Area	Source	Study Focus Area
Crevecour et al. (2014)	-Between subjects, ELL and non-ELL, data dissagregated -20-25 minutes per day, 18 weeks -In the treatment condition: ELLs (n= 31); monolinguals (n = 49) -In the business-as-usual condition: ELLs, (n = 17); monolinguals (n = 25)	-Group differences indicated by statistically significant proportion of the variance in posttest on the Peabody Picture Vocabulary Test–III (PPVT–III) for treatment -No differences were found for language status	Kinder-garten	General academic	Peer-reviewed journal	-Tier II intervention drew on shared interactive reading that set the context for robust and leaner vocabulary instruction -Meanings of words were explained in the study along with engaging students in discussions with words, and connecting the words to visuals
Gunn et al. (2005)	-Between- subjects, Hispanic and non-Hispanic data disaggregated by ELL status -159 (53.2%) were Hispanic, and 140 (46.8%) were non-Hispanic.	Gains ($t = 12.56$, $p < .0001$) on a curriculum-based measure of vocabulary knowledge in the intervention condition	Kinder-garten-3rd	5 components of reading	Peer-reviewed journal, National Literacy Panel	Tier II reading intervention with *Reading Mastery* and *Correct-ive Reading* centered on scientifically-based reading research that coincided with daily reading instruction in their classroom

Study by Program	Research Design	Findings	Grade	Vocabulary Area	Source	Study Focus Area
Fillipini (2007) Fillipini et al. (2012)	-Within-subjects three group design, no instruction versus phonological awareness (PA) instruction only, PA plus vocabulary -20-25 minutes per day -71 Spanish-speaking English Learners (EL)	-Group differences indicated as outcomes on a researcher developed measure for vocabulary with moderate effects (*d*=.54), compared to the no instructed group -Effects were significant for both treatment groups on the phonological decoding metrics	1st	General academic vocabulary enhanced reading instruction	Peer-reviewed journal, doctoral dissertation	Tier II intervention where the vocabulary portion consisted of direct instruction of general academic words with word learning strategy instruction
Kieffer & Lesaux (2012)	-Between- subjects design; treatment condition versus a no-instructed condition -45 minutes, daily for 18 weeks -349 ELLs and 133 non-ELLs	Moderate effects on a morphological decomposition task for the treatment (*d*=.61)	6th	Word learning strategies	Peer-reviewed journal,	Intervention focused on academic language instruction and the use of word learning strategies, such as contextual analysis and morphology
Howard & Artegotia (2013)	-Between- subjects -1-hour, daily over 16 weeks -350 ELL students in middle school	Moderate effects for vocabulary in the treatment group compared to the control.	7th-8th	Word learning strategies	Peer-reviewed conference paper	Intervention, EVOCA, focused on morphology, context clues and cognate awareness

Study by Program	Research Design	Findings	Grade	Vocabulary Area	Source	Study Focus Area
Lawrence et al. (2012)	-Unmatched Quasi-experimental study -15 minutes per day -328 students, 10% ELL	Increased scores on the wave-to-wave vocabulary knowledge assessment ($\gamma10 = 0.371$, $p < .00$) in the treatment condition	Middle school	Word learning strategies, explicit instruction	Peer reviewed journal	Word Generation intervention at the middle school level focused on word learning strategies and explicit instruction
Lesaux et al. (2010)	-Between-subjects with intervention treatment versus business as usual -Intervention for 45 minutes per day for 20 weeks -Implemented with 476 ELLs	Group differences for the treatment where the post-intervention outcomes on a morphological decomposition task yielded an effect size of *d*=.22	6th	Word learning strategies academic vocabulary	Peer-reviewed journal, IES practice guide	Intervention focused on academic language instruction and the use of word learning strategies, such as contextual analysis and morphology
Lesaux et al. (2014)	-Between-subjects with intervention treatment versus business as usual -Intervention for 45 minutes per day for 20 weeks -Implemented with 1,365 ELLs	-Academic Word Mastery Test yielded an effect size of .69 -Word Association Test yielded an effect size of .27, and a context clues test garnered an effect size of .22 -Effect sizes for reading (*d*=.09 on a norm-referenced reading test)	6th	Word learning strategies academic vocabulary	Peer-reviewed journal, IES practice guide	Intervention focused on academic language instruction and the use of word learning strategies, such as contextual analysis and morphology

Study by Program	Research Design	Findings	Grade	Vocabulary Area	Source	Study Focus Area
Paez et al. (2011)	-Between- subjects intervention versus intervention plus home connection and business as usual -16-week program for four months, 20-minute segments per day -148 students (61 control condition, 43 classroom-only condition, and 44 home-plus-classroom condition)	The home-plus-classroom intervention had significantly bigger gains (moderate effects) compared to the control condition in the researcher-developed KLS receptive skills for target vocabulary words	Kinder-garten	General academic words, home connection	Peer-reviewed conference paper	The longitudinal intervention focused on vocabulary instruction in kindergarten connected with a home component that included parent meetings and workshops and home resources, such as Spanish books

Study by Program	Research Design	Findings	Grade	Vocabulary Domain	Source	Study Focus Area
Peercy et al. (2015) **Silverman et al. (2017)**	-Between- subjects design, ELL and non-ELL and treatment and business as usual -30 minutes per day -196 kindergarteners (106 in intervention and 90 in comparison) and 239 fourth graders (131 in intervention and 108 in comparison). 50% of the kindergarteners (48% in the intervention and 51% in the comparison group) and 20% of the fourth graders (18% in the intervention and 23% in the comparison group) were ELLs	-For Kindergartners, small effects on PPVT (g=.20), with no intervention by ELL status interaction. -Strong effects for the researcher-developed expressive and receptive measures (g=.86 and .77, respectively). -For 4^{th} grade, no effects on th Gates MacGinitie Vocabulary subtest, moderate effects on the researcher developed vocabulary measure (g=.56) and small effects for the researcher developed expressive vocabulary test (g=.27). -No differential effects for ELLs with respect to the receptive measure, but not for the expressive measure	K and 4^{th} grade	General academic and domain-specific vocabulary	Peer reviewed book chapter, peer review journal	-Intervention consisted of a cross-age peer tutoring multimedia program where participants watched videos or read digital texts -Students took part in vocabulary and reading strategy instruction as part of that cross-age peer tutoring program

Study by Program	Research Design	Findings	Grade	Vocabulary Domain	Source	Study Focus Area
Proctor et al. (2009) **Dalton & Grisham (2007)**	-Between-subjects, treatment and business as usual -310 students, 78% ELL	Students in the treatment group had a gain of 2% on a standardized vocabulary measures, compared to the control	4th - grade students in two class-rooms	English vocabulary and reading comprehension	Peer reviewed journal, IES practice guide	Intervention consisted of a digital, universal design platform with in situ parenthetical explanations and robust vocabulary instruction
Nelson et al. (2011)	-Between-subjects design, intervention versus business as usual -20-minute lessons, 5 days a week, 20 weeks -185 ELLs	-Group differences indicated on a root word researcher-developed measure that yielded an effect size of *d*=.68 -Standardized vocabulary measure garnered an effect size of *d*=.15 -Standardized reading measure yielded an effect size of *d*=.19	Kinder-garten	English vocabulary and beginning word reading	Peer-reviewed journal, IES practice guide	Small group intervention for struggling readers that focused on the components of scientifically-based reading research, study conducted as part of an MTSS framework

Study by Program	Research Design	Findings	Grade	Vocabulary Domain	Source	Study Focus Area
Silverman & Hines (2009)	-Between-subjects design, ELLs and non-ELLs -45-minute lessons, three days a week, for 12 weeks -Eighty-five children across the four grade levels participated in the study. Of these 85 children, 15 were in pre-kindergarten, 28 in kindergarten, 25 in first grade, and 17 in second grade	-Between subjects group differences (*d*=.97) for the multimedia condition for ELLs -No differential effects for the non-ELLs -Standardized vocabulary measure (PPVT) yielded an effect size gain of *d*=.55	PK & 1st grade	General academic vocabulary	Peer reviewed journal	-Intervention compared traditional and multimedia-vocabulary instruction in the context of read aloud activities -Participants watched videos with target vocabulary -In traditional condition, students read books with the target words

Study by Program	Research Design	Findings	Grade	Vocabulary Domain	Source	Study Focus Area
Taboada & Rutherford (2011)	-Between-subjects -Mixed methods, two conditions: contextual vocabulary instruction and intensive vocabulary instruction -8 weeks -20 students classified as ELL	Significant differences found with effects for explicit vocabulary instruction in the intensive vocabulary condition at eight weeks post intervention for both higher and lower readers	4th	General academic vocabulary, science vocabulary, explicit vocabulary instruction	Peer reviewed journal article IES practice guide	Explicit vocabulary instruction was one element explored in the Intensified Vocabulary Instruction intervention condition of the study
Townsend & Collins (2009)	-Within-subjects design -20-day after school program -52 ELL students	Effect sizes for the instructed versus words that were not instructed was moderate at $n^2 = .36$	7th and 8th	General academic vocabulary, explicit instruction	Peer reviewed journal	Afterschool intervention for extended vocabulary learning that consisted of daily discussion related to four target words and student engagement with vocabulary games

Study by Program	Research Design	Findings	Grade	Vocabulary Area	Source	Study Focus Area
Uchikoshi (2005)	-Between-subjects design -Both groups watched one 30-minute episode, three times per week, treatment group had additional scaffolds of captioned word meanings -108 ELLs	Growth in the treatment condition related to the sample who viewed additional TV programs in the home, but no differences between the no instruction and school viewing conditions	Kinder-garten	General academic vocabulary	Peer reviewed journal	Participants in the study watched educational television programs, *Arthur* or *Between the Lions* with embedded in situ explanations for vocabulary and captioned support for the treatment group
Uccelli & Rosenthall (2013) **Artzi et al. (2015)**	-Within- subjects design, incidental exposure -15 min per day, 4 days per week -186 ELLs in transitional bilingual program	Significantly moderate effects for the words that were instructed compared to the words that were not instructed	2nd	General academic connectives	Peer reviewed conference papers	-Study explored direct instruction on words of grammatical cohesion -Words were taught using picture cards and through kinesthetic activities

Study by Program	Research Design	Findings	Grade	Vocabulary Area	Source	Study Focus Area
Vadasy et al. (2010) **Vadasy et al. (2013)**	-Between- subjects design -20 minutes a day, five days for 20 weeks -93 treatment, 92 control; after attrition, 74 in treatment and 66 in control	-Students in the treatment scored higher than those in the no-instructed condition on reading and vocabulary (*d* = 0.23) and (*d* = 0.29) (curriculum aligned and standardized) as well as word reading (*d* = 0.35) -Treatment benefits maintained in delayed follow up administered one year post intervention, lessened magnitude for effects	Kindergarten	English vocabulary and beginning word reading	Peer-reviewed journal, IES practice guide	Study focused on a Tier II supplemental vocabulary intervention within MTSS

Study by Program	Research Design	Findings	Grade	Vocabulary Area	Source	Study Focus Area
Vadasy et al. (2015)	-Two-cohort cluster-randomized trial -20 minutes a day, five days for 20 weeks -324 EL students	-Intervention garnered significantly greater gains in reading vocabulary (*d*= .64) and decoding (*d*=45) -In a one-year follow up, students in treatment maintained differences but effect sizes were lessened: *d*=.29 for vocabulary and *d*=.27 for decoding	Kindergarten	English vocabulary and beginning word reading	Peer-reviewed journal	Tier II supplemental vocabulary intervention within MTSS

Study by Program	Research Design	Findings	Grade	Vocabulary Domain	Source	Study Focus Area
Vaughn et al. (2009)	-Between-subjects design, intervention versus business as usual -50-minute lesson, 5 days a week, 9-12 weeks -97 ELLs in study 1	Intervention yielded gains in the treatment condition compared to those in the no-instruction condition; (*d*=.57) on a researcher developed measure of vocabulary	7th	Social studies vocabulary and general academic words	Peer-reviewed journal, IES practice guide	Vocabulary as part of a Tier II supplemental intervention within MTSS
Vaughn et al. (2006)	-Between-subjects design, intervention versus business as usual -50-minute lessons, 5 days a week, for 35 weeks -91 ELLs	General measure of vocabulary administered before and after intervention (Woodcock Language Proficiency Battery-Revised), post intervention effect of *d*=.15 for treatment students	1st	Social studies vocabulary and general academic words	Peer-reviewed journal, IES practice Guide	Small group intervention for struggling readers focused on components of scientifically-based reading, study as part of an MTSS framework

Study by Program	Research Design	Findings	Grade	Vocabulary Domain	Source	Study Focus Area
Wijukamer et al. (2016)	Matched control group design	Results indicate that the treatment classroom children outperformed the matched sample on signaling word task	Grades 4-6	Content vocabulary	Peer reviewed conference paper	Program within a universally designed web-based platform used to support instruction in reading, including vocabulary

Some general themes come to light with the studies highlighted in Tables 1 and 2. First, compared to other research reviews (i.e., NRP, 2000) there is paucity in the vocabulary research base regarding ELL vocabulary instruction studies. Although there is at least one study per grade level referenced in Table 2, vocabulary instruction for kindergarten students is examined in the most studies, followed by upper elementary and early middle school and then first grade. There are few studies that investigated vocabulary instruction with second grade ELLs—a formative year for ELLs in a transitional bilingual program, as second grade is the year before they transition to English and the year before students tend to shift from learning how to read to using text to acquire content information.

As seen across the studies highlighted in Tables 1 and 2, the vocabulary instructional methods for non-ELLs outlined by the NRP (2000) and Graves (2006) were shown to be effective with ELLs. Those practices tended to yield effect sizes in the moderate range for proximal curriculum aligned measures of vocabulary and small to moderate effects for more distal standardized metrics of vocabulary. However, the studies in Table 2 depict additional instructional considerations that specifically account for the learning needs of students acquiring English as well as their learning attribute of bilingualism (also outlined in Baker et al., 2014; Gersten et al., 2007; Graves et al., 2013). Additional techniques shown in the Table 1 and 2 studies to be particularly supportive for ELL word learning include:

- Using vocabulary instructional approaches that include the use of non-linguistic support, such as: visual aids, demonstrations, realia, and multimedia

(August et al., 2016; Howard & Artegotia, 2013; Uccelli et al., 2013; Proctor et al., 2009; Silverman & Hines, 2009; Uchokoshi, 2005; Uccelli et al., 2013; Vaughn, et al., 2006; Wijukamer et al., 2016);

- Providing opportunities for reinforcement of vocabulary through all the modalities of language (reading, writing, speaking and listening) (August et al., 2015; August et al., 2009; August et al., 2014; Baker et al., 2014; Peercy et al., 2015; Paez, 2013; Taboada & Rutherford, 2011);
- Drawing on language-based scaffolds to support students' use of academic vocabulary, such as sentence starters and frames (August et al., 2016; August et al., 2018; Peercy et al., 2015; Paez, 2013; Uccelli et al., 2013);
- Leveraging the first language and cultural background as a learning resource, drawing on principles of culturally and linguistically responsive instruction and utilization of the first language (August et al., 2014; Carlo et al., 2004, Fillipinni, 2007; Proctor & Mo, 2009; Kieffer & Lesaux, 2012; Vaughn et al., 2006; Paez, 2013; Uccelli et al., 2013).

The following section details each of the essential components of the four-pronged approach to vocabulary instruction described in Figure 2.

Explicit Vocabulary Instruction

Vocabulary words are learned through explicit instruction. Robust, intensive direct instructional approaches for vocabulary, also called rich, explicit and extended approaches, provide the learner with a plethora of activities that lead to deep processing of a word.

Across elementary and intermediate grades, robust explicit vocabulary instruction should include: (1) multiple exposures to words over several days using the four modalities of language; (2) opportunities for student engagement with use of word meanings in reading, writing, speaking, and listening, (3) utilization of child-friendly definitions with contextual information; and (4) use of examples, non-examples and visual aids and multimedia associated with the word (recommendations put forth in Baker et al., 2014).

For example, Carlo et al. (2004) executed a program with elementary school-aged ELLs using robust instruction. In that intervention study, teachers' pre-taught vocabulary words by explaining their meaning. Students were encouraged to infer word meanings from the context of the reading passage. Students also took part in other word learning tasks such as: word association tasks, synonym/antonym determination, and semantic feature and morphological analyses.

In another study with lower elementary ELL learners, Silverman (2007) drew on three methods of vocabulary instruction where words in children's books were discussed using child-friendly definitions. Questions and prompts for the words were used to promote student discussion about the word meanings with their peers. Students also acted out the meanings of words and learned about examples and non-examples of the program words. Instruction, called anchored instruction, was used to instruct students on letter-sound correspondences with the target words and was meant to fostering students understanding of sublexical skills such as spelling patterns and sound relationships of the target vocabulary words.

In a later study with kindergarten and fourth grade ELLs and their non-ELL peers, Silverman et al. (2017), students in a peer-tutoring program read texts together that included high-utility academic words. The teacher explained the meanings of those words, showed pictures that illustrated the word, led students in discussions with the word, provided examples and nonexamples of the word, and prompted word reinforcement through the peer discussions about the text and unit theme.

Baker et al. (2014) recommend that teachers: robustly teach a set of academic vocabulary words across several days using a variety of activities drawing on multiple language modalities. According to Baker et al. this may include: showing visuals connected to the word, providing child-friendly meanings for the word, using morphology and context clues with the word, engaging students in conversation with the word, conducting word association tasks with the word, and acting out the word meaning. Baker et al. (2014) also recommend that practitioners: use interesting text to contextualize the intensive academic vocabulary instruction, select a small set of high-utility academic vocabulary for extended instruction, and include domain-specific words as well as general academic vocabulary that supports the domain-specific words in word selection for more robust instruction.

The use of nonlinguistic supports is a core element of robust vocabulary instruction and this technique has been used across the ELL studies highlighted in Table 2. These kinds of supports include acting out words, drawing activities with the word, and using pictures to visualize words. Additionally, multimedia has been shown to be an important element for ELLs because it links oral language with visual support. Using captioned TV programs centered on school-relevant content, Neuman and Koskinen (1992) found that ELLs in the middle elementary grades acquired vocabulary from those captioned TV programs significantly better than a group of students that read a text on the same topic with the same target words. Uchikoshi (2005) found similar effects for elementary learners. Silverman and Hines (2009) investigated the use of multimedia with students who were ELLs and non-ELLs. Their results indicate that there were positive effects for the ELL group.

With variation across studies in Tables 1 and 2, techniques for first language support within robust and explicit vocabulary instruction draw on students' first language capabilities and include: defining target words in English and the students' first language, promoting peer discussions in students' home language with the target vocabulary, pointing out whether the word is a cognate in the students first language and English, and fostering students cognate awareness knowledge so they can figure out the meanings of unknown English cognates through contrastive analysis activities (August, Artzi, & Barr, 2016; August, Artzi, Barr, & Francis, 2018; Carlo et al., 2004).

Baker et al. (2014) acknowledge that an important limitation to extended robust vocabulary instruction is the number of words that can be realistically targeted. Vocabulary intervention studies in Tables 1 and 2 have largely followed a figure of roughly 8-15 words per week. For example, Carlo et al. (2004) targeted 10-15 words per week through explicit instruction in classrooms comprised of ELLs. Taboada and Rutherford (2011) taught 13 words per weekly unit. Beck and McKeown (2007) targeted 8-10 words per week in their project with non-ELLs and ELLs. Based upon these figures, only about 10-15% of the estimated 3,000 words children need to acquire each year to keep up with grade level text can likely be targeted through robust means. Therefore, the main limitation to robust vocabulary instruction is the limited number of words that can be targeted at that intensity level.

Other research on explicit vocabulary instruction investigated classroom-based activities through what Graves et al. (2014) refer to as "brief explanations," where words are explained through parenthetical student-friendly meanings (p.340). In the research literature, this method is also referred to as embedded approach to vocabulary instruction (Coyne, McCoach, & Kapp, 2007; Penno, Wilkinson, & Moore, 2002). Just clarifying the meaning of a word through embedded in situ methods with brief explanations has garnered effects in the research. For example, Biemiller and Boote (2006) found that elementary-aged learners, ELLs and non-ELLs, learned nearly twice the number of words targeted through brief in situ explanations compared to words that students encountered in the texts without instruction. Penno, Wilkinson, and Moore (2002) found similar effects with non-ELLs.

Exposure Only

Words are acquired through incidental exposure. For vocabulary instruction that is balanced, indirect learning methods, such as providing plentiful opportunities for exposure to school-relevant vocabulary, are used in tandem with direct methods of explicit instruction described above (see August et al., 2014; Coyne et al., 2007). Given the vocabulary acquisition demands that are required to keep up with grade-level text in the U.S., incidental exposure is a critical component in a vocabulary program since the majority of words, roughly 90% of the words students need each year, will be acquired through this method.

A key principle to promoting students' incidental exposure to vocabulary is that students are immersed in a language-rich environment because words that are acquired through incidental exposure are not explicitly highlighted or directly instructed (Krashen, 2013). Unlike the explicit vocabulary instructional methods (whether extended or embedded), where there is attention to the meanings of specific words, the focus in incidental exposure is on the exposure to a language-rich environment and on the meaning of and attention to the text and discourse in that environment (Graves, 2006; Graves, August, & Mancilla-Martinez, 2012).

A number of studies have investigated word learning from exposure within shared book reading. Some researchers claim that children's story books provide a good platform for vocabulary acquisition from exposure because research finds that there are ample academic, school-relevant words in books that may not necessarily be present in everyday conversations or through television (see Graves, August, & Mancilla-Martinez, 2013; NASEM, 2017). Book floods, where a large assortment of books is brought into the child's environment (with encouragement to engage with the books in their own way), and silent sustained reading (SSR), where children engage in free reading, are examples of these indirect methods in literacy instruction (Krashen, 2013).

Incidental exposure to new vocabulary with the use of multimedia activities is another element to vocabulary instruction reported through the research. Neuman and

Koskinen (1992) reported on a study where ELLs in middle grades watched captioned multimedia. The experimental group outperformed a comparison business-as-usual group of students on measures of word recognition and other receptive measures of vocabulary. Findings from this study also indicate that ELLs who started out with more English proficiency learned more than students with less English proficiency. Silverman et al. (2017) used multimedia digital text and video and peer-mediated learning models to enhance students cognitive processing of the words. In that study the video condition garnered the most effects.

For ELLs, strengthening their first language helps support English vocabulary acquisition. Paéz (2013), for example, conducted a study where half of the ELL participants heard the stories encountered in their school classrooms at home in their home language, which was Spanish in that study. The other half of students heard stories read at home in English. Findings from that study reveal that students who had heard stories read at home in their home language acquired significantly more words than those who heard stories read at home in English--although positive pre- to post-test (wave 1 to wave 2) effects were found for both groups. These results demonstrate that continuing to promote the first language for an ELL student has positive effects on English acquisition and providing access to bilingual books or books in students' home language benefits positive learning outcomes for the student.

To help support ELL students' acquisition of vocabulary through incidental exposure, the studies in Tables 1 and 2 used practices that ensured that instructional input was understandable, drew on first language support, and engaged students in activities to promote oral language proficiency (August & Shanahan, 2006; Baker et al., 2014; Gersten et al., 2007). The studies in Tables 1 and 2 also supported and fostered peer-mediated discussions where students who share a common home language worked together in peer-to-peer learning situations (Peercy et al., 2015; Silverman et al., 2017).

Word Learning Strategy Instruction

Since 85-90% of the words needed for acquisition will be independently acquired by the child, instructing students on word learning strategies is vitally important in that it helps augment the exposure to a language-rich environment that sets the context for incidental vocabulary exposure. Research finds that instruction in context clues and morphology as well as the use of reference tools, such as dictionary skills, can have significant effects in elementary aged children (Bauman, Edwards, Boland, Olejink, & Kamèenui, 2003; Lesaux et al., 2010). For a review of the instructional strategy related to context clues and morphology see Graves (2006) and Graves, August, and Mancilla-Martinez (2013)[6].

Facilitating ELL students' cognate awareness has shown to be effective (Carlo et al., 2004; Dressler, 2000; De'Angelo, Hipfner-Boucher, & Chen, 2017). Dressler (2000) investigated cognate awareness instruction with a sample of fifth-grade Spanish-speaking ELLs. The students who had been shown how to determine a word as a cognate were better able to determine the meanings of cognates that were not taught compared to a group that did not receive that cognate awareness instruction.

[6] The word learning strategies of morphology, context clues, and dictionary skills is not investigated in the study and will not be focused on in this review.

Dressler also found that cognates with more orthographic and phonological overlap were more likely to be inferred by the students and that students with a greater degree of oral language proficiency in their first language were more successful at transferring the cognate knowledge. De'Angelo et al. (2017) investigated cognate awareness in French/English bilinguals and found that cognate awareness was a contributor to reading comprehension in the second grade. Carlo et al. (2004) implemented a vocabulary program with Spanish-English bilinguals and cognate awareness instruction was included in the larger program. They found significant effects for word learning. The word learning strategy of cognate awareness was included in the AVE vocabulary program reported on in this study.

Word Reinforcement in Vocabulary Instruction

Research is mixed on whether the more a word is encountered, the better one will acquire it. August, Artzi, and Barr (2016) investigated robust instruction with third and fourth grade English learners. Outcomes were explored by domain-specific words specific to the study of science and general academic words that are characterized by high cross-subject dispersion. Authors also examined the role of reinforcement across both conditions and did not find effects for the condition that included an additional reinforcement game at the end of the program.

In Elley's (1989) study, the students who just listened to the story three times with embedded instructional support gained significantly more words, compared to students who read the book one time with the embedded instructional support. Penno, Wilkinson, and Moore (2002) found similar effects to Elley's study.

McKeown, Beck, Omanson, and Perfetti (1983) studied a program with elementary aged children in the third through fifth grade. That program incorporated two frequency conditions: 26 and 40 encounters in a condition called many words and between 10 and 18 encounters in a condition called some words. The authors found significant growth in word knowledge and lexical access for both conditions, though the words with the additional encounters in the many words condition presented with the greatest amount of growth. In a subsequent study, (McKeown, Beck, Omanson, & Pople, 1985) the researchers explored whether students learned words better with four or eighteen encounters. In the study, more encounters yielded better results and more complete knowledge of words. However, for the semantic decision task, in the extended instructed condition, encounters did not yield greater results.

These research studies indicate that words that are instructed through robust methods and are encountered more than ten times in a curriculum are more likely to be learned compared to words that are not encountered as frequently or instructed through robust methods. However, research has not yet investigated encounters in relation to type of word (stratified by cognate and abstractness status) or student-level factors of components of English and Spanish proficiency, which is where this study is situated.

Chapter 3 describes the AVE vocabulary program in more detail. It was developed using the four prongs to vocabulary instruction described by Graves (2006) and Graves et al. (2013) and detailed in this chapter. In the AVE program, words were explicitly taught through robust and extended vocabulary instructional methods and some word learning strategies (dictionary skills and cognate awareness) were also explicitly taught.

To promote a language-rich environment, the AVE program used children's books within the CCSS-ascribed stretch text lexile band at the target grade level, as the context for the vocabulary instruction. The program also included the use of multimedia by the way of animated PowerPoint shows, utilized visual environmental enhancements such as a visual word wall that included pictures cards with the word in Spanish and English and a visual image of the word. Techniques to promote academic talk among the peer learners and promote peer-mediated learning were also employed in the AVE vocabulary program.

Theoretical Underpinnings

This study is situated in perspectives from two overarching vocabulary acquisition and learning theories. The *Interactive Word Learning Burden* helps to explain why it is important to consider word-level attributes in vocabulary selection for instruction. Nagy and Hiebert's *Theory towards Word Selection* positions the way in which word-level attributes should factor into word selection for vocabulary instruction. Each of these theories is explained in more detail below.

Interactive Word Learning Burden

Nation (2001) describes a "learning burden," or learning load, of a word as "the amount of effort required to learn it" (p. 20). "Different words have different learning burdens for learners with different language backgrounds," Nation points out (p. 23). The learning burden Nation describes is a theoretical means to explain learner factors in vocabulary acquisition and learning, stipulating that a learner's background impacts the learning load of a word. For second language learners, language background is one important component in their overall background. According to Nation (2001):

> The general principle of learning burden is that the more a word represents patterns and knowledge that learners are already familiar with, the lighter the learning burden. These patterns can be from the first language, from knowledge of other languages, or from previous knowledge of the second language… for learners whose first language is closely related to the second language, the learning burden will be light. (p. 24)

While Nation addresses the learning burden related to person-level factors, as in language background and relatedness between the first and second language, the notion of learning burden can also be extended to word-level attributes through Nagy and Hiebert's (2011) recent theory toward word selection.

Nagy and Heibert published a chapter in the *Handbook of Reading Research* (Kamil et al., 2011) in which they present a theory regarding word selection. According to Nagy and Hiebert, word selection should consist of three elements: 1) identify the attributes of words when selecting words for instruction, 2) use multiple criteria and determine how word attributes work together to inform word selection, and 3) use that information to inform instruction. Nagy and Hiebert's chapter addresses the first two and the work of this proposed study provides a starting point for unpacking the third component of their word selection theory.

Nagy and Heibert (2011) discuss word-level properties in their word selection theory. They describe form properties and semantic properties that comprise the conceptual difficulty of a word. Form properties include components such as word length, number of letters, and degree of affixation. These properties may be more associated with visual word recognition, or our ability to decode a word form from text and then associate that decoded word form with a meaning. Semantic word properties entail elements of a word that impact how one understands the meaning of that word. Semantic properties discussed in the literature often include whether a word is concrete or abstract, able to be mentally imaged, and the level a word is related to other words. Polysemy, or the number of senses associated with a word, is another semantic attribute. For second language learners, cognate status, or the degree in which a word in the second language overlaps with a word in the first language via orthographic and phonological elements, is another component of semantic word attributes.

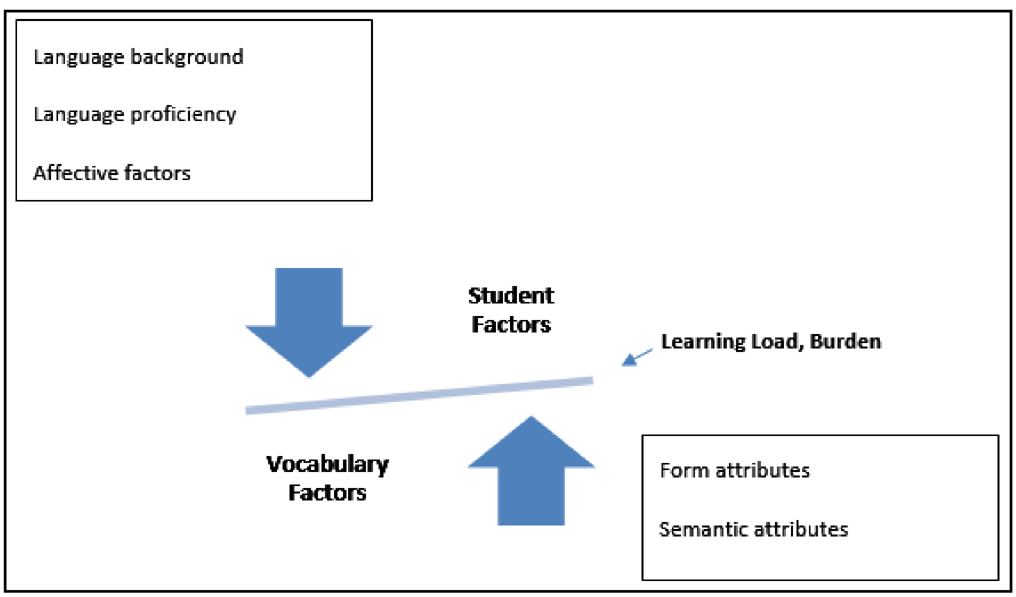


Figure 3. Interactive word learning burden.

For some word-level attributes, the learning burden of the word may interact with the learner's background. For example, the learning load for semantically, orthographically and phonologically related words across two languages (i.e., cognates) will shift based on whether a learner knows another language in their background and the degree of his or her vocabulary proficiency in that other language.

In combining the word-level considerations from Nation's notion of *Word Learning Burden* with the person-level considerations drawing on Nagy and Hiebert's (2011) notion of *Vocabulary Conceptual Difficulty* into what I call an *Interactive Learning Burden of Vocabulary* (see Figure 3). The *Interactive Learning Burden of Vocabulary* is a means to theoretically illustrate how a word's learning burden interacts with person-level factors, such as language proficiency as well as word-level factors, such as abstractness and conceptual complexity. Based on the work on cognate transfer knowledge (i.e, Genesee et al., 2006), educators may be able to strategically leverage a learner's backgrounds and learning funds, such as another language, to accelerate word learning in English. Therefore, researchers may be able to more precisely pinpoint the words that should be targeted through more robust vocabulary instruction.

Word Selection

Nagy and Hiebert (2011) propose selecting words for school vocabulary instruction by determining the word's attributes while assessing the importance of the word for the academic learning of the child. Teachers should determine the word's role in the lesson (i.e., usefulness of the word in text and for acquisition of the content), the type of lexical relationship to other words (i.e., morphological and semantic relatedness), a word's function with respect to language (i.e., frequency and dispersion across content areas). Delineating word attributes that interact together to impact word acquisition and learning encompasses an additional prong of the word selection theory and the *Interactive Learning Burden Theory* rests here.

Also, Nagy and Hiebert point to the *frequency familiarity disparity* as one element to inform judicious word selection for school-based vocabulary instruction. In the frequency familiarity disparity, according to Nagy and Hiebert (2011), words that appear frequently in academic text but are unknown or likely misunderstood by students may be good targets for more extended explicit vocabulary instruction. Since estimates posit that adequate reading comprehension depends on understanding the meaning of 90-95% of words in text (Nagy & Scott, 2000; Nation, 2001; Lauffer, 2001), it stands to reason that knowledge of words that appear frequently in text is an important factor in students' ability to comprehend school text, especially for ELLs (see Graves, August, & Mancilla-Martinez, 2014), as purported by Nagy and Hiebert.

Thus, for general academic vocabulary, word frequency is a critical factor in strategic word selection of general academic words for ELL vocabulary instruction. According to Hiebert (2005), five thousand high-frequency words account for 80% of the words in text. The utility of frequent words is what Brysbaert, Warriner, and Kuperman (2014) call the *word frequency effect*. This makes the case that investigating word frequencies might be useful and word frequency tools like the *Educator's Word Frequency Guide* (WFG, Zeno, Ivens, Millard, & Duvvuri, 1995) that principally provide data on the frequency of words in U.S. school texts by content area and grade level, might be important tools for instructors to utilize. The WFG was used to draw the words targeted for instruction in the AVE program.

Word Attributes to Inform Word Selection

While frequency is an important consideration in general academic word selection, Nagy and Hiebert (2011) also suggest that word-level attributes should also be taken into account so the high-utility vocabulary requiring more intensive explicit instruction is chosen. According to Nagy and Hiebert (2011) "some words are harder to learn than other words…sometimes the problem lies in the form and sometimes in semantics" (p. 220). Judicious word selection should inform strategic vocabulary instruction that consists of domain-specific as well as general academic terms. The following sections describe each word-level component in more detail.

Form Properties. Form properties regard the word form itself and may influence the speed of word recognition when one reads. Usually the number of letters in a word serves as an index of a word form property. With lexical decision tasks, Hudson and Bergman (1985) found a greater degree of association between word length and performance on the lexical response task than for word frequency. Bradley (1992) found similar effects with foreign language learners. However, Nagy, Anderson, and Herman (1987) did not find effects for word length; rather, there were significant effects for semantic properties. August and Barr (2011) found that the number of letters in a word predicted performances of ELLs and non-ELLs on a receptive measure of vocabulary in grades three and four.

Semantic Properties. Semantic properties are those concerned with the word's meaning. They are important with respect to how one retrieves a word meaning from their cognition in order to aid their reading comprehension and how that person uses their cognitive processes to understand a new word or phrase in academic text or discourse.

Cognate Status. Cognate status is a semantic attribute that interacts with a learner's background. For some learners only (bilinguals), cognate status is relevant, while for others (monolinguals), cognate status is not a relevant factor. This conjecture relates back to the interactive facet of the *Interactive Word Learning Burden Theory*.

Depending on the first language and the degree of orthographic and phonological relatedness, cognate status can shift for each learner. Cognates are words from two different languages that share a common historical origin. Many will therefore be orthographically and/or phonologically similar across the two languages, but not necessarily. For example, some words are historical cognates like the Spanish word *pesos* and English word *fish*, which do not exhibit a great degree of orthographic or phonological overlap. For instructional purposes, cognates like *pesos* and *fish* should not be considered cognates and the orthographic and phonological relatedness of the cognates should be taken into account, which this work did when coding for cognate status.

Research indicates that second language learners are better able to acquire cognates if their home language shares the cognate with the target language and that a higher degree of orthographic and phonological overlap between the two words helps promote the acquisition of the cognates (de Groot & Keijzer, 2000; Genesee et al., 2006; Tonzar, Lotto, & Job, 2009). Recent research with Spanish-English ELL students shows a strong association between English reading comprehension outcomes and Spanish cognate knowledge (Proctor, Carlo, August, & Snow, 2005; Proctor & Mo, 2009) and French cognate knowledge (White & Horst, 2012). In those studies, the Spanish-English or French-English bilinguals that presented with higher scores on English reading comprehension assessments also presented with differentially higher levels of Spanish or French cognate knowledge.

However, school-aged children who are bilingual learners may not always draw on their cognate awareness knowledge (Nagy, Garciá, Durgunogĺu, & Hancin-Bhatt, 1993). To that end, research investigating best methods for vocabulary instruction for ELLs demonstrates that ELL learners benefit from cognate awareness instruction because cognate awareness instruction keys students to noticing those words shared across both languages in school text and helps them make better use of their cognate knowledge (August et al., 2009; 2014; White & Horst, 2012). For example, Carlo et al. (2004) executed cognate awareness training as one part of a literacy program and found significant post-intervention effects for cognate recognition. White and Horst (2012) implemented cognate awareness instruction with French-English bilinguals and found a treatment advantage on a cognate awareness measure.

Abstractness (Conceptual Complexity) Status. A word's conceptual complexity is often conceptualized in the research to include degrees of a word's concreteness, imageability, and relatedness (as discussed in Nagy & Heibert, 2010); although some prior work has suggested that these components can be reduced to a single factor (Schwanenflugel, 1991) and are referred to in some cases as the abstractness factor.

Brysbaert et al. (2014), for example, envision a word's abstractness as a language-based element; whereas concreteness is situated within one's direct experience. While we can see, touch, and feel a concrete word such as *pencil*, physically feel *furry*, or smell *putrid*, the word *efficiently* does not have a physical representation for the word that we can sense; rather, it requires language to convey the word's meaning.

Recently, Brysbeart, Warriner, and Kuperman (2014) obtained concreteness ratings on a Likert Scale of 1-5 for over 37,000 words appearing in SUBTLEXus with at least 25 observations per word. In this database, the word *sled* appears at the concrete end of the scale and the word *irrespectively* at the abstract end.

Prior research has demonstrated that less concrete words, those less perceived through senses or direct actions, present with a greater learning burden in that they are less likely to be known and more difficult to acquire and learn in general (Schwanenflugel, 1991) and by second language learners (Crossley, Salsbury, & McNamara, 2011; Ellis & Beaton, 1993). Compared to concrete words, the more abstract words were acquired or learned later and less likely to be retained. This work suggests that words are easier to acquire when learners connect physically with words through senses.

Prior research also shows that concrete words are more able to kindle a mental image of the word, referred to as imageability (Paivo, 2013). For example, a word such as *pencil* is highly imageable, as we can quickly conceive a mental image of the object it describes; however, the word *theory* is not considered to be very imageable because it does not easily promote a mental image and requires language to sufficiently convey the true sense of the word. Bird, Franklin and Howard (2001) studied imageability as a factor in age-of-acquisition ratings obtained through self-report. In their study, participants self-reported the age at which they acquired a word. Findings reveal that study participants reported that the less-imageable words were more likely to be acquired at later ages.

This investigation is focused on the investigation of instructional outcomes related to semantic word attributes of cognate status and conceptual complexity, defined as whether the word is abstract or concrete. Prior research indicates that those are among the most salient semantic factors that impact the learning burden of words in elementary-aged ELLs (August & Barr, 2011).

Significance of the Study

This study explores interactions between level of instructional reinforcement (encounters with the word) and type of word—abstract cognates, abstract noncognates, concrete cognates, and concrete noncognates. This study also considers generalized English language vocabulary and Spanish language proficiency of each student participant. There is a paucity of experimental research focused on ELL vocabulary instruction in the second grade, and though there is work that compares conditions that differ on levels of reinforcement with respect to teacher-directed instruction in non-ELL studies (e.g., Elley, 1989; McKeown, et al., 1985; Penno, et al., 2002), there is no work that explores the interactions between reinforcement and word-level properties stratified by the dimensions of cognate status and conceptual complexity. Further, there are limited studies that consider student-factors related to both their first and second language as additional possible influencers in the word learning of each reinforcement condition.

This study is a contributor to the field in that the premise of the work takes into account the theoretical notion that words display different learning demands and that the learning demand might shift based on the learner factors, such as language background and language proficiency. This shifting learning burden can be taken into account and inform strategic vocabulary instruction of various intensity levels for ELLs. There was no research uncovered in this review that has studied the effects of word learning by level of reinforcement as well as word and student factors, such as words' conceptual complexity and student language proficiency in their first and second. Classroom instructional time is a limited commodity so teaching time related to vocabulary must be as efficient as possible. Accordingly, it is important to determine how much reinforcement is needed for an ELL student to aptly learn a word and whether components of word-type and Spanish and English language proficiency interacts with word learning in those conditions.

Chapter 3: Methods

Introduction

This chapter describes the present study, AVE 3B, a within-subjects substudy of a larger principal study called the Acquisition of Vocabulary in English (AVE). As described in the literature review (Chapter 2), both research and theory support the importance of exploring word and student factors on ELL word learning with varying levels of instructional reinforcement. This chapter details the research design, data analyses, research methods, and analytic approach to address this dissertation's research questions.

Study Overview

The study reported in this dissertation is a within-subjects subset study (substudy) in the third year of a larger study, called the AVE project. AVE was an intervention project designed to investigate breadth of vocabulary learning with lexicalized words as well as words of grammatical cohesion when teachers used evidence-based practices for vocabulary instruction in the context of shared reading. The 3B substudy investigates varying levels of instructional reinforcement of the lexicalized words in the treatment condition only. AVE3B research using the AVE program explores the relationship between instruction, with varying levels of instructional reinforcement (or encounters) and student achievement on measures of vocabulary knowledge among Spanish-English ELL children in second grade who were participating in the AVE intervention as part of their ELD instruction within a district-wide transitional bilingual program.

AVE Project Design

The principal AVE project took place over three years. The first year was not an intervention project, but rather a year in which the research team assessed a large sample of ELLs in the Mid-Atlantic and Southwestern regions related to various components of literacy. This precursor work informed the subsequent intervention years. Therefore, in Year 1 the research was designed to determine what types of academically-salient words students, including ELLs, struggle with the most. Findings from that year informed the development of a subsequent intervention from those words. The intervention was developed for instruction of lexicalized words as well as words of grammatical cohesion.

Year 2 of the principal AVE project was the first intervention year and conducted with a sample of ELL students across ten classrooms in the Southwestern region (reported in August et al., 2018). The research team applied findings from the assessment outcomes of Year 1 to determine word selection (explained in the intervention section of this chapter and in August et al., 2018) and conducted a within-subjects research study to investigate how extended methods of vocabulary instruction versus embedded methods (brief parenthetical explanations during text readings) of vocabulary instruction interact with type of word (defined by conceptual complexity/abstractness status and cognate status). The research team also explored instruction of words of grammatical cohesion as part of the program.

The final year of the project was designed as the culminating study and entailed a between-subjects study examining AVE program outcomes in a treatment and business-as-usual design.

As part of that culminating study in Year 3, there was an incorporation of a substudy, a within-subjects component in the treatment classrooms that would allow for the investigation of varying levels of word reinforcement connected to type of word (stratified by dimensions of cognate status and conceptual difficulty/abstractness status). As such, the project investigators and research organization where the study took place agreed to my use of the deidentified assessment data for this dissertation work. Accordingly, this dissertation investigates the within-subjects substudy, called AVE 3B (third year of the project and "B" denotes the sub-study), of the principal main-effects intervention (AVE3) that took place in that final year of the three-year AVE project.

In AVE 3B, robust vocabulary instruction was implemented during shared interactive reading through two instructional conditions. Both conditions included some initial extended/robust vocabulary instruction, which I refer to in this dissertation as *base instruction*. One condition consisted of base instruction plus six encounters with the word through reinforcement activities (called the reinforcement condition) another condition consisted of base instruction plus fourteen encounters with the word through reinforcement activities (called reinforcement plus condition). Other words were embedded in the curriculum and assessed but not instructed. This condition is referred to as the incidental exposure condition. The incidentally exposed words, therefore, serve as a within-subjects proxy for a business-as-usual, no-instruction condition.

Across all three instructional conditions (reinforcement, reinforcement plus, and exposure), words were stratified by cognate status and level of abstractness (also referred to as conceptual complexity). Stratifying the pool of targeted intervention words by those two dimensions resulted in four types of words: abstract cognates, concrete cognates, abstract non-cognates, and concrete non-cognates. Although the larger AVE study utilized a between-subjects design, the substudy (AVE 3B) drew on a within-subjects methodology. In that AVE3B substudy design the four word-types were assigned to each of the three instructional conditions using methods to equate the words across the three conditions and control for time confounds.

Accordingly, words in each condition were theoretically equated using variables such as the type of word (abstractness and cognate status), grade level knownness and percent correct from the *Living Word Vocabulary*[7] (Dale & O'Rouke, 1981), frequency of words in school text according to the WFG (Zeno et al, 1995), and part of speech (POS). The methods used for equating the words across conditions and then implementing a pool of equally stratified words for each instructional condition allows for the investigation of vocabulary-level outcomes related to the types of word by instructional condition and by the student's English and Spanish language proficiency in a concurrent within-subjects design.

Student Sample

Two hundred and twenty-eight students across twelve classrooms participated in the AVE 3B substudy. All student participants were second grade Spanish-speaking ELLs attending schools in a large urban district in the Southwestern region

[7] The Living Word Vocabulary word knownness data details the percentage of students at grades 4,6,8,10,and 12 who knew the meaning of a word. See Dale and O'rouke (1981) for more details related to determining word knownness in the LWV.

of the United States. All students participating in the study were receiving school instruction through a transitional bilingual program model. Thus, the daily school instruction was conducted in Spanish, except for the 90-minute English language development (ELD) block (the block where the teachers implemented the AVE program). Accordingly, all participants in this study were transitioning to English instruction the following year.

Twelve teachers participated in the AVE3B substudy and carried out the AVE program instruction. All participating teachers instructed in a transitional bilingual program and spoke Spanish as their first language. In the study schools, over 75% of the students received free and reduced lunch.

Although all students received the AVE program instruction in their ELD blocks, only data from students who participated in the intervention and returned guardian consent forms indicating guardian consent are included in data analyses reported in Chapter 4. To that end, over 90% of the total students who received the AVE program turned in confirmatory consent forms.

Because the study took place in a Spanish-English transitional bilingual program, all the students spoke Spanish as their home language and were classified as ELLs by the district. The AVE intervention was delivered during the ELD block in core instruction, meaning that the AVE vocabulary instruction program was given to all students, regardless of the students who were receiving intervention or special education services.

AVE Intervention

The AVE instructional program (used in the AVE3B study) consisted of multidimensional vocabulary instruction, drawing on a range of evidence-based vocabulary instructional procedures that used interactional techniques related to the four prongs of vocabulary instruction discussed in Chapter 2 (see Graves, 2006; Graves et al., 2013). The AVE program's instruction occurred within the context of shared interactive reading. The shared reading used children's literature lexiled within the CCSS-ascribed stretch bands and were also aligned with district content standards[8].

Teachers who were delivering the AVE program instruction took part in one professional development training session prior to the AVE program implementation. They also received biweekly job-embedded coaching from experienced instructional technical assistance (TA) specialists classified as district-level master teachers. Therefore, the TA providers possessed wide-ranging knowledge in coaching teachers in transitional bilingual program classrooms.

Intervention Vocabulary. Because each year of the principal AVE studies built on top of one another, the same methods for word selection in the AVE3B study were used in Year 2 of the AVE project. These methods are reported in August et al. (2018) and are also overviewed below.

Target vocabulary in the AVE program comprised 144 words stratified by cognate status and abstractness status (also referred to as conceptual complexity)

---

[8] A lexile is a quantitative index of text complexity that includes the sentence length as well as word frequency in the text complexity formula. Lexiles are widely used as metrics for determining text difficulty by grade level.

resulting in four 12-word pools for each instruction condition-- 36 words of each type (abstract cognates, abstract non-cognates, concrete cognates, and concrete non-cognates). Appendix B shows the AVE program words by unit stratified by word-type and instructional condition. The AVE intervention words were randomly selected from the larger AVE Database described below.

AVE Word Database. Target words in the AVE program were randomly drawn from the AVE Database, developed in Year 1 of the project and also reported in August et al. (2018).

Development of the AVE Database started with the Educator's Word Frequency Guide (WFG) database (Zeno et al., 1996). The words in the WFG that appeared with a high frequency in elementary school texts (U value of 10 or more in the WFG) comprised the first component of the AVE Database. Theoretically, this pool of words is high-utility in that the words appear frequently in elementary school texts in the United States, according to WFG indices. In this first step, 522 total word forms were pulled from the WFG into the AVE Words Database. At this stage, the words were only word forms with no associated meanings.

The 522 word forms from the WFG pulled into the AVE Database were assigned grade-level meanings based on grade-level definitions in the *Living Word Vocabulary* (LWV) (Dale & O'Rouke, 1981). The LWV is a database of word knowness indicators, meaning the LWV displays the percentage of assessed students who knew the word by the following grades: 4, 6,8,10, and 12 (see Dale & O'Rouke, 1981 for LWV development procedures). All applicable meanings from the LWV were used, ranging in one to five entries per word—five being the most polysemous word in the LWV database.

Matching LWV meanings to WFG word forms resulted in 1335 cases for the 522 word forms, accounting for entries for polysemous word meanings. All 1335 cases in the AVE Database were then coded for cognate and abstractness/conceptual complexity status at the meaning level.

To code for cognate status, two Spanish-English bilinguals with advanced degrees in a linguistics-related field each coded half of the words in the AVE Database. The words were coded at the word meaning level because some meanings of polysemous English word forms may share cognate status with Spanish while other meanings may not. Accordingly, the coders initially scored each entry for cognate status using three indicators: *cognate, noncognate*, or *false cognate*. After initial training regarding the coding rubric and procedures, the coders calibrated by jointly coded 200 words together and met after the first round of coding to discuss their codes. A third coder, also a Spanish-English bilingual with an advanced degree in a linguistics-related field, refereed when coding difference arose. Comprehensive coding procedures and the full cognate coding rubric is reported in August et al. (2018).

For abstractness status (also referred to as conceptual complexity status), two coders with advanced degrees in a linguistics-related field each coded the AVE Word Database using criteria informed by Nagy et al. (2001). Accordingly, the coders rated level of abstractness on a five-point Likert scale with 1 being the most concrete and 5 the most abstract. After initial training regarding the coding rubric and procedures, the coders calibrated by jointly coded 200 word meanings together to determine congruence. Between the two coders, inter-rater reliability using Kendall's coefficient of stability was .75. Detailed coding procedures and the full conceptual complexity coding rubric is reported in August et al. (2018).

Word Selection for the Intervention. Words in the AVE Database (with codes for cognate status and abstractness status) were then dichotomously grouped by cognate status and abstractness status. For cognate status, words that were coded as cognates were placed in the cognate pool of words and words that were coded as non-cognate were placed in the non-cognate pool. For abstractness status, words receiving an abstractness code of 1- 2.5 were placed in the concrete category and words coded as 3.5-5 were put in the abstract category. Words with a combined concreteness rating of three were omitted since those words were directly in the middle of abstractness and concreteness and did not clearly fall into any one of those categories (see Appendix B for the intervention words by condition stratified by word-type).

Words were randomly selected in a stratified fashion to allow for the selection of 12 of each word-type for the four pools of words: concrete cognate, abstract cognate, concrete noncognate, abstract noncognate. The selected words were then imported into the books chosen for the program. Importing is a strategy where words are placed into books that replace other semantically congruent words (Graves et al., 2014). Because of importation, 10% of words were exchanged due to semantic incongruence with the selected book. When incongruence arose, words were randomly sampled a second time until a word fit with the book and condition.

Attention was used to control for time effects by type of word, so that each word-type was equally dispersed across the entire curriculum. Given the selection strategy of words, the type of word (concrete cognate, abstract cognate, concrete noncognate, and abstract noncognate) was controlled in each condition, so all word-types were taught an equal number of times, in each of their respective reinforcement conditions, for every two units. This procedure mitigates confounds related to time effects.

Conditions in the AVE 3B study. Once words were randomly selected in a random stratified fashion from the AVE Word Database and grouped into the three conditions (exposed, reinforcement, and reinforcement plus) the word pool of each condition was randomly assigned to one of the three conditions: reinforcement, reinforcement plus, or incidental exposure. Therefore, within word-type, 12 of the 36 words were randomly assigned to one of the three conditions: reinforcement, reinforcement plus, and a within-subjects "proxy" for a business-as-usual condition (the exposed condition). In the exposed condition, the words were embedded in the curriculum with no instruction or reinforcement with that set of words.

Fidelity of implementation of the AVE program was also collected twice over the course of the program to assess instructional adherence to the program as well as instructional quality using a protocol developed the prior year and reported in August et al. (2018). To assess the effects of the intervention, students were given curriculum-aligned and standardized assessments before and after the implementation of the AVE program.

AVE3B Program Instruction

Each AVE unit consisted of eight lessons and each lesson was delivered for 25 minutes per day, eight days per every two weeks. There were eight units total, each centering on a theme connected to a book and state content standards. A ninth unit consisted of activities that reviewed the program words. The unit book used for the read-aloud was an oversized book that the research team chose for its robust language, pictures and story theme, and that the book was appropriate in terms of grade level. Both instructional conditions (reinforcement condition and reinforcement plus) consisted of *base instruction* and *reinforcement*. The *reinforcement plus condition* included additional reinforcement. Base instruction was consistent across those two instructed conditions with reinforcement varying as a function of the condition. See Appendix A for a curriculum chart of each unit.

Base instruction. In base instruction, the target vocabulary was introduced by way of animated PowerPoint slides which displayed a picture of the target word, a definition of the word in English and in Spanish, an explanation if the word was a cognate, and a description of how the picture demonstrated the target word's meaning (see Figure 4). The teacher was prompted to use the word in a sentence in order to model the use of the word. The teacher then gave students a prompt to discuss the word with a partner. At the end of that pre-teaching, base instruction segment, students were prompted to listen for the target word in the story while the teacher read the story and touch their nose if they heard the word. This technique was embedded in the curriculum to promote word consciousness of the target word. Figure 4 shows the script for the base instruction related to the word *informed.*

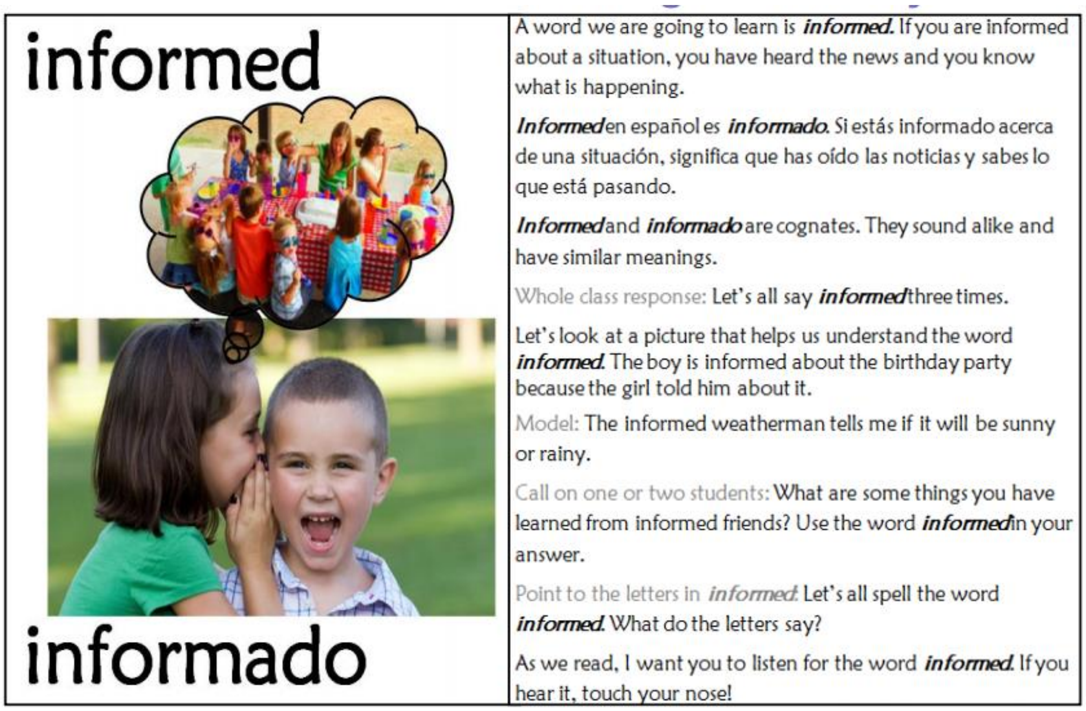


Figure 4. Example of base instruction, retrieved from www.cal.org, used with permission.

After the pre-teaching vocabulary portion of the lesson, the teacher read a quarter of the book every two days using interactive reading techniques, also as part of base instruction. After reading each page, the teacher summarized the text using embedded explanations for the target vocabulary words. The teachers also asked students' questions about the story themes with the target vocabulary to engage students in conversation about the text. In order to review the text with the students, on days 2-4 the teacher started the interactive reading portion of the lesson by conducting a picture walk of the previous day's reading. The picture walk included embedded explanations for the targeted vocabulary as well. Figure 5 shows an example of the interactive reading for the book *A Chair for My Mother*.

Point to the picture on page 9: On this page, Rosa is telling about what happened not long ago. Rosa and her mother are walking down the street towards their house. An informed neighbor tells them their house is on fire. An informed person is someone who knows what is happening.

Call on one student: How do you think Rosa and her mother felt when they found out that their house was on fire? [Responses will vary.]

Call on one student: What did the informed neighbor tell Rosa and her mom? Remember, ***informed*** means "knows what is happening." Start your answer with, "The informed neighbor..." [Anticipated response: The informed neighbor tells Rosa and her mom that their house is on fire.]

Figure 5. Interactive reading example with target word: delicate. Retrieved from www.cal.org, used with permission.

Reinforcement Activities. The base instruction was augmented with reinforcement activities to promote word learning. Reinforcement activities consisted of stratified oral and written activities. These activities included group active response activities, games using the target vocabulary, matching activities, songs, and glossary writing activities. Figure 6 shows a reinforcement glossary where students completed a sentence with the target word and were prompted to draw their own picture that shows the word. Figure 7 shows an example of an active response technique where students decipher between examples that show the target word and nonexamples.

The final day of each unit was spent on review activities and a weekly progress monitoring assessment. Figure 8 shows the progress monitoring assessment and Figure 9 shows a review song. Appendix A shows the lesson scope and sequence for AVE3B program.

Figure 6. Example reinforcement glossary. Retrieved from www.cal.org, used with permission.

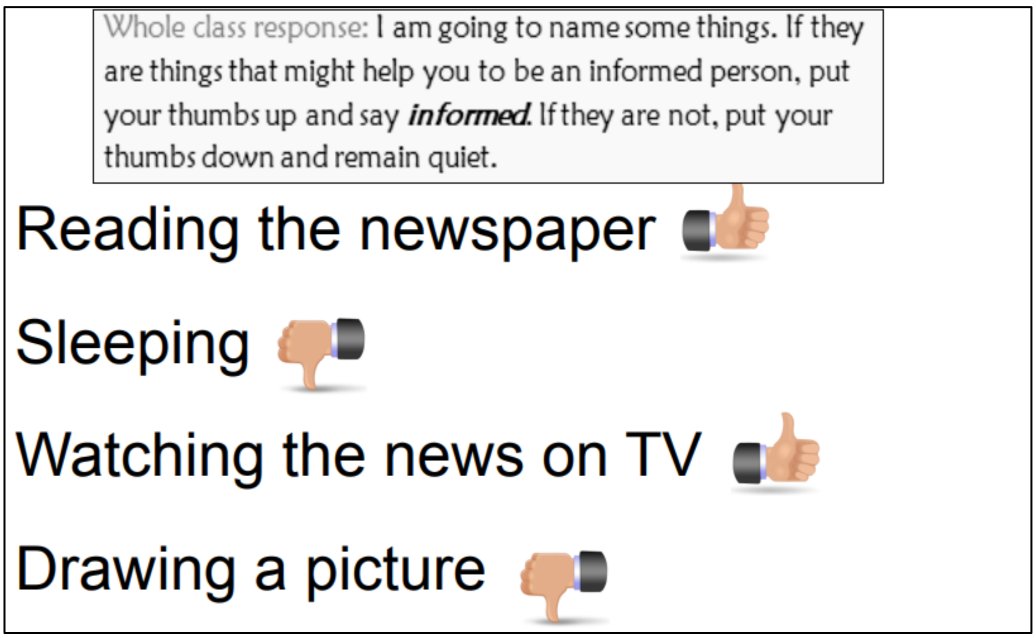
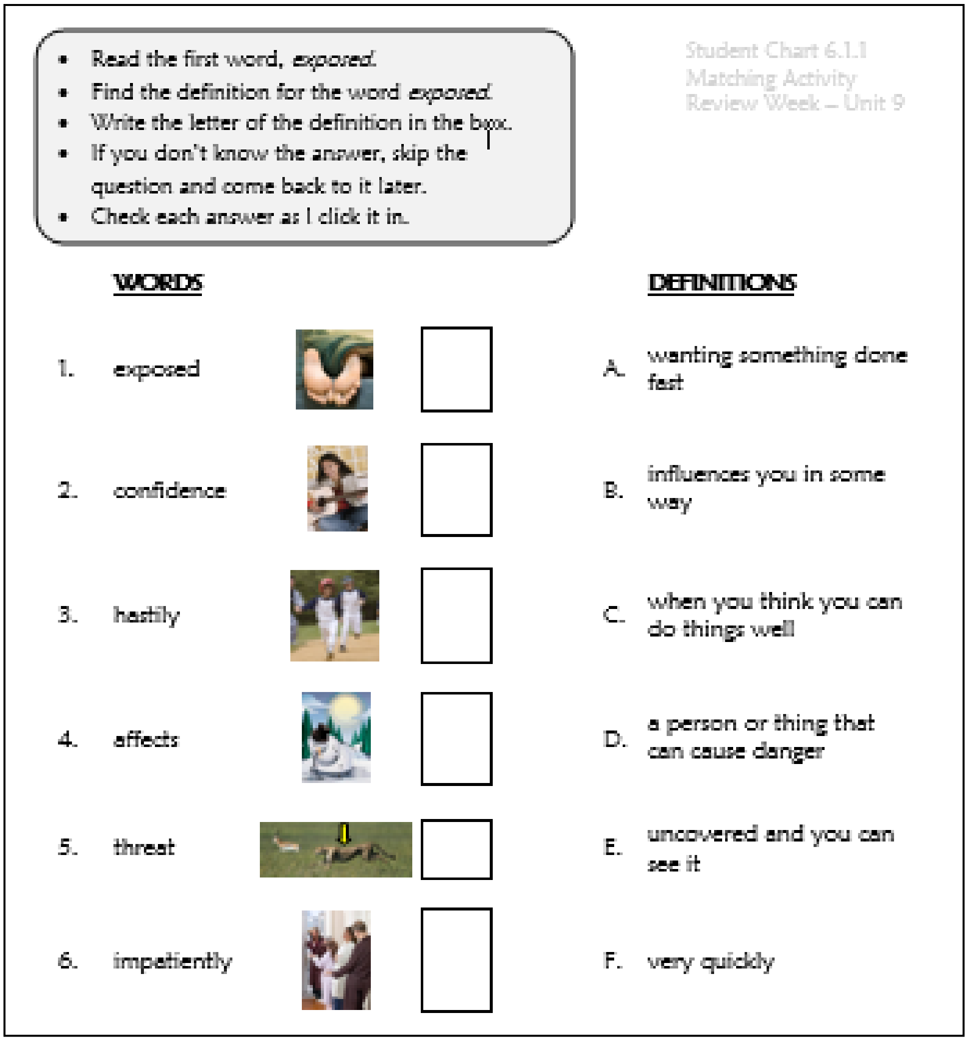


Figure 7. Example oral reinforcement activity. Retrieved from www.cal.org, used with permission.

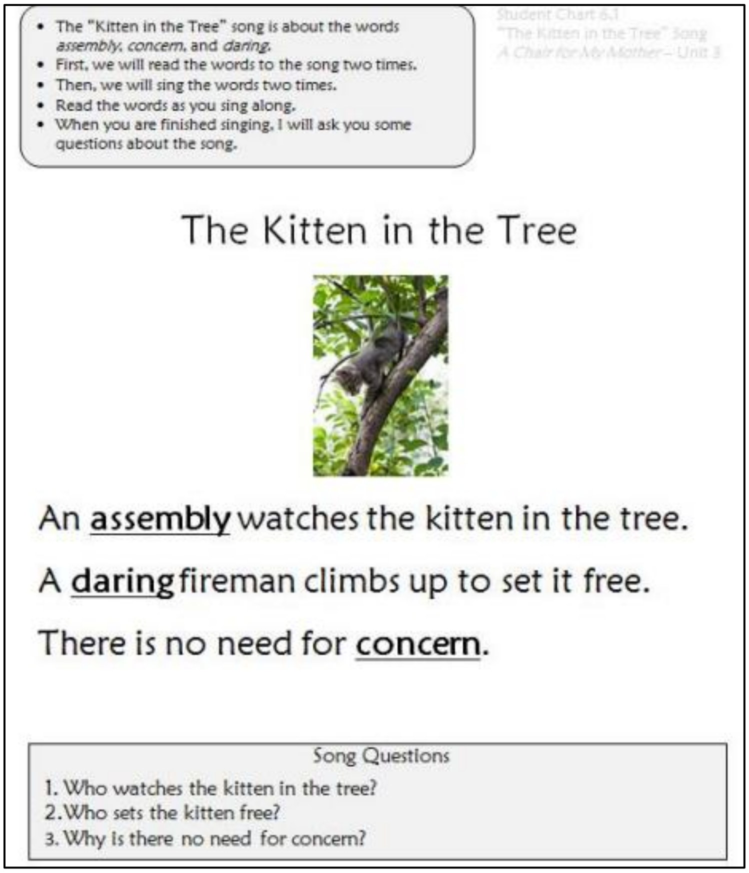


- Read the first word, *exposed.*
- Find the definition for the word *exposed.*
- Write the letter of the definition in the box.
- If you don't know the answer, skip the question and come back to it later.
- Check each answer as I click it in.

Student Chart 6.1.1
Matching Activity
Review Week – Unit 9

WORDS			DEFINITIONS	
1.	exposed		A.	wanting something done fast
2.	confidence		B.	influences you in some way
3.	hastily		C.	when you think you can do things well
4.	affects		D.	a person or thing that can cause danger
5.	threat		E.	uncovered and you can see it
6.	impatiently		F.	very quickly

Figure 8. Example progress monitoring assessment, used with permission from Center for Applied Linguistics.



Student Chart 6.1
"The Kitten in the Tree" Song
A Chair for My Mother – Unit 3

- The "Kitten in the Tree" song is about the words *assembly*, *concern*, and *daring*.
- First, we will read the words to the song two times.
- Then, we will sing the words two times.
- Read the words as you sing along.
- When you are finished singing, I will ask you some questions about the song.

The Kitten in the Tree

An **assembly** watches the kitten in the tree.

A **daring** fireman climbs up to set it free.

There is no need for **concern**.

Song Questions

1. Who watches the kitten in the tree?
2. Who sets the kitten free?
3. Why is there no need for concern?

Figure 9. Example reinforcement song. Retrieved from www.cal.org, used with permission.

Reinforcement activities took place in three areas: within reinforcement activities of that unit, embedded in reinforcement activities in subsequent weeks, and within a final review unit that took place weeks 18-20 of the program. In order to control for time effects, post-unit reinforcement was spiraled through the curriculum, meaning the words were encountered at equal intervals through the curriculum.

The level of reinforcement was varied across the reinforcement and reinforcement plus condition within the weekly unit as well as post-unit (spiraled within other units) or in a two-week review unit. Table 3 provides an explanation for the number of encounters by condition.

Table 3. Encounters within Reinforcement and Reinforcement Plus Conditions

Condition	**Within Unit Reinforcement**	**Post-Unit Reinforcement**	**Final Review Week Reinforcement**	**Totals**
Reinforcement Condition	2 encounters	2 encounters	2 encounters	6 total reinforcement encounters
Reinforcement Plus Condition	6 encounters	4 encounters	4 encounters	14 total reinforcement encounters

As Table 3 shows, the encounters with the target word included 6 total encounters (after the base instruction) for the reinforcement condition and 14 total encounters for the reinforcement plus condition. Figure 10 shows this design in graphical form and Table 4 breaks the encounters down by type of encounter.

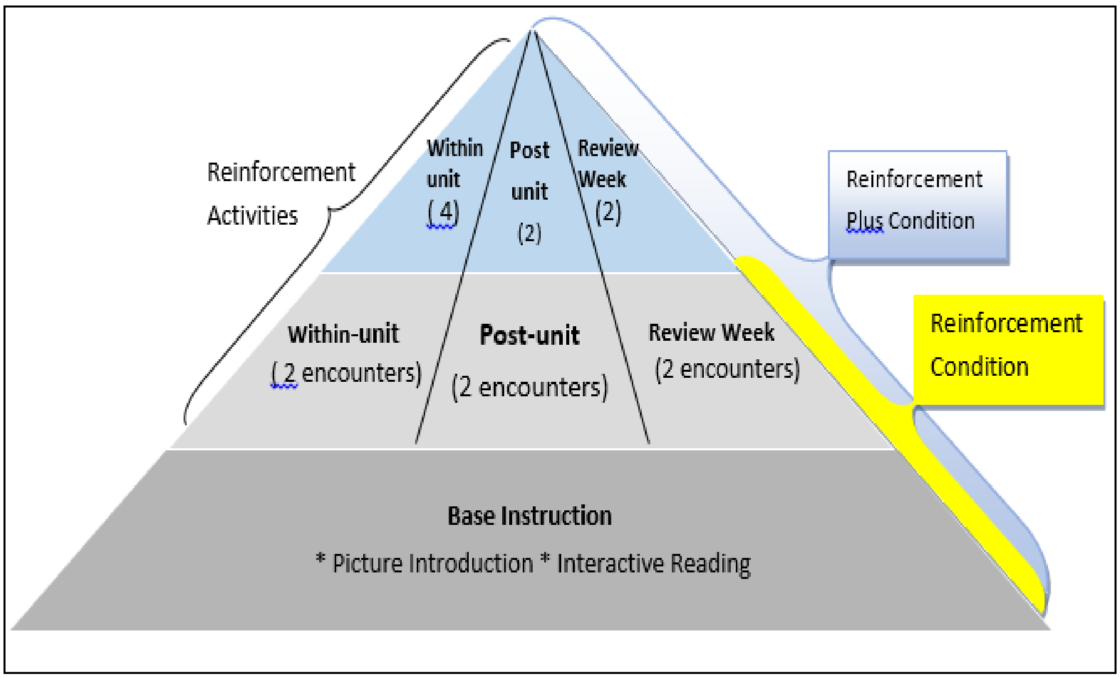


Figure 10. Encounters by condition

Table 4. Encounters by Language Modality

	Within-Unit Reinforcement	Post-Unit Reinforcement	Review Week	Total written Encounters After Base Instruction	Total Oral Encounters After Base Instruction	Total
Reinforcement Condition	1 written 1 oral	1 written 1 oral	1 written 1 oral	3	3	6
Reinforcement Plus Condition	2 written 2 oral	1 written 1 oral	1 written 1 oral	7	7	14

As Table 4 and Figure 10 depict, the reinforcement plus condition received eight more written and oral encounters compared to the reinforcement condition.

Professional Development

Since the program was delivered in the core instruction via the classroom teachers and not trained research assistants, a professional development workshop provided participating teachers a training session prior to the start of the program. That professional development session was delivered over one full day and designed and led by the AVE project investigators, myself, and the two project coaches. During the workshop, the teachers engaged in discussions about the research-based practices utilized in the AVE program and practiced implementing the AVE program itself. At the end of the workshop, teachers received the first unit.

Classroom teachers implementing the AVE program were provided job-embedded technical assistance through coaching with instructional technical assistance (TA) providers who were classified as master teachers in the district and therefore had extensive experience with instructional coaching in that district. The project coaches met with the implementing classroom teachers every two weeks to observe the lessons and provide feedback on implementation. When needed, the coaches modeled certain instructional components of the AVE program lessons. The project coaches met weekly with the research team to debrief the observations and discuss any outstanding issues. Coaches also received the next unit at those meetings to disseminate to the implementing teachers.

Measures and Protocols

In addition to the fidelity instrument, two norm-referenced metrics were utilized to assess standardized literacy gains in the AVE 3B substudy. One researcher-developed measure was administered to assess students' receptive knowledge of intervention vocabulary.

Test of Oral Language Development-Primary: Fourth Edition (TOLD-P:4). To assess generalized English vocabulary development, the Oral Vocabulary subtest of the TOLD-P:4 was administered at pretest and posttest. The Oral Vocabulary subtest of the TOLD-P:4 assesses students expressive understanding of target words. In this assessment students were asked to provide a definition for a target word. All subtests of the TOLD-P:4 have been found to be reliable and valid and that data is reported in Newcomer and Hamill (2005).

Aprenda 3: La prueba de logros en español, Tercera edición. The Aprenda 3 is modeled after the Stanford 10 and measures various aspects of Spanish-speaking K-12 students' academic achievement, including language proficiency. The test is conducted in Spanish and administered to students whose home language is Spanish. Therefore, the Aprenda 3 is a norm referenced measure for Spanish language proficiency. Reference norms are provided for by the test developer, Pearson, for students in the U.S. who come from Spanish-speaking homes. The Aprenda 3 is documented as reliable and valid (Pearson, 2005).

Acquisition of Vocabulary in English Vocabulary Assessment (AVEVA). For the curriculum-aligned measure, the AVEVA was administered pre-and-post intervention. The AVEVA is a researcher-developed curriculum-based measure of the intervention vocabulary and developed in Year 2 of the larger study. Development procedures and psychometric data for the AVEVA are reported in detail in Artzi, Massoud, August, and Barr (2011) and August et al. (2018).

Each item on the AVEVA consisted of a picture that demonstrated the meaning of the target word, a sentence that described the picture, and a child-friendly definition of the target word (see Figure 11).

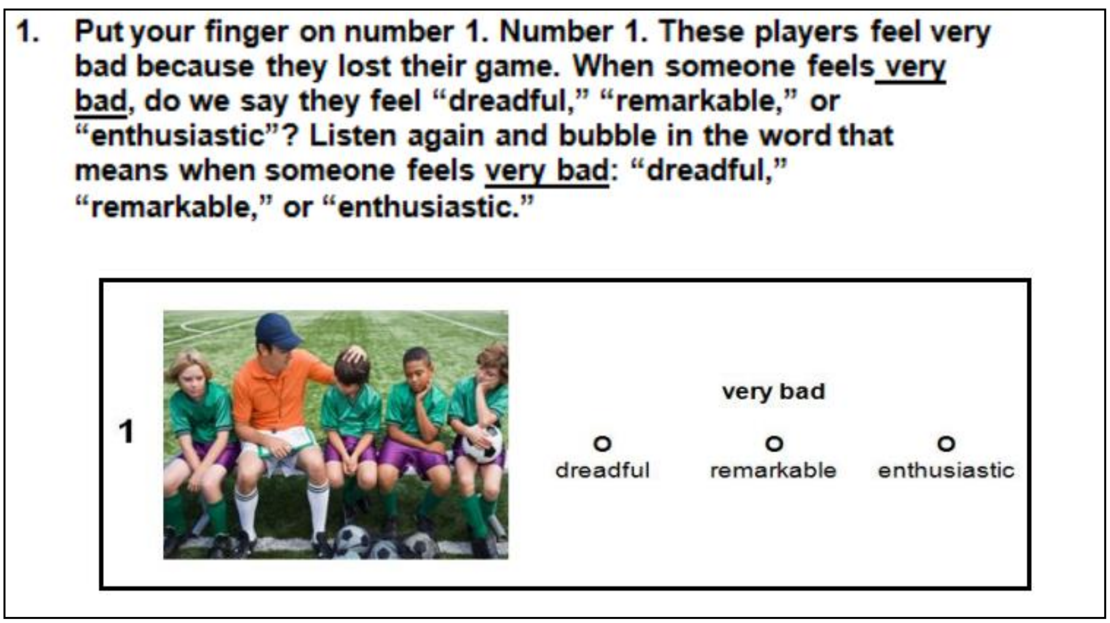


Figure 11. Example of AVEVA item. Retrieved from www.cal.org, used with permission.

During the AVEVA assessment sessions, the student participants were prompted by the test administrator to bubble in the correct target word. To investigate reliability of the AVEVA, coefficient alphas were computed with the following results: pre-test AVEVA, α = .78; Post-test 2 AVEVA, α = .91. The correlation with TOLD Oral Vocabulary subtest (r=.66, p<.01) indicates a good degree of concurrent validity.

Intervention Fidelity Protocol. Fidelity data were collected twice by the trained research assistants through the duration of the program. AVE's fidelity protocol was designed to assess instructional adherence to the AVE program, or that the program was implemented as conceptualized. The protocol also probed for the quality of the instructional delivery. The instrument was also developed in Year 2 of the project and reported in detail in August et al. (2018). It is also described below.

In order to assess instructional adherence, instructional components were noted on the fidelity protocol when implemented. Instructional quality for each activity consisted of four components: maintaining proper pacing, managing the activity materials without interrupting instruction, eliciting student responses effectively, and appropriately responding to students. Instructional quality was rated on a 4-point Likert scale with 4 indicating excellent level of instructional quality. Fidelity data across all levels illustrates that the AVE program was executed with a high degree of how it was intended and with a high degree of quality with ratings averaging between 3 and 4 each session.

Analytic Approach

This investigation involves four questions pertaining to the ELL students' word learning with respect to the AVE vocabulary program in the AVE3B substudy:

1. What is the effect of the three conditions (exposure, reinforcement, and reinforcement plus) on word learning?

2. Does type of word (abstract cognates, concrete cognates, abstract noncognates, concrete noncognates) influence students' acquisition and learning in each condition?
3. Does ELL word acquisition and learning across the three conditions vary as a function of generalized English vocabulary knowledge?
4. Does ELL word learning across the three conditions vary as a function of Spanish language proficiency?

I draw on an analytic modelling technique using a special case of hierarchical generalized linear models (HGLM). Three models are presented in Chapter 4. Model 1 is presented in relation to research questions 1 and 2 as the first research topic, Model 2 is presented in relation to research question 2 as the second research topic, and Model 3 relates to research question 4 as the third research topic. Therefore, the first two questions are unpacked with one model, then I modeled separately for the research questions related to English and Spanish language proficiency.

For multilevel data with non-normally distributed error terms and nonlinear structural models, HGLM models allow for an integrated modeling procedure. HGLM models, therefore, extend generalized linear models by allowing error terms to vary and therefore account for nested research situations often found in education research. For this reason, HGLM models are frequently utilized in situations such as this study because this modelling technique allows for error terms to be introduced at different data levels (e.g., student-level and classroom-level). In the case of this study the nested modeling structure is appropriate as the AVE program took place across twelve different classrooms as opposed to one classroom.

In this study, the level-1 outcome of interest is whether a student answered a test item correctly. This outcome is dichotomous; items are either scored as 1 (correct answer by student) or as 0 (incorrect answer). The multi-level analytic model for this study includes test items nested in students and students nested in classrooms.

The level-1 model in HGLM consists of three parts: (1) a sampling model, (2) a link function, and (3) a structural model. With dichotomous outcomes, such as the outcome of interest in this study, the use of a standard level-1 HLM model would be inappropriate for a few reasons. First, there are no restrictions on the predicted values of the level-1 outcome in a standard HLM model. In contrast, the predicted value of a dichotomous outcome, if viewed as a probability, is limited to the 0 to 1 interval. Therefore, a nonlinear transformation of the predicted value, such as a logit transformation, ameliorates the issue. Given the predicted value of the outcome, the level-1 random effect can take on only one of 2 values, and therefore cannot be normally distributed. Thus, this scenario would not satisfy the HLM assumption that errors are normally distribute and does not meet the assumption of linearity. In this instance, therefore, the level-1 random effect cannot present with homogenous variance. Therefore, the linearity and normality assumptions will not apply.

Because the level-1 outcome is dichotomous and the assumptions of linearity and normality do not apply, I used an HGLM with a binomial (or Bernoulli) sampling procedure and a logit link function. The logit link function is expressed as η_{ikj} = $\log(\varphi_{ijk}/1\text{-}\varphi_{ijk})$, where η_{ijk} is the log of the odds of a correct item, and φ_{ijk} is the probability of success, or a correct item-- ηikj is the log odds for the $_i$th item for the $_j$th student in the $_k$th classroom. This link function transforms the data from probabilities to logit scores.

Instead of using a probit link function, the logit link function is preferred for a few reasons. First, it is the traditional function used to fit dichotomous data using the Bernoulli. Also, compared to the probit link, the logit link displays better interpretability since the probit link requires additional data transformation maneuvers.

I used a nonlinear model because the outcome is dichotomized as correct and incorrect. The model is mixed because the conditions and word-types are treated as fixed effects and the student language proficiency is treated as a random effect; therefore, the outcomes will be predicted by the variables measuring instructional conditions and also by word-type (as measured by cognate status and concreteness level of the word). While both conditions and word-types are within-subject's predictors, the predictors of generalized English vocabulary and Spanish language proficiency are between-subject's predictors.

To execute the three HGLM models, I used the GLIMMIX procedure on Statistical Analysis Software (SAS). The GLIMMIX procedure fits statistical models presenting with correlations that are not normally distributed. GLIMMIX maximizes power because it takes into account non-normal distributions and correlations of multiple outcomes (Little et al., 2006).

According to SAS Institute the estimation strategy used by GLIMMIX for each model is an open-form estimation, with a maximum likelihood. Therefore, the probability of answering an item correct is based on a theta scale on an X axis of a person's ability, with the probability of the item correct on the Y axis. Item difficulty is always defined at a 50% point, which is the point of maximum discrimination. Essentially, GLIMMIX determines the pattern of correct and incorrect responses for the entire sample and examines different possible ability values for each case multiple times. This includes the most likely theta ability level for any case, and the other possible ability levels. By means of iterations, GLIMMIX calculates different possibilities of abilities and item difficulty and determines the most likely predicted value probabilistically.

Using the GLIMMIX procedure I ran three-level HGLM models, with items nested within students, and students nested within classrooms. I did not reflect the nesting of classrooms within schools given the small number of schools in the study. I fit the hierarchical nonlinear mixed-effects models using the logit link function of GLIMMIX. To address the four research questions, I utilized the parameter estimates associated with the model predictors. Item performance on the AVEVA is considered the outcome variable with item-level (level 1) predictors of word attributes and AVEVA pretest item scores, and student-level (level 2) predictors of Spanish and English language proficiency. Random error terms at each level account for the nesting of items within students and students within classrooms.

The equations for the three analytic models are presented below.

Model 1, level-1 equation:

$$\eta_{ijk} = \pi_{0jk} + \pi_{1ijk} \text{ (Cognate status}_{ijk}) + \pi_{2ijk} \text{ (Concreteness status}_{ijk}) + \pi_{3ijk} \text{ (AVEVA pretest score}_{ijk}) + e_{ijk}$$

where:

- η_{ijk} is the log of the odds of success on item i for student j in classroom k;
- π_{0jk} is the mean success across all items for student j in classroom k;
- Cognate $status_{ijk}$ is the cognate status for item i;
- Concreteness $status_{ijk}$ is the concreteness status for item i;
- AVEVA pretest $score_{ijk}$ is the AVEVA pretest score for item i for student j; and
- e_{ijk} is a random item-level error term.

Model 1 level-2 equation:

$\pi_{0jk} = \beta_{00k} + r_{0jk}$

where:

- β_{00k} is the mean success across all students in classroom k; and
- r_{0jk} is a random student-level error term or "student effect".

Model 1 level-3 equation:

$\beta_{00k} = \gamma_{000} + u_{00k}$

where:

- γ_{000} is the the mean success across all classrooms (grand mean); and
- u_{00k} is a random classroom-level error term or "classroom effect".

Model 2, level-1 equation:

$\eta_{ijk} = \pi_{0jk} + \pi_{1ijk}$ (Cognate $status_{ijk}$) + π_{2ijk} (Concreteness $status_{ijk}$) + π_{3ijk} (AVEVA pretest $score_{ijk}$) + e_{ijk}

where:

- η_{ijk} is the log of the odds of success on item i for student j in classroom k;
- π_{0jk} is the mean success across all items for student j in classroom k;
- Cognate $status_{ijk}$ is the cognate status for item i;
- Concreteness $status_{ijk}$ is the concreteness status for item i;
- AVEVA pretest $score_{ijk}$ is the AVEVA pretest score for item i for student j; and
- e_{ijk} is a random item-level error term.

Model 2 level-2 equation:

$\pi_{0jk} = \beta_{00k} + \beta_{01k}$(TOLD language proficiency $category_{jk}$) + r_{0jk}

where:

- β_{00k} is the mean success across all students in classroom k,
- TOLD language proficiency $category_{jk}$ is the TOLD language proficiency category (high or low) for student j; and,
- r_{0jk} is a random student-level error term or "student effect".

Model 2 level-3 equation:

$\beta_{00k} = \gamma_{000} + u_{00k}$

where:

- γ_{000} is the the mean success across all classrooms (grand mean); and
- u_{00k} is a random classroom-level error term or "classroom effect"

Model 3, level-1 equation:

$\eta_{ijk} = \pi_{0jk} + \pi_{1ijk}$ (Cognate $status_{ijk}$) + π_{2ijk} (Concreteness $status_{ijk}$) + π_{3ijk} (AVEVA pretest $score_{ijk}$) + e_{ijk}

where:

- η_{ijk} is the log of the odds of success on item i for student j in classroom k;
- π_{0jk} is the mean success across all items for student j in classroom k;
- Cognate $status_{ijk}$ is the cognate status for item i;
- Concreteness $status_{ijk}$ is the concreteness status for item i;
- AVEVA pretest $score_{ijk}$ is the AVEVA pretest score for item i for student j; and
- e_{ijk} is a random item-level error term.

Model 3 level-2 equation:

$\pi_{0jk} = \beta_{00k} + \beta_{01k}(\text{Aprenda language proficiency category}_{jk}) + r_{0jk}$

where:

- β_{00k} is the mean success across all students in classroom k,
- Aprenda language proficiency $category_{jk}$ is the TOLD language proficiency category (high or low) for student j; and,
- r_{0jk} is a random student-level error term or “student effect”.

Model 2 level-3 equation:

$\beta_{00k} = \gamma_{000} + u_{00k}$

where:

- γ_{000} is the the mean success across all classrooms (grand mean); and
- u_{00k} is a random classroom-level error term or “classroom effect”

For each model, I present the fit statistics, to assess item fit, the Type III Test for Fixed Effects, and the Least Square Means, which include the parameter estimates on item difficulty. In the Fit Statistics, the log likelihood is the residual log likelihood and the chi-squared statistic measures the residual sum of squares in the model.

I computed effect sizes from the arithmetic means using the Cohen's d procedure (see Rice & Harris, 2005). Cohen's d is expressed as: $d = \frac{M_1 - M_2}{SD_{pooled}}$. In some cases, I computed a simple average of effect size in order to explore combined components (such as outcomes aggregated by instruction versus exposed). This procedure follows parameters outlined in Turner and Bernard (2006) which specifies that a simple average for effect size is adequate to report effects within a study; though a weighted average is used in instances across studies to pool effect size difference. Since PROC GLIMMIX only provides the standard error in the Least Squares Means, I conducted PROC MEAN to obtain the simple means and then PROC T TEST to secure the mean difference and pooled standard deviation.

Assumptions Related to the Models

Assumptions of HGLM models include the following: (1) the expected outcome at each level can be represented as a linear function and (2) that the random effects at each level can be assumed to be normally distributed with homogenous variance. However, in the models reported in this work, the assumptions of linearity and normality do not apply because the outcomes are dichotomous and not linearly distributed. With each model, I investigated the relevant statistical assumptions and report findings at the end of Chapter 4.

Variables Used in the Models

Because of the complexity of the research design, there are a number of variables presented in the models and figures shown in Chapter 4. EX refers to the exposure condition, R refers to the reinforcement condition, and R+ refers to the reinforcement plus condition. WAVE 1 refers to the scores from the test administration at pretest and WAVE 2 are the posttest scores. COGNATE is the cognate status with COG meaning the words were coded as Spanish-English cognates and NOT the words coded as non-cognates. COMPLEXITY relates to the word's abstractness status with CONCRETE the indicator for those concrete words that are conceptually simple and ABSTRACT the conceptually complex words.

When presenting figures in Chapter 4, variables in the legend of each figure are shown with the condition first (as EX, R, or R+), followed by the conceptual complexity status (as ABSTRACT or CONCRETE) and then cognate status (noted as COG for cognate and NOT for words that are not cognates). Therefore, EXCONCRETECOG is the variable for a word in the exposed condition that is concrete and a cognate and R+ABSTRACTNOT is the variable for a word in the reinforcement plus condition that is an abstract non-cognate.

Chapter 4: Results

In this chapter, I detail the results of the analyses that I conducted using SAS's PROC GLIMMIX in relation to each research question, providing three models. The first two questions are unpacked with one model in research topic 1. Then I modeled a separate model for the research questions related to English and Spanish language proficiency, as research topics 2 and 3.[9]

<u>Research Topic 1: What is the effect of reinforcement on word learning? Does type of word (concrete cognates, concrete noncognates, abstract cognates, abstract noncognates) influence students' acquisition and learning in each condition?</u>

To answer the first two research questions, I used SAS GLIMMIX to model word type (cognates and abstractness) and instructional condition by item with the dichotomous outcome, correct and incorrect. I refer to this model as Model 1. Model 1's response distribution is binary, the link function is logit, and the variance matrix is not blocked. I used a residual PL estimation technique, and the method for degrees of freedom is containment.

The descriptive data for the Model 1 variables are presented in Table 5 and Table 6 shows the fit statistics for Model 1. Because there were seven comparisons in the model (word type and instructional conditions), I utilized a family-wise correction of $p<.007$ to evaluate tests of significance (calculated as .05/7).

---

[9] I ran a model initially with gender and gender was not significant with ($p=.26$); therefore gender was not investigated further.

Table 5. Simple Means by Condition and Word-type for Model 1

Condition	WAVE	Cognate Status	Abstractness Status	N Obs	N	Mean Percentage correct	Std Dev
EX	1	NOT	ABSTRACT	1547	1508	0.3932361	0.4886305
			CONCRETE	1326	1298	0.3713405	0.4833495
		COG	ABSTRACT	2210	2156	0.5199443	0.4997180
			CONCRETE	2873	2803	0.5957902	0.4908260
	2	NOT	ABSTRACT	1484	1446	0.4585062	0.4984477
			CONCRETE	1272	1251	0.4412470	0.4967347
		COG	ABSTRACT	2120	2071	0.6059874	0.4887556
			CONCRETE	2756	2684	0.6818182	0.4658573
R	1	NOT	ABSTRACT	1326	1299	0.4357198	0.4960418
			CONCRETE	1326	1290	0.4139535	0.4927314
		COG	ABSTRACT	1326	1298	0.5423729	0.4983933
			CONCRETE	1326	1290	0.4364341	0.4961353
	2	NOT	ABSTRACT	1272	1249	0.7309848	0.4436255
			CONCRETE	1272	1238	0.7859451	0.4103309
		COG	ABSTRACT	1272	1249	0.8118495	0.3909889
			CONCRETE	1272	1237	0.7647534	0.4243244
R+	1	NOT	ABSTRACT	1326	1286	0.4331260	0.4957005
			CONCRETE	1326	1297	0.3731689	0.4838330
		COG	ABSTRACT	1326	1295	0.5583012	0.4967812
			CONCRETE	1326	1299	0.4672825	0.4991206
	2	NOT	ABSTRACT	1272	1228	0.8379479	0.3686487
			CONCRETE	1272	1249	0.8590873	0.3480709
		COG	ABSTRACT	1272	1245	0.8650602	0.3417965
			CONCRETE	1272	1250	0.8624000	0.3446175

Table 6. Fit Statistics for Model 1

-2 Res Log Pseudo-Likelihood	162776.8
Generalized Chi-Square	33996.78
Gener. Chi-Square / DF	0.96

The data in Table 6 indicate that Model 1 is a suitable fitting model with a generalized chi-squared degrees of freedom ration of .96. Methods for meeting assumptions across the models are described at the end of this chapter. Table 7 shows the Type III Test for Fixed Effects for Model 1. Table 7 provides information related to the Type III Test of Fixed Effects.

Table 7. Type III Test of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
Condition	2	35197	10.67	<.0001
WAVE	1	35197	2546.55	<.0001
Condition*WAVE	2	35197	436.10	<.0001
Cognatestatus	1	35197	9.51	0.0020
Condition*cogstatus	2	35197	2.04	0.1306
WAVE*cogstatus	1	35197	6.37	0.0116
Condit*WAVE*cogstatus	2	35197	6.92	0.0010
Complexity	1	35197	0.03	0.8687
Condition*Complexity	2	35197	0.52	0.5956
WAVE*Complexity	1	35197	21.30	<.0001
Condit*WAVE*Comple	2	35197	6.56	0.0014
cogstatus*Complexity	1	35197	0.04	0.8505
Condit*cogstatus*Comple	2	35197	0.88	0.4136
WAVE*cogstatus*Comple	1	35197	0.68	0.4102
Condit*WAVE*cog*Compl	2	35197	1.38	0.2509

As can be seen in Table 7, there is a significant interaction between WAVE and CONDITION (F=436.10; p<.001), WAVE and COMPLEXITY (F=21.30; p<.001). CONDITION, WAVE, and COMPLEXITY of a word just missed significance (F=6.56; p=.0014). Although CONDITION and WAVE were both highly significant variables (p<.0001), and COGNATE STATUS was a moderately significant factor (p=.0002), COMPLEXITY missed significance (p=.8687).

Table 8 reports the Least Square Means for the variables in Model 1. In Table 8, the conditions are depicted as EX for the incidental exposure condition, R for the reinforcement condition, and R+ for the reinforcement plus condition. WAVE 1 is the variable for the pretest scores of the AVEVA before the intervention and WAVE 2 are the posttest sores. Words that are noted as COG in the COGNATE STATUS column are cognates and the words noted as NOT in the cognate column are not cognates. The estimate reported in the table below refers to the parameter estimate related to item difficulty.

Table 8. Least Square Means for Model 1

Cond-ition	Conceptual Complexity	Wave	Cog-nate	Estimate	Standard Error	t Value	Pr > |t|
EX	ABSTRACT	1	NOT	0.5738	0.2772	2.07	0.0385
EX	CONCRETE	1	NOT	0.6071	0.2973	2.04	0.0412
EX	ABSTRACT	1	COG	-0.07439	0.2339	-0.32	0.7504
EX	CONCRETE	1	COG	-0.4763	0.2079	-2.29	0.0220
EX	ABSTRACT	2	NOT	0.2351	0.2772	0.85	0.3964
EX	CONCRETE	2	NOT	0.2855	0.2972	0.96	0.3367
EX	ABSTRACT	2	COG	-0.4857	0.2344	-2.07	0.0382
EX	CONCRETE	2	COG	-0.9220	0.2087	-4.42	<.0001
R	ABSTRACT	1	NOT	0.2997	0.2970	1.01	0.3130
R	CONCRETE	1	NOT	0.4095	0.2971	1.38	0.1682
R	ABSTRACT	1	COG	-0.1789	0.2971	-0.60	0.5471
R	CONCRETE	1	COG	0.2930	0.2969	0.99	0.3236
R	ABSTRACT	2	NOT	-1.1244	0.2988	-3.76	0.0002
R	CONCRETE	2	NOT	-1.4479	0.3002	-4.82	<.0001
R	ABSTRACT	2	COG	-1.6451	0.3012	-5.46	<.0001
R	CONCRETE	2	COG	-1.2740	0.2992	-4.26	<.0001
R+	ABSTRACT	1	NOT	0.3178	0.2971	1.07	0.2848
R+	CONCRETE	1	NOT	0.5979	0.2973	2.01	0.0443
R+	ABSTRACT	1	COG	-0.2459	0.2968	-0.83	0.4074
R+	CONCRETE	1	COG	0.1548	0.2970	0.52	0.6022
R+	ABSTRACT	2	NOT	-1.8154	0.3023	-6.01	<.0001
R+	CONCRETE	2	NOT	-1.9875	0.3033	-6.55	<.0001
R+	ABSTRACT	2	COG	-1.9930	0.3033	-6.57	<.0001
R+	CONCRETE	2	COG	-2.0623	0.3038	-6.79	<.0001

*Degrees of freedom for all variables is 35197.

As can be seen students made significant gains on words instructed in the three conditions. Given these findings, I investigated the three-way interaction (CONDITION, WAVE, and COMPLEXITY) further by conducting follow-up pairwise comparisons. Table 9 shows those estimates.

Table 9. Follow Up Pairwise Analyses

Label	Estimate	Standard Error	t Value	Pr > |t|
Taught versus (vs.) untaught	-11.1880	2.5495	-4.39	<.0001
Gains in R vs. R+ at cognates	-6.9974	0.2016	-34.72	<.0001
Gains in R vs. R+ at non-cognates	-7.9999	0.1995	-40.10	<.0001
Gains in cognates vs. non-cognates at R	0.2483	0.1881	1.32	0.1869
Gains in cognates vs. non-cognates at R+	0.7543	0.2086	3.62	0.0003
Gains in R vs. R+ at concrete	-1.3781	0.1999	-6.89	<.0001
Gains in R vs. R+ at abstract	-0.9900	0.1974	-5.02	<.0001
Gains in concrete vs. abstract at R	-0.5341	0.1881	-2.84	0.0045
Gains in concrete vs. abstract at R+	-0.9222	0.2086	-4.42	<.0001

* Degrees of Freedom for all variables is 35197.

As can be seen in Table 9, gains of the instructed conditions are highly significant compared to the exposed words (t= -4.39; p<.001). The taught items were -11.1880 points easier on posttest (wave 2) compared to pretest. Gains in cognates versus non-cognates were only significant at the REINFORCEMENT PLUS condition (t= 3.62; p=.0003), but not the REINFORCEMENT condition (t=1.32; p=.1869). However, these differences are very minimal with estimates of .75 and .24, respectively. Similarly, gains with CONCRETE versus ABSTRACT words were more significant at the REINFORCEMENT PLUS condition (t= -4.42; p<.0001) compared to the REINFORCEMENT condition (t= 2.84; p>.0045). The estimates indicate very minimal differential effects by conceptual complexity (concrete versus abstract words).

The following figures depict the gains in Table 7. The negative slope found in Figure 12 is due to the fact that the items from wave 1 (pretest) to wave 2 (posttest) became easier, reflecting the vocabulary growth of the student sample.

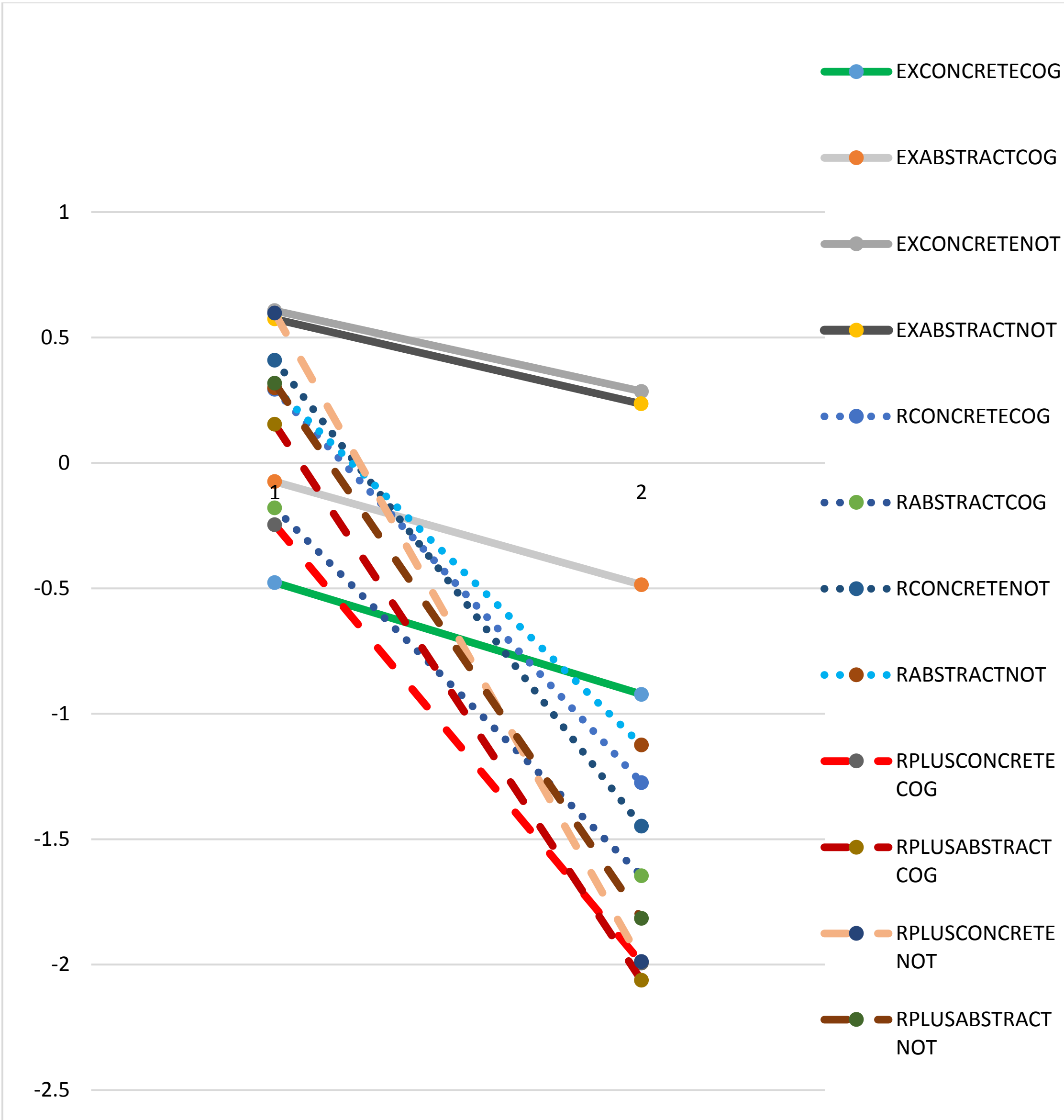


Figure 12. Pre- to post-test estimates by word type and condition.

Using the estimates, I calculated the growth from wave 1 (pretest) to wave 2 (posttest) and plotted the growth through the following figures (Figures 13-17). These figures show the growth on a Y axis of item ease. Figure 13 shows all conditions and word-types disaggregated. Figures 14 and 15 show the words at each level of cognate status with conceptual complexity disaggregated. Figures

16 and 17 show the words at each level of conceptual complexity with cognate status disaggregated.

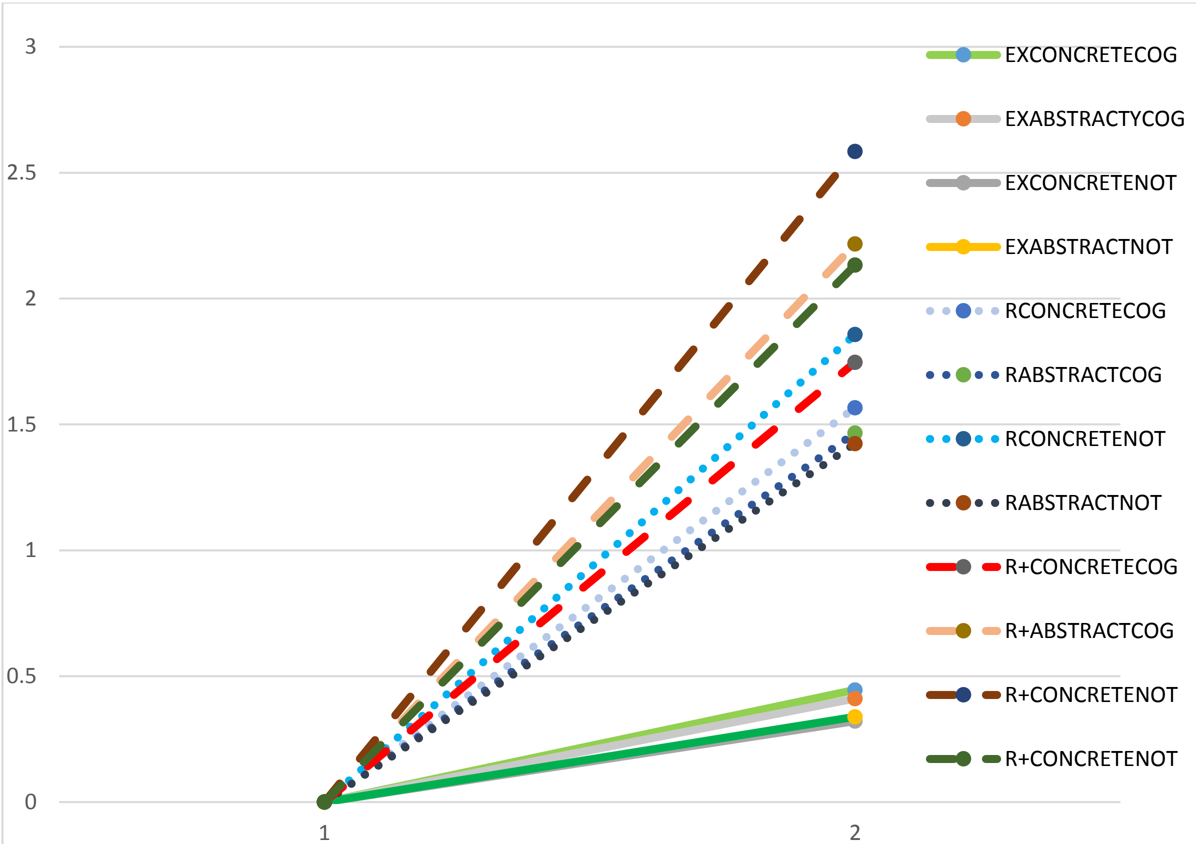


Figure 13. Growth of estimates by condition and word-type.

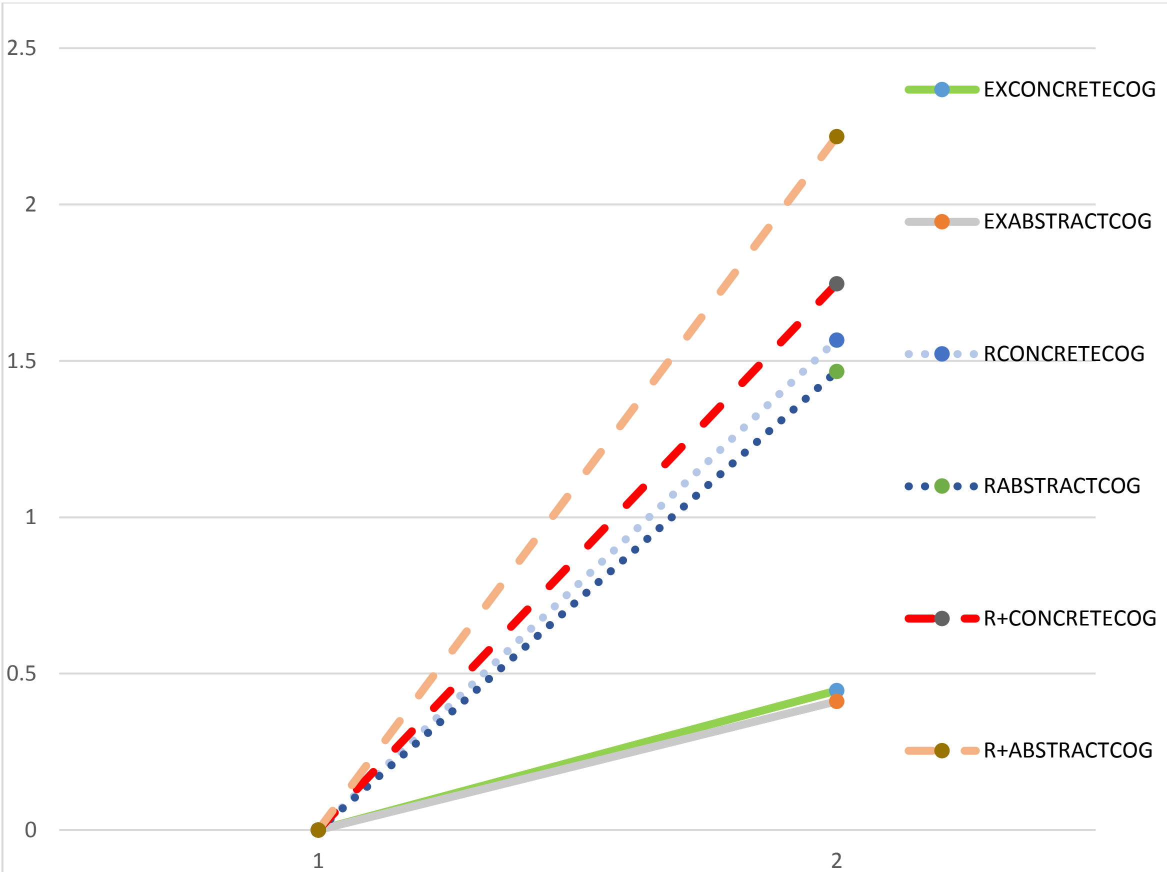


Figure 14. Growth of cognates by condition with conceptual complexity disaggregated.

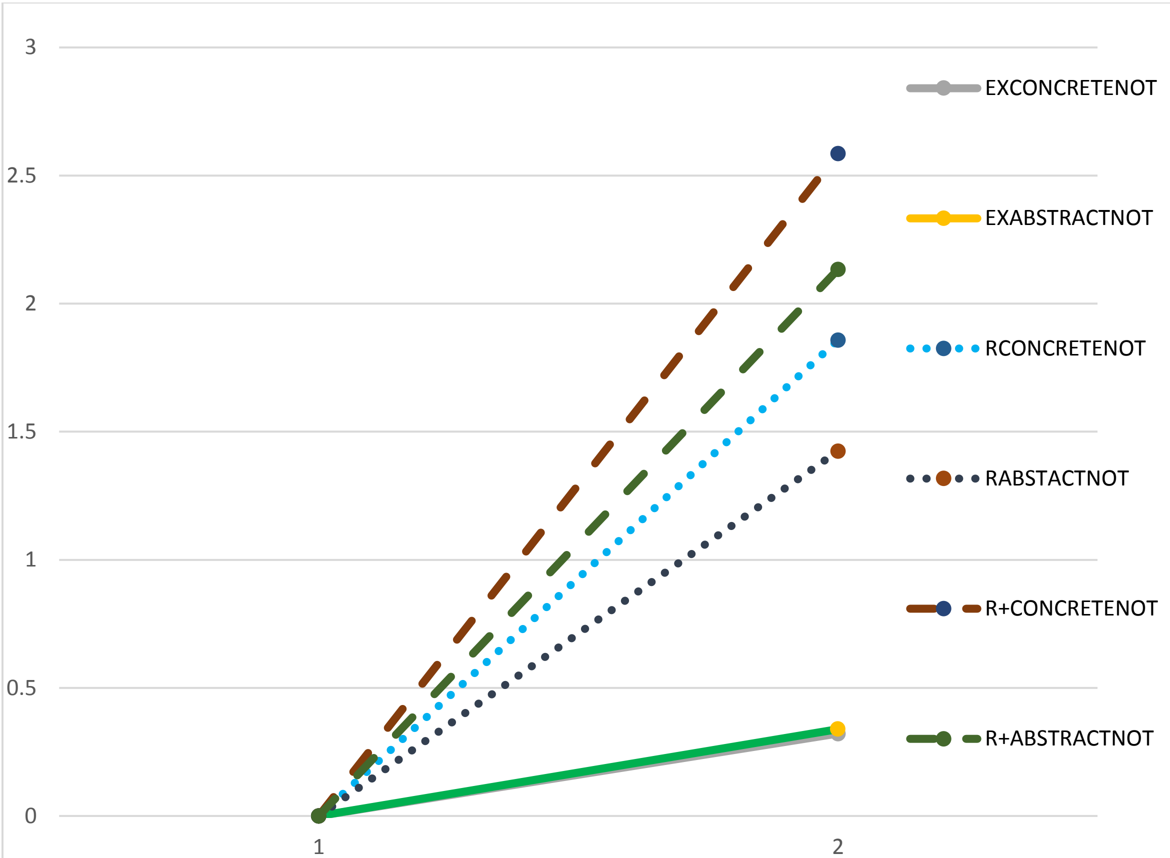


Figure 15. Growth of noncognates by condition, conceptual complexity disaggregated.

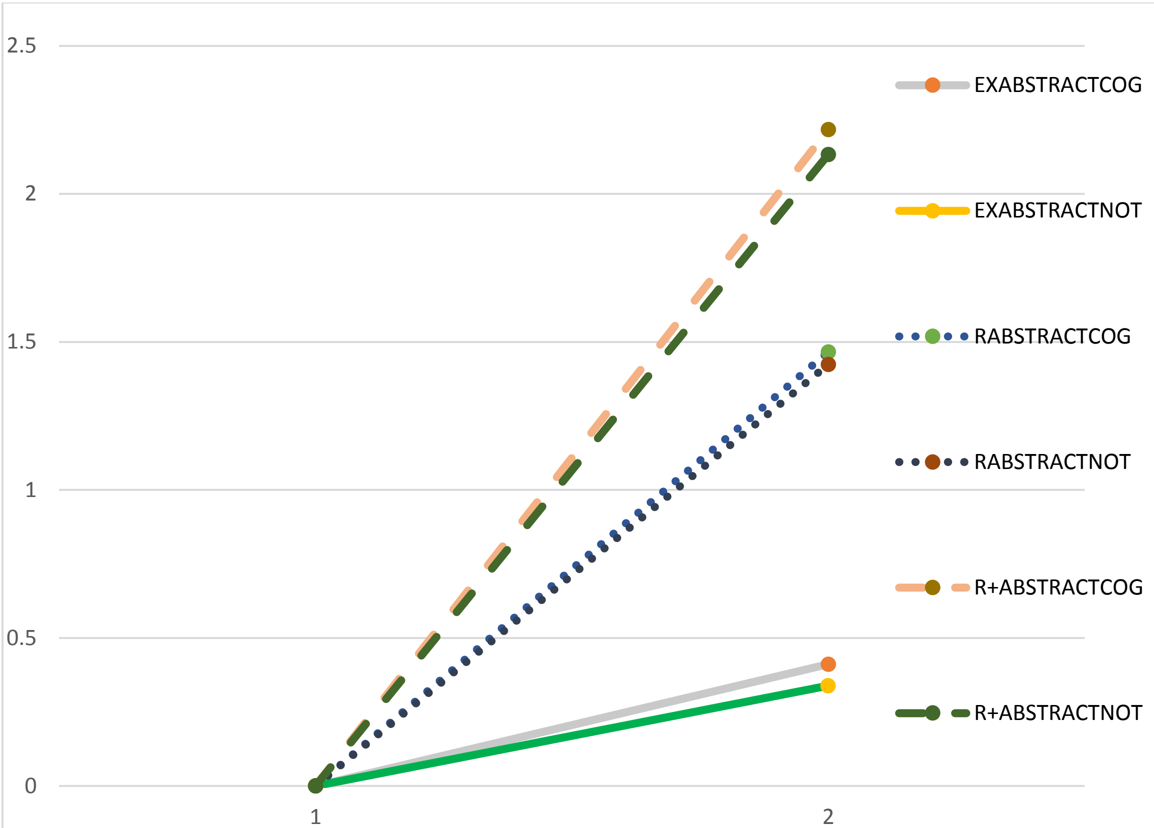


Figure 16. Growth of abstract words by condition, cognate status disaggregated.

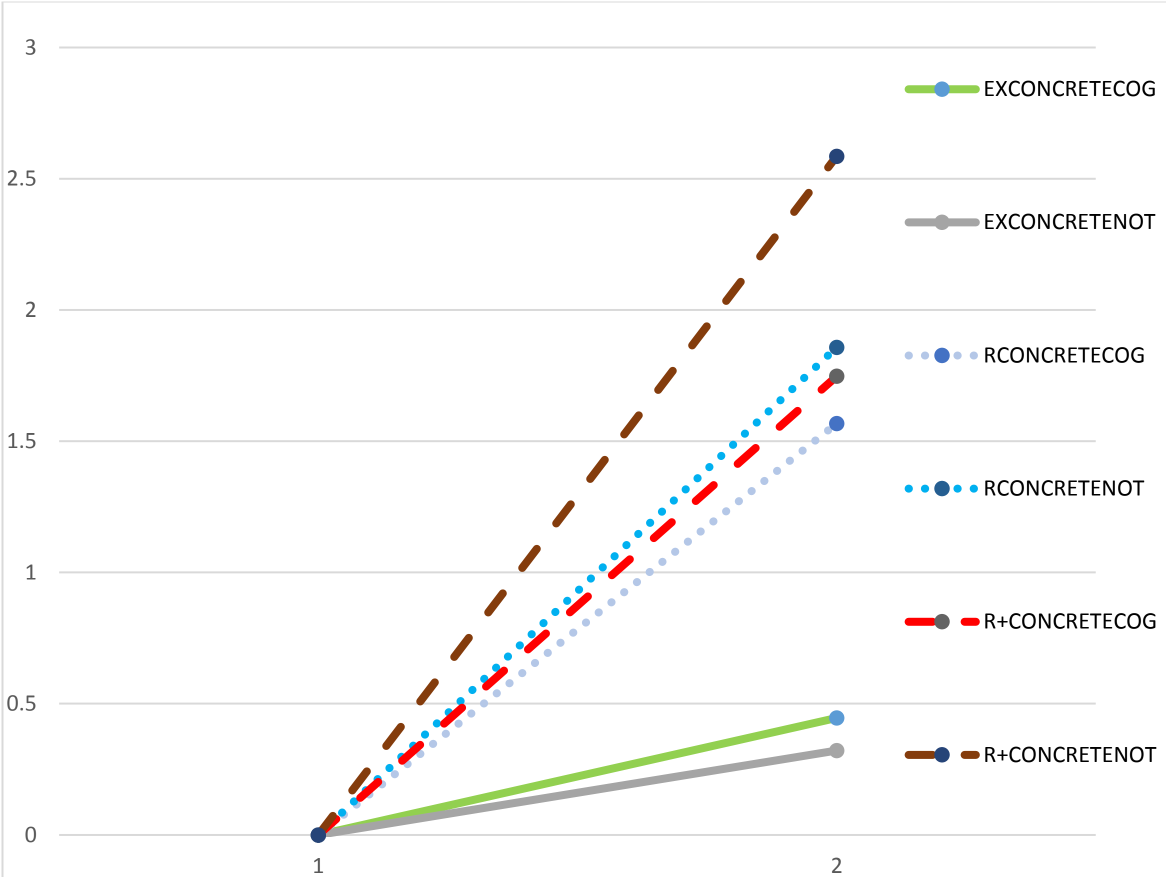


Figure 17. Growth of concrete words by condition, cognate status disaggregated.

Research Topic 2: Does ELL word acquisition and learning across the three conditions vary as a function of generalized English vocabulary knowledge?

To answer the third research question, I used PROC GLIMMIX to model a second model. First, I took a split half of the Test of Oral Language Development IV scores from the Oral Vocabulary subtest. From the median, the scores in the top half of the sample were in the high category, with the variable HIGH in the tables and HI in the figures. The bottom half of the TOLD Oral Vocabulary scores were coded as in the low category, with the variable LOW in the tables and figures.

I used SAS GLIMMIX to model word type (cognates and abstractness), instructional condition, and TOLD category by item with the dichotomous outcome, correct and incorrect. I refer to this model as Model 2.

The model's response distribution is binary, the link function is logit, and the variance matrix is not blocked. I used a residual PL estimation technique, and the method for degrees of freedom is containment. Because there are nine comparisons in the model (TOLD category, word type, and instructional conditions), I applied a family-wise correction of .05/9, or p<.005, to evaluate tests of significance.

Table 10 shows the descriptive data for the variables in this model. In Table 10, TOLD CAT refers to the TOLD category and HI are those scores that are in the top half of the sample, relative to each other, and LOW are those scores in the bottom half of this sample.

Table 10. Means and Standard Deviations for Model 2 Variables

told_cat	Condition	WAVE	N	Mean	Std Dev
Hi	EX	1	2986	0.56	0.50
		2	2870	0.67	0.47
	R	1	1991	0.53	0.50
		2	1915	0.88	0.33
	R+	1	1991	0.50	0.50
		2	1916	0.93	0.25
Low	EX	1	2931	0.44	0.50
		2	2751	0.50	0.50
	R	1	1951	0.40	0.49
		2	1836	0.69	0.46
	R+	1	1954	0.43	0.50
		2	1833	0.79	0.41

Table 11 shows the fit statistics for Model 2. This information demonstrates that Model 2 is a sufficient fitting model.

Table 11. Fit Statistics for Model 2

-2 Res Log Pseudo-Likelihood	124576.1
Generalized Chi-Square	25623.39
Gener. Chi-Square / DF	0.95

Table 12 shows the Type III Test for Fixed Effects for Model 1. In this table, COND is the condition variable (incidental exposure, reinforcement, reinforcement plus), WAVE is the wave variable (pretest and posttest), COGNATE is whether a word is a cognate or not, COMPLEXITY is whether the word was coded as abstract or concrete, and TOLD CAT is the told category with the split half group, scores in the HIGH category the top half and the LOW category the bottom half of the TOLD Oral Vocabulary scores.

Table 12. Type III Tests of Fixed Effects for Model 2

Effect	Num DF	F Value	Pr > F
Condition	2	11.80	<.0001
WAVE	1	1945.68	<.0001
Condition*WAVE	2	343.56	<.0001
Cogstat	1	9.16	0.0025
Condition*cogstatus	2	1.88	0.1528
WAVE cogstatus	1	2.45	0.1176
Condit*WAVE*cogstatus	2	4.77	0.0085
Complexity	1	0.01	0.9260
Condition*Complexity	2	0.38	0.6866
WAVE Complexity	1	22.34	<.0001
Condit* WAVE Comple	2	5.97	0.0025
cogstatus*Complexity	1	0.07	0.7915
Condit*cogstatus*Comple	2	0.92	0.4000
WAVE *cogstatus*Comple	1	2.63	0.1046
Cond*WAVE*cogn*Compl	2	0.87	0.4177
told_cat	1	91.17	<.0001
Condition*told_cat	2	7.78	0.0004
WAVE *told_cat	1	116.11	<.0001
Condit*WAVE*told_c	2	15.04	<.0001
cogstatus*told_cat	1	15.26	<.0001
Condit*cogstatus*told_c	2	0.62	0.5365
WAVE*cogstatus*told_c	1	0.33	0.5635

Cond*WAVE*cogn*told_	2	0.53	0.5863
Complexity*told_cat	1	2.51	0.1130
Condit*Comple*told_c	2	0.52	0.5940
WAVE*Comple*told_c	1	0.68	0.4093
Cond*WAVE*Comp*told_	2	0.26	0.7697
cognum*Comple*told_c	1	5.84	0.0157
Cond*cogn*Comp*told_	2	1.01	0.3648
WAVE*cogn*Comp*told_	1	0.13	0.7167
Con*WAVE*cog*Com*told	2	1.28	0.2778

*Denominator degrees of freedom for all variables is 26640.

As can be seen in Table 12, there is no significant four-way interaction between CONDITION, WAVE, COMPLEXITY and TOLD CATEGORY (F= .26; p=.7697). There is no significant five-way interaction between CONDITION, WAVE, COGNATE, COMPLEXITY, and TOLD CATEGORY (F= 1.28; p=.2778). There is, however, a significant interaction between CONDITION, WAVE and TOLD CATEGORY (F=15.04; p<.0001). Table 13 shows the Least Square Means for that significant three-way interaction.

Table 13. Least Squares Means for Condition, Wave, and TOLD Category

Condition	**told_cat**	**WAVE**	**Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
EX	HIGH	1	-0.1685	0.1430	-1.18	0.2388
	LOW	1	0.3831	0.1435	2.67	0.0076
	HIGH	2	-0.6971	0.1438	-4.85	<.0001
	LOW	2	0.1290	0.1438	0.90	0.3699
R	HIGH	1	-0.1355	0.1631	-0.83	0.4059
	LOW	1	0.4491	0.1635	2.75	0.0060
	HIGH	2	-2.2079	0.1730	-12.77	<.0001
	LOW	2	-0.8759	0.1650	-5.31	<.0001
R+	HIGH	1	-0.01905	0.1631	-0.12	0.9070
	LOW	1	0.3124	0.1633	1.91	0.0558
	HIGH	2	-2.8637	0.1829	-15.65	<.0001
	LOW	2	-1.4310	0.1672	-8.56	<.0001

*Degrees of freedom for all variables is 26640.

The following figures depict the gains in Table 12. In these figures, R refers to the reinforcement condition, RPLUS is the reinforcement plus condition, EX is the incidental exposure condition, ABSTRACT refers to conceptually complex (abstract) words, CONCRETE is conceptually simple (concrete) words. COG refers to whether the word is a cognate, and NOT means the word was not a cognate. HI refers to those students who scored in the top half of the TOLD Oral Vocabulary subtest and LOW refers to those students who scored in the bottom half of the TOLD Oral Vocabulary subtest.

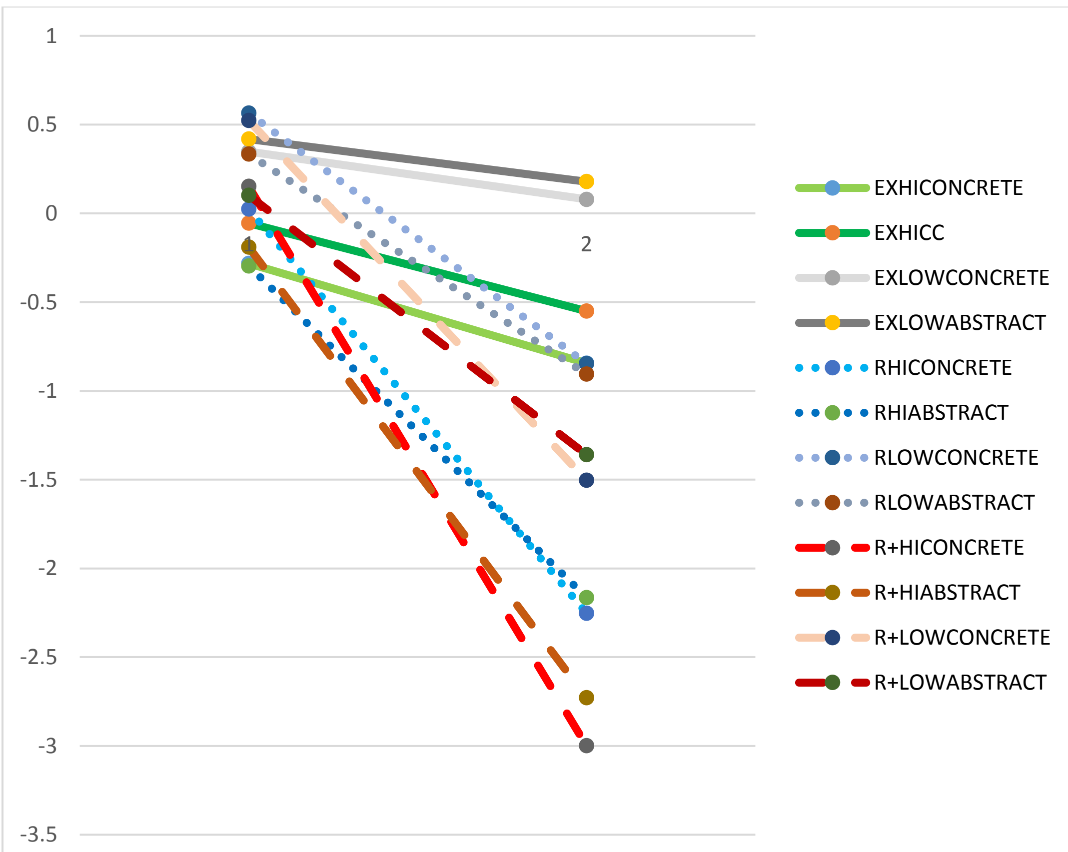


Figure 18. Least square means from Wave 1 to Wave 2 by condition, conceptual complexity, and TOLD category.

Using the estimates, I calculated the growth from wave 1 (pretest) to wave 2 (posttest) and plotted the growth through the following figures (Figures 19-21). These figures show the growth on a Y axis of item ease.

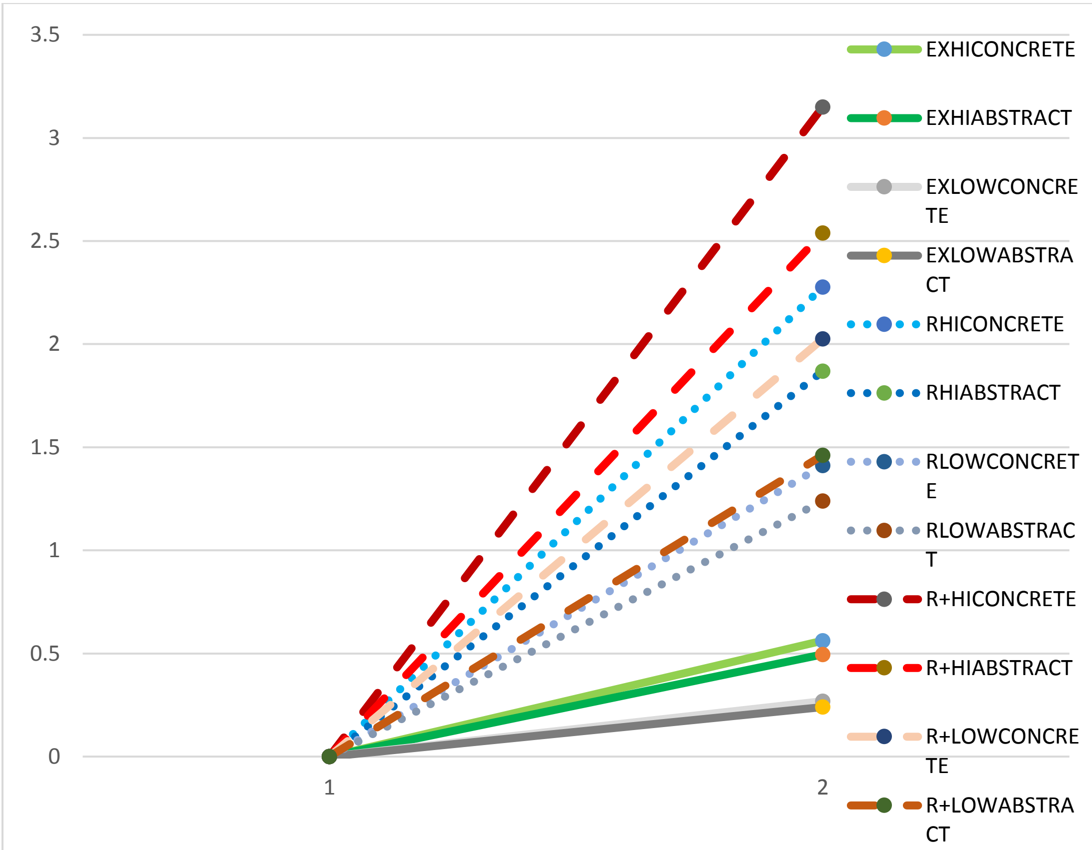


Figure 19. Growth by condition, conceptual complexity, and TOLD category.

The following figure shows the findings related to the TOLD categories (HIGH and LOW) in the exposed condition as well as the reinforcement conditions aggregated. Therefore, the variable INSTRUCT are the outcomes from the reinforcement and reinforcement plus condition.

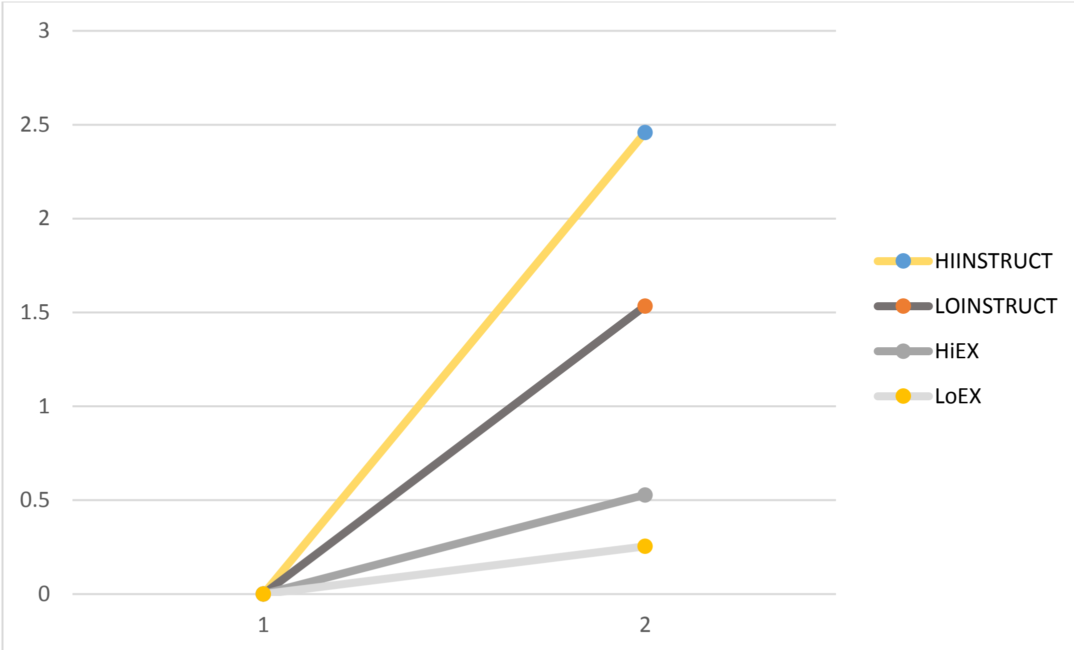


Figure 20. Growth of estimates by TOLD category in exposure and instruction.

As shown in Figure 20 the difference between the high TOLD group (HI) and the low TOLD group (LO) is significant with large effects across exposure as well as with instruction.

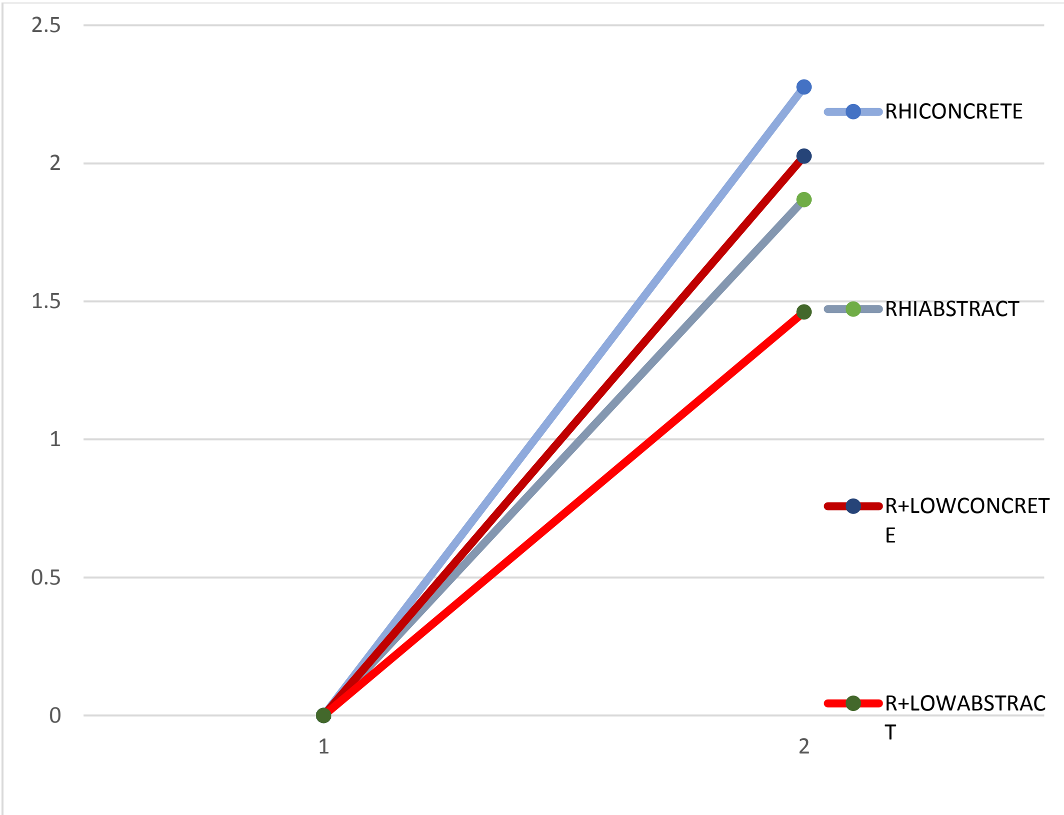


Figure 21. Gains of students in the HIGH TOLD category in the REINFORCEMENT condition versus gains of students in LOW TOLD category in the REINFORCEMENT PLUS condition.

As shown in this figure, reinforced high TOLD conceptually simple (concrete) words compared to reinforced plus low conceptually simple words with moderate effects. Collapsing conceptual complexity word type, students in the high told group showed gains of 2.07; whereas students in the low TOLD group showed gains of 1.75. The difference between the groups presents with moderate effects.

Research Topic 3: Does word learning of ELLs in this study vary as a function of Spanish language proficiency?

To answer the fourth research questions, I used PROC GLIMMIX to model word type (cognate status and abstractness status/conceptual complexity) and instructional condition. In a similar procedure used with the third research question, I took the split half of the Aprenda total score, with the scores above the median in the high category (HI) and scores below the median in the low category (LOW). I refer to this model as Model 3.

Model 3's response distribution is binary, the link function is logit, and the variance matrix is not blocked. I used a residual PL estimation technique, and the method for degrees of freedom is containment. Because there are nine comparisons in the model (Aprenda category, word type, and instructional conditions), I applied a family-wise correction of .05/9, or $p<.005$, to evaluate tests of significance.

Table 13 displays the descriptive statistics for the variables in Model 3. In Table 14 APPRENDA_CAT refers to the category of student scores on the Aprenda. The HI category refers to the variable of the Aprenda scores in the top half of the sample and the LOW category are those scores in the bottom half of the sample.

Table 14. Simple Means for Model 3

apprenda_cat	Condition	WAVE	Cognum	N	Mean	Std Dev
HIGH	EX	1	NOT	1306	0.42	0.49
			COG	2310	0.65	0.48
		2	NOT	1317	0.51	0.50
			COG	2329	0.75	0.43
	R	1	NOT	1206	0.46	0.50
			COG	1205	0.55	0.50
		2	NOT	1215	0.87	0.34
			COG	1216	0.93	0.26
	R+	1	NOT	1204	0.44	0.50
			COG	1207	0.59	0.49
		2	NOT	1214	0.95	0.22
			COG	1218	0.96	0.20
LOW	EX	1	NOT	1337	0.35	0.48
			COG	2363	0.48	0.50
		2	NOT	1367	0.39	0.49
			COG	2403	0.55	0.50
	R	1	NOT	1233	0.40	0.49
			COG	1233	0.43	0.50
		2	NOT	1260	0.65	0.48
			COG	1258	0.65	0.48
	R+	1	NOT	1231	0.37	0.48
			COG	1236	0.44	0.50
		2	NOT	1251	0.75	0.43
			COG	1265	0.77	0.42

Table 15 presents the Model 3 Fit Statistics. These data indicate that Model 3 fits the data sufficiently. The Type III Test of Fixed Effects for Model 3 are in Table 16.

Table 15. Model 3 Fit Statistics

-2 Res Log Pseudo-Likelihood	159563.1
Generalized Chi-Square	32882.79
Gener. Chi-Square / DF	0.96

Table 16. Type III Test of Fixed Effects for Model 3

Effect	Num DF	F Value	Pr > F
Condition	2	13.62	<.0001
WAVE	1	2442.24	<.0001
Condition*WAVE	2	455.33	<.0001
Cogstatus	1	10.74	0.0010
Condition*cogstatus	2	1.73	0.1771
WAVE*cogstatus	1	0.81	0.3672
Condit*WAVE*cogstatus	2	4.84	0.0079
Complexity	1	0.00	0.9496
Condition*Complexity	2	0.52	0.5932
WAVE*Complexity	1	18.71	<.0001
Condit*WAVE*Comple	2	6.91	0.0010
cogstatus*Complexity	1	0.07	0.7893
Condit*cogstatus*Comple	2	1.00	0.3663
WAVE*cogstatus*Comple	1	1.19	0.2754
Cond*WAVE*cogn*Compl	2	2.09	0.1241
apprenda_cat	1	174.20	<.0001
Condition*apprenda_c	2	28.91	<.0001
WAVE*apprenda_cat	1	270.03	<.0001
Condit*WAVE*appren	2	53.18	<.0001
cogstatus*apprenda_cat	1	40.78	<.0001
Condit*cogstatus*appren	2	1.88	0.1531
WAVE*cogstatus*appren	1	0.37	0.5420
Cond*WAVE*cogn*appre	2	2.92	0.0540
Complexit*apprenda_c	1	0.50	0.4791
Condit*Comple*appren	2	2.26	0.1044
WAVE*Comple*appren	1	0.34	0.5593
Cond*WAVE*Comp*appre	2	1.84	0.1593
cogstatus*Comple*appren	1	0.37	0.5435
Cond*cogn*Comp*appre	2	0.28	0.7553
WAVE*cogn*Comp*appre	1	0.10	0.7541
Cond*WAVE*cog*Com*app r	2	0.34	0.7106

*Denominator degrees of freedom for all variables is 34055.

As can be seen on Table 16, COGNATES and APRENDA are significant at (F=40.78; p<.0001). COGNATE and APRENDA is a significant interaction (F=40.78; p<.0001) and CONDITION by WAVE and APRENDA is also significant (F=53.18; p<.0001). However, the four-way interaction between CONDITION, WAVE, COGNATE, and APRENDA is not significant (F=2.92; p<.0540). Below I present the Least Squares Means for CONDITION, WAVE, COGNATE STATUS and APRENDA CATEGORY.

Table 17. Least Squares Means for Condition Wave Cognate Status and Aprenda

Cond-ition	Apren-da Cat	Wave	Cog Stat	Estimate	Standard Error	t Value	Pr > |t|
EX	HIGH	1	NOT	0.3781	0.2113	1.79	0.0735
	LOW	1	NOT	0.7412	0.2113	3.51	0.0005
	HIGH	1	COG	-0.6839	0.1649	-4.15	<.0001
	LOW	1	COG	0.08750	0.1636	0.53	0.5928
	HIGH	2	NOT	-0.03781	0.2110	-0.18	0.8578
	LOW	2	NOT	0.5094	0.2106	2.42	0.0156
	HIGH	2	COG	-1.2611	0.1662	-7.59	<.0001
	LOW	2	COG	-0.2255	0.1635	-1.38	0.1679
R	HIGH	1	NOT	0.2132	0.2180	0.98	0.3280
	LOW	1	NOT	0.4516	0.2176	2.08	0.0380
	HIGH	1	COG	-0.2217	0.2181	-1.02	0.3095
	LOW	1	COG	0.2882	0.2174	1.33	0.1850
	HIGH	2	NOT	-2.0362	0.2272	-8.96	<.0001
	LOW	2	NOT	-0.7131	0.2179	-3.27	0.0011
	HIGH	2	COG	-2.7511	0.2396	-11.48	<.0001
	LOW	2	COG	-0.6956	0.2179	-3.19	0.0014
R+	HIGH	1	NOT	0.2868	0.2181	1.32	0.1884
	LOW	1	NOT	0.5602	0.2179	2.57	0.0101
	HIGH	1	COG	-0.3851	0.2182	-1.77	0.0775
	LOW	1	COG	0.2518	0.2173	1.16	0.2467
	HIGH	2	NOT	-3.0880	0.2484	-12.43	<.0001
	LOW	2	NOT	-1.2102	0.2197	-5.51	<.0001
	HIGH	2	COG	-3.2648	0.2538	-12.86	<.0001
	LOW	2	COG	-1.3449	0.2202	-6.11	<.0001

*Degrees of freedom for all variables is 34055.

On Table 18 I present the Least Squares means table for Aprenda category (HI and LOW) by cognate status (COGSTAT), with NOT the words that are not cognates, and COG the words that are cognates.

Table 18. Least Squares Means by Aprenda Category and Cognate Status

Aprenda Cat	Cog num	Estimate	Standard Error	t Value	Pr > |t|
HIGH	NOT	-0.7140	0.1333	-5.36	<.0001
LOW	NOT	0.05651	0.1307	0.43	0.6655
HIGH	COG	-1.4279	0.1272	-11.23	<.0001
LOW	COG	-0.2731	0.1233	-2.21	0.0268

*DF for all variables is 34055.

The following figures depict the gains in Tables 16 and 17. In these figures, R refers to the reinforcement condition, R+ is the reinforcement plus condition, EX is the incidental exposure condition. ABSTRACT refers to conceptually complex (abstract) words, CONCRETE is conceptually simple (concrete) words, COG refers to whether the word is a cognate, and NOT means the word was not a cognate. HI refers to the top half of the Aprenda total scores, LOW is the bottom half of the Aprenda total scores.

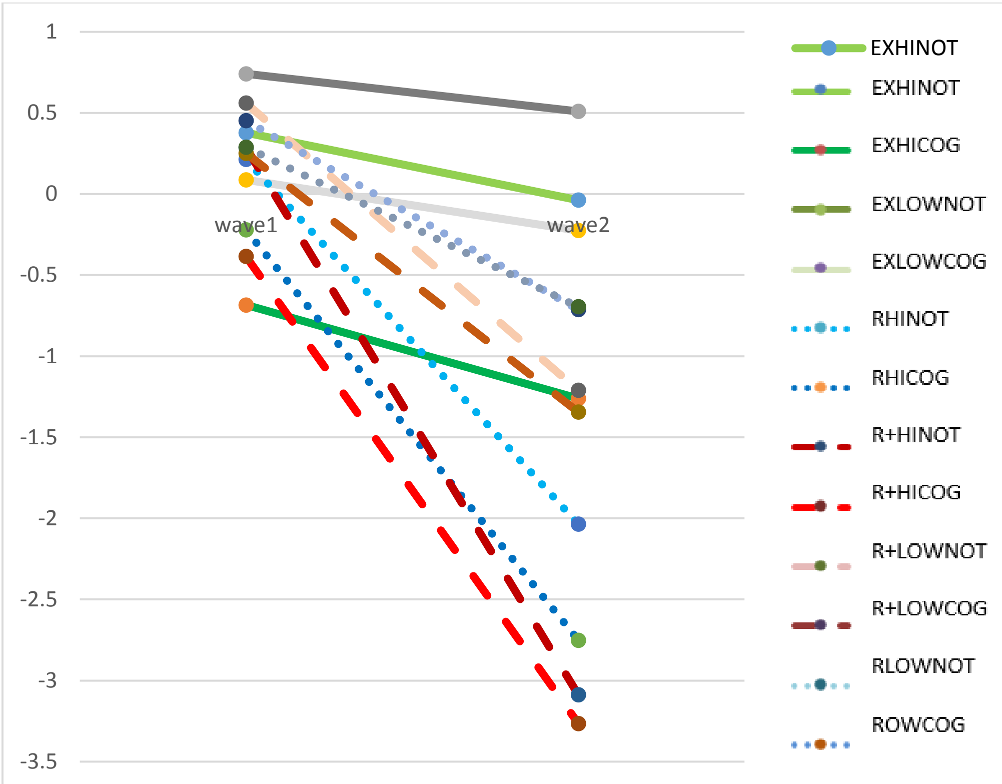


Figure 22. Wave 1 to Wave 2 estimates across condition, Aprenda category and cognate status.

Using the estimates, I calculated the growth from wave 1 (pretest) to wave 2 (posttest) and plotted the growth through the following figures (Figures 23-26). These figures show the growth on a Y axis of item ease.

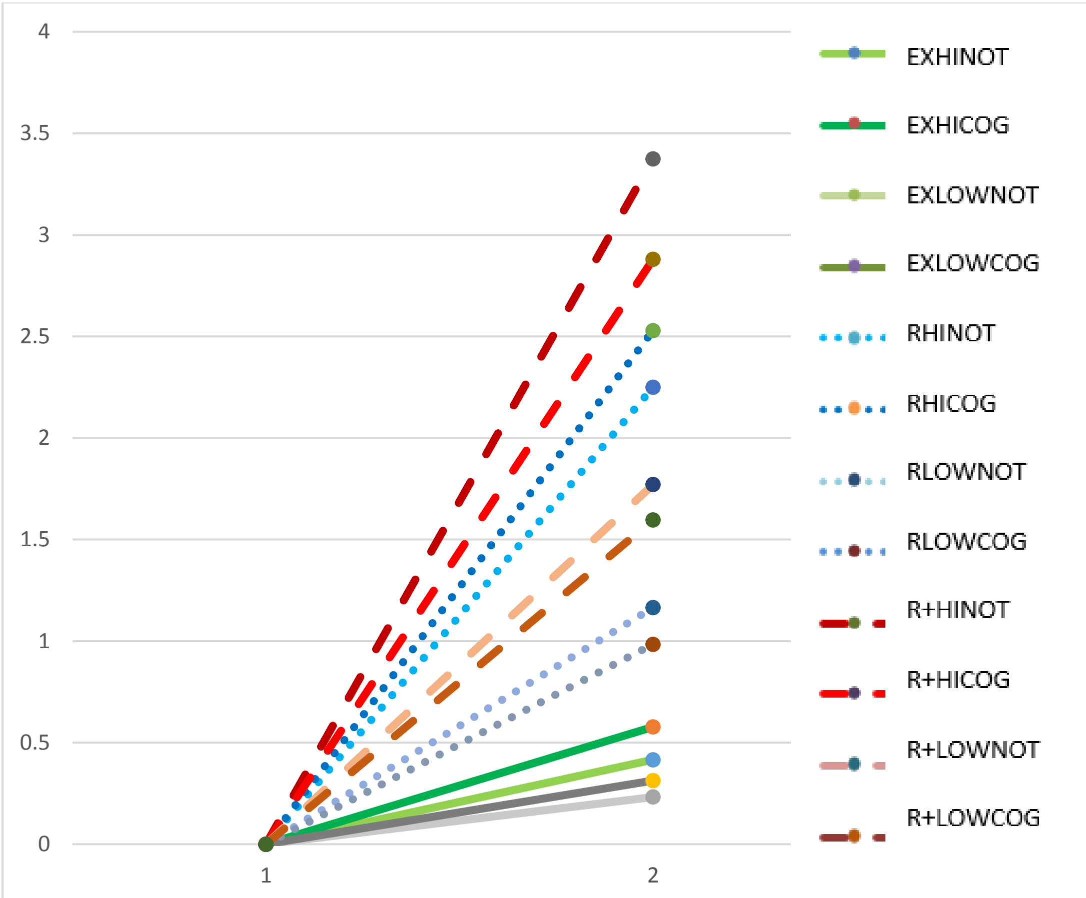


Figure 23. Wave 1 to Wave 2 growth by condition, Aprenda category, and cognate status.

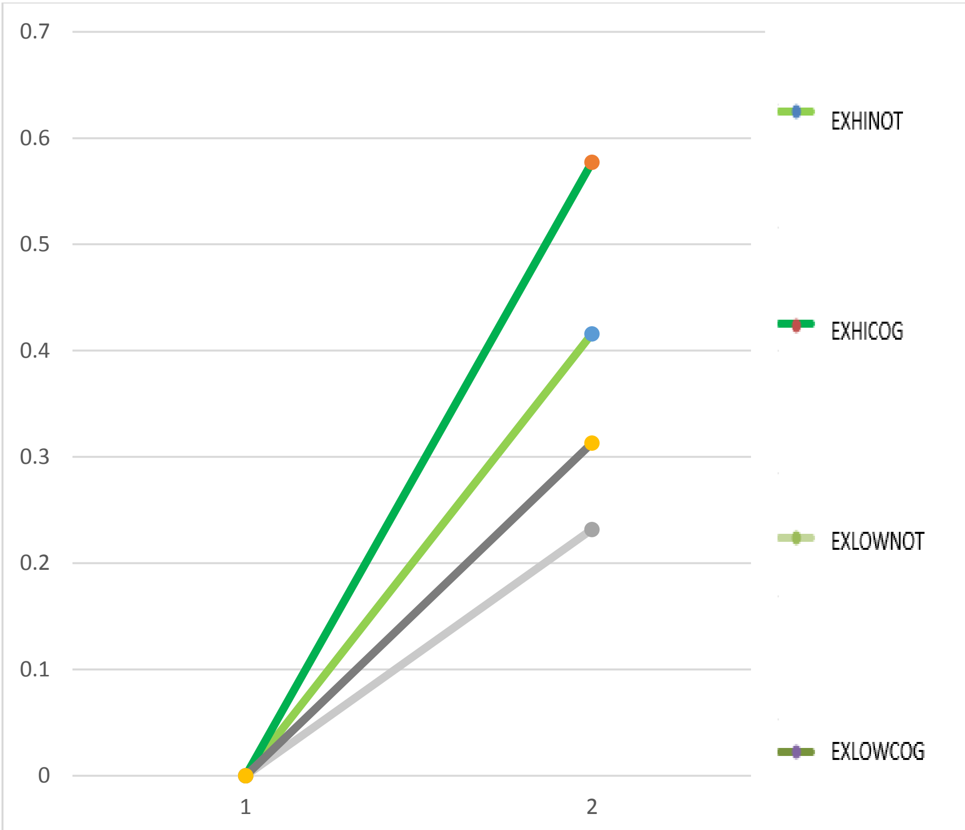


Figure 24. Wave 1 to Wave 2 growth of cognates versus non-cognates at exposed.

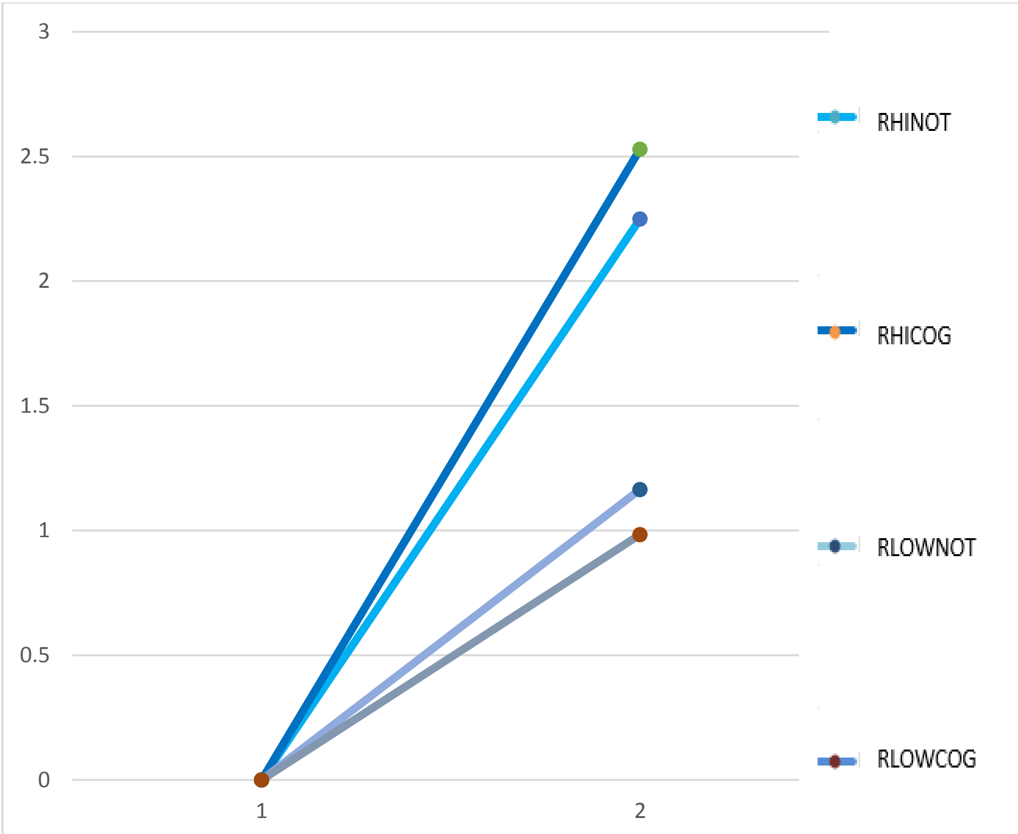


Figure 25. Wave 1 to Wave 2 growth of cognates and noncognates at REINFORCEMENT.

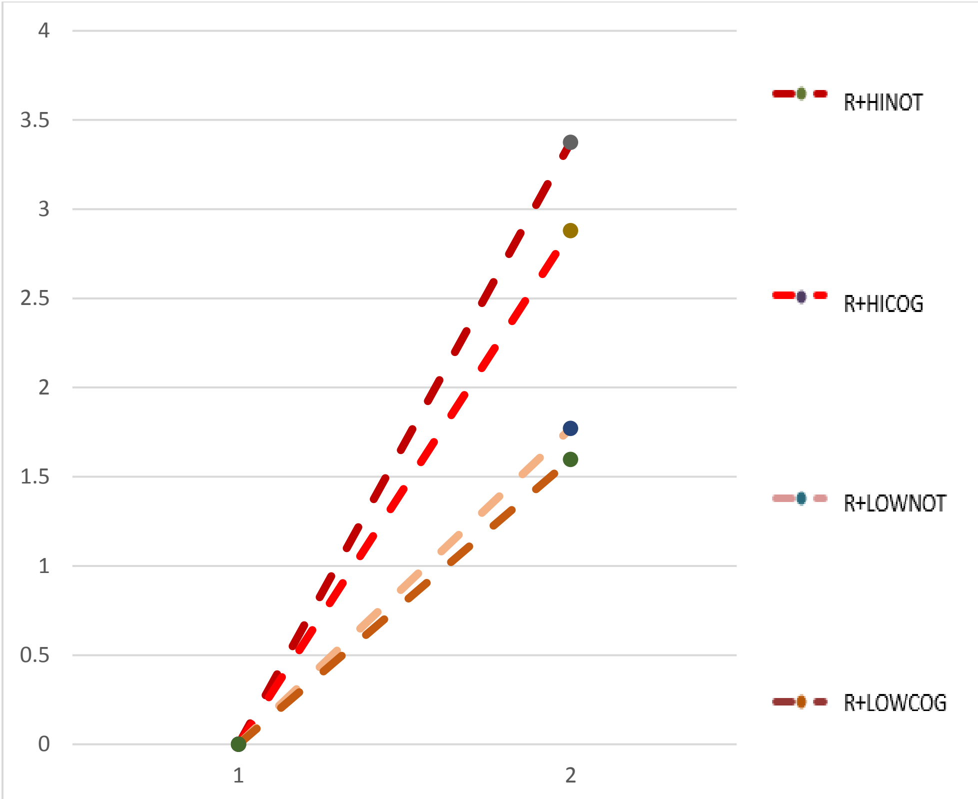


Figure 26. Wave 1 to Wave 2 growth of cognates versus non-cognates at REINFORCEMENT PLUS.

In order to assess the relationship between TOLD Oral Vocabulary scores and Aprenda total scale scores I conducted a Pearson correlation. The correlation was small (r=.25915; p>.05), indicating that while some of the students scored similarly on the TOLD and Aprenda there was little overlap between the two scores.

Model Assumptions

I explored the assumptions of linearity, normality, and homoscedasticity for each model presented in this chapter. I also explored whether the data distribution is conditional on the random effects.

With the binary logistic model, the assumptions of linearity and normality do not apply because the outcome is binary and the model is logistic. A compound metric covariance structure is not imposed on binary data. Accordingly, the assumption of the probability of the event occurring, Y=1, is the case that an event is a correct answer for this model. For binary data the factor level one of the dependent variable is the outcome with independent variables linearly related to log odds (logit) and continuous variables. Therefore, with a categorical outcome this assumption does not apply.

Because the design of this research employs a within-subjects methodology, the assumptions of homoscedasticity and independence do not apply. Homoscedasticity does not apply to this model because it assumes that the correlation between the repeated measures and the magnitude of correlation is equal across the entire correlation matrix. For all relevant predictors included in the model, there are not variables that would increase the effect size of a given predictor. The assumption of independence does not apply because this study employs a within-subjects design with items repeated in students and in students across time. However, the analytic models include random effects that account for the dependency within students and for the nested structure.

The three models presented in this chapter have little residual. According to SAS Institute (2015), the generalized chi-square statistic measures the residual sum of squares of the model and is quadratic from the marginal residuals. It, therefore, accounts for correlations among the data into account. Accordingly, the ratio of that statistic with its degrees of freedom indicates the degree of variability in the observations of the model mean. As such, the fit statistics reported with each model

(Table 5, 10, 14) reveal that there is little residual related to under or over dispersion for each model and that each one of the models fit well.

Chapter 5: Discussion

In the AVE3B substudy, I investigated the differential effects of instructional reinforcement on vocabulary learning and acquisition in the AVE program considering the influence of type of word and student-level factors on the learning and acquisition of the words in the AVE program. The AVE3B study included three conditions. In the first condition, words were embedded in the AVE curriculum with no instruction (exposed only condition). Another set of words were instructed using the robust and extended vocabulary instructional methods and reinforced six times (reinforcement condition). A third set of words were instructed using the same techniques as the second set and reinforced fourteen times (reinforcement plus condition). In this chapter, I describe considerations for the results presented in Chapter 4.

Type of Instruction

The instructional conditions garnered significant effects in this study. At wave two (posttest) there was a 16% change gain in the percent correct scores of exposed only words, while gains in the reinforcement condition represented a 70% growth change in the percent correct scores. The percent correct scores in the reinforcement plus condition changed by 87%. At wave two (posttest), the instructed conditions were highly effective compared to the no instructed exposure condition (d=.64). The reinforcement plus condition saw the most gains compared to the exposure condition with an additional small effect over the reinforcement condition (d=.24). As such, all students benefitted from the multimodal instruction and reinforcement in their core, Tier I, instruction.

The level of growth of the curricular words instructed through the robust, multimodal methods that characterize the instructed conditions in the AVE program is well-founded in the research across general and special education (i.e., Beck & McKeown, 2007; Coyne et al., 2007) and studies centered on ELLs (i.e., Baker et al., 2014; Gersten et al., 2007). The level of effects of the vocabulary instruction is similar to the growth that is reported across the studies summarized in Table 2 of Chapter 2. However, an important consideration with robust vocabulary instruction relates to the time factor of teaching words at the instructional intensity required for robust instruction. Therefore, the number of words that can be targeted through a robust level of extended vocabulary instruction is a limitation of that instructional approach.

This study taught twelve words per unit, similar to prior vocabulary research studies (see Table 2 and Baker et al., 2014). Using the figures from this study, if teachers were to implement the AVE program through the duration of a typical ten-month school year and target the same number words, with the same degree of retention rate as seen in this study, students will learn ~188 words, or 6% of the 3,000 words estimated necessary to keep pace with grade-level text. Therefore, it is critical for teachers to also promote conditions optimal for acquiring words through exposure, such as immersing students in a language-rich environment and providing instruction that helps students become better independent word learners, such as the methods of word learning strategy instruction and fostering students' word consciousness that are described in Chapter 2.

In addition to immersing students in language rich environments, instruction in word learning strategies, and promotion of word consciousness, robust extended vocabulary instruction should be carried out on the words that are very high utility for the students' academic learning, words likely not known by the students, and words that require the intensity level of robust instruction. To that end, examining effects of the instruction by word-level attributes is important. This study explored the word properties of conceptual complexity (i.e., concreteness and abstractness) and cognate status.

Word-type on Instruction

Across all conditions and across the entire sample, differences in type of word by conceptual complexity were minimal but present at pretest (*d*=.08). Concrete versus abstract words presented with even less difference in the posttest at the instructed conditions (*d*=.03). Interestingly, concrete words were more readily acquired compared to abstract words in the exposed condition (*d*=.17). There was a 6% change in the percent correct scores of the concrete words from wave 1 to wave 2, compared to the abstract words. With instruction intervening, it appears students in this sample learned the concrete and abstract words. However, with the natural process of word acquisition, as indicated by the exposed condition, students more readily acquired the concrete words, by extremely slight margins.

More pronounced differences in cognate status, the cognate versus noncognate words, were found. Cognates were more known at pretest, compared to noncognates (d=.24), with small but significant effects. Across the entire sample, students knew 27% more cognates compared to noncognates at pretest. In the instructed conditions by the posttest the small difference in knowledge between cognates and noncognates dissipated (d=.05), representing a difference of only 2.5%. However, the difference between cognates and noncognates in the exposure condition was more pronounced at posttest (d=.36). In the exposure condition, students were more likely to know the cognates at posttest with a difference of 43% between the percent correct scores for cognates and noncognates.

Though posttest scores in the exposure condition remained significantly lower compared to the instructed conditions, without the intervention of instruction students across this entire sample more readily acquired the cognates. However, the cognate "advantage" is not as evident in either condition that included instruction. Perhaps the cognate advantage lessened in the instruction conditions because the robust vocabulary instruction intervened and students made similar gains on the cognates and noncognates alike. This finding indicates that when there is no scaffolding of instruction students in this sample, second graders in a transitional bilingual program, likely bootstrap onto their cognate awareness to a significant degree.

With this sample, it would appear that the differences between the concrete and abstract words were far less pronounced compared to cognate status, but extremely small differences were present nonetheless. Exploring word attributes as influencers on word learning in order to inform strategic word selection in vocabulary instruction is an important area that future research should explore in more depth, especially when one considers the time of instruction issue many prior researchers have raised with robust instruction (Graves et al., 2013).

Impact of Generalized English Vocabulary on Exposure and Learning

This study used the TOLD-P:4 Oral Vocabulary subtest as a generalized metric of vocabulary knowledge. Using a split half procedure, the top half of the TOLD Oral Vocabulary scores were placed in the category HIGH and the bottom half of the scores in the category LOW. As such, the outcomes by TOLD scores were examined with respect to student scores relative to other students in this sample and not an external benchmark.

The students in the HIGH TOLD category presented with an advantage at pretest compared to the students in the LOW TOLD category, with small effects (d=.22). The students in the HIGH TOLD group answered 26% more of the items correct compared to their peers in the LOW TOLD category.

The difference between the two groups, HIGH TOLD and LOW TOLD, persisted with instruction. At WAVE 2 (posttest) in the REINFORCEMENT PLUS condition students in the HIGH TOLD category showed a significant advantage for learned words compared to those in the LOW TOLD group, with moderate effects (d=.45). Students in the HIGH TOLD group answered 21% more of the items correct at WAVE 2 (posttest) in the instructed conditions compared to their LOW TOLD peers. Similarly, in the exposed condition the HIGH TOLD group acquired significantly more words compared to the LOW TOLD group (d=.33), answering

34% more of the items correct. It appears that students in the HIGH TOLD category may have hit the ceiling on the AVEVA assessment within the reinforcement plus condition.

When comparing the students in the HIGH TOLD category in the REINFORCEMENT condition with the students in the LOW TOLD category in the REINFORCEMENT PLUS condition, the disparity between the two groups starts to fade. The difference between the two groups moves from moderate (*d*=.45) to small effects (*d*=.21), which is only an 11% difference. It appears that students in the LOW TOLD category benefitted from the additional reinforcement, allowing them to approach the levels that students in the HIGH TOLD category garnered with the lesser reinforced condition. However, students in the HIGH TOLD category essentially hit ceiling in the most intensive instructional condition; therefore, they may be served well enough through the less intensive instructional condition.

Results in Relation to the Matthew Effect

Students in the HIGH TOLD category more readily learned the instructed vocabulary words and more easily acquired the exposed words, compared to their peers in the LOW TOLD category. This finding empirically illustrates the Matthew Effect for both acquisition and learning. Stanovich (1986) first described the Matthew Effect in relation to the natural acquisition of words in conditions similar to this study's exposed condition. In this study, however, there is evidence for the Matthew Effect not only for acquisition in the exposed condition but there is also a Mathew Effect with respect to the learning in the instructed conditions. This finding lends support to the need to intervene in the early elementary years and before elementary in preschool in order to promote ELL English vocabulary learning for the benefit of reading comprehension by the middle elementary school grades.

In their study with lower elementary-aged students, Penno et al. (2002) also found a Matthew Effect with vocabulary learning. Using composite scores from classroom-based literacy assessments, Penno et al. (2002) determined composite ability scores and in doing so were able to investigate ability level by outcomes. Findings in Penno et al. (2002) reveal an explanation by ability interaction effect where students with higher ability levels learned more words that were explained through embedded means. However, there was no effect for ability by the frequency the word was encountered through the repeated readings in Penno et al. (2002). The authors conclude that increased encounters with the word or embedded explanations did not overcome the Matthew Effect. They surmise that additional strategies are needed to support vocabulary learning and acquisition of struggling learners.

However, in this study the Matthew Effect started to dissipate when comparing the HIGH TOLD group in REINFORCEMENT with the LOW TOLD group in REINFORCEMENT PLUS. In this study, word encounters were interwoven in a systematic way throughout the AVE program curriculum; therefore, students encountered the words through an array of engaging activities using all four language modalities. The curricular activities in this study drew heavily on visual cues, language-based scaffolds, songs, and games across the four language modalities (see Appendix A for curricular overview). According to Peercy et al. (2015) reading, listening, speaking, and writing activities are vital when considering ELL immersion in language-rich environments. Each language modality is characterized by different linguistic structures and may require different skills. In this study, it appears that students presenting with more limited generalized English vocabulary knowledge, students in the LOW TOLD category, benefited from the additional encounters with the words reinforced through all language modalities.

Potential Instructional Implications for Tiered Instructional Frameworks

Scores related to the students in the HIGH TOLD category indicate that the level of reinforcement in the REINFORCED condition was sufficient for adequate word learning, but that the students in the lower TOLD category significantly benefitted from the added reinforcement in the REINFORCEMENT PLUS condition. The finding in this study with respect to the TOLD categories also has potential implications for overcoming the vocabulary Matthew Effect through vocabulary instruction as part of tiered instructional frameworks, such as Responsive to Intervention (RTI) and Multi-Tiered Systems of Supports (MTSS). It points to the likelihood that students who present as struggling readers with limited vocabularies who may be receiving Tier 2 intervention may also benefit from the reinforcement of the words encountered and instructed in core instruction.

Coyne, Capozzoli, Ware, and Loftus (2010) and Pullen et al. (2010) present a case for the facilitative effect of targeting explicit vocabulary and word learning skills in small group settings with struggling learners, such as Tier 2 intervention situations. Authors claim that targeting vocabulary instruction in intervention settings provide benefit for sublexical, lexical, and supralexical skills. Pullen, Tuckwillar, Konald, Maynard, and Coyne (2010) executed a tiered vocabulary intervention with struggling ELL and non-ELL readers in early elementary school. In their study, Tier 2 instruction was implemented in small groups of four to five students for 20 minutes a day. The instruction consisted of reviewing target word meanings encountered in core instruction, and providing additional multimodal deep processing activities characteristic of robust vocabulary instruction. When comparing the intervention group with a no-intervention control group, the researchers found initial significant effects for the group receiving the vocabulary intervention with respect to both word learning and text-level skills. In their study, however, the effects dissipated in the delayed posttest, pointing to the need for sustained vocabulary instruction in intervention settings.

Although reinforcing words using multidimensional methods is helpful in core instruction and with intervention settings, there is also a need to consider ways to sustain that instruction over the long term. However, there may be sparsity with respect to interventions and vocabulary-related outcomes. A recent large synthesis found paucity with respect to vocabulary-related outcomes for Tier 2 interventions in early elementary. In a 2017 policy report, Gersten, Newman-Gonchar, Haymond, and Dimino (2017) uncovered only one to two Tier 2 interventions in first and second grade that reported vocabulary-related outcomes.

Given the findings presented in this dissertation and the promise shown by prior research with exploring vocabulary instruction in intervention settings, future research should investigate vocabulary instruction in the intervention settings for ELLs who present with vocabulary limitations and reading challenges. Future work could investigate more systematically the word learning of ELLs in Tier 2 intervention instruction and the benefit with augmenting extant Tier 2 interventions with vocabulary instruction through extended and embedded instructional methods for reinforcing the words from the core instruction.

Impact of Spanish Language Proficiency on English Vocabulary Learning

This study used the Aprenda 3 as a generalized metric of Spanish language proficiency. Using a split half procedure, the top half of the Aprenda scores were placed in the category HIGH and the bottom half of the scores in the category LOW. Similarly to the TOLD, the outcomes by Aprenda scores were examined with respect to student scores relative to other students in this sample.

In this sample, students in the HIGH APRENDA group answered 36% more items correctly for the cognates compared to noncognates. Students in the LOW APRENDA group also showed an advantage with cognates at pretest, answering 21% more of the cognate items correct, compared to noncognate items. The HIGH APRENDA group knew 33% more of the cognate items than the students in the LOW APRENDA category. Effect sizes for cognates versus noncognates at pretest was moderate (d=.3) for the HIGH APRENDA group and small for the LOW APRENDA group (d=.11).

Differences in learning the cognates versus noncognates in the instructed conditions for the HIGH APRENDA group was minimal (d=.11). A potential ceiling effect was present with respect to the cognates for students in the HIGH APRENDA category in the REINFORCEMENT and REINFORCEMENT PLUS condition. The simple means in Table 13 show that the means for that group are very high and standard deviations low. This ceiling effect perhaps explains why less growth was seen in the instructed conditions for that group and thereby less of a differential effect was uncovered with respect to cognates and noncognates. Perhaps without a ceiling effect, the HIGH APRENDA group would present with a more significant cognate advantage in the instructed conditions at posttest. For that HIGH APRENDA group there were moderate differences in cognate acquisition at the exposed condition, compared to non-cognates (d=.5). In the exposed condition, the students in the HIGH APRENDA category answered 47% more of the cognate items correctly, compared to the noncognates.

Those same patterns are evident in the LOW APRENDA group, but to a lesser magnitude. At posttest in the LOW APRENDA group, differences in the instructed conditions for cognates versus noncognates were almost nonexistent (d=.02), representing only an additional percent advantage in the cognate items. However, there were differential effects approaching moderate with respect to the exposed condition (d=.32), a 41% advantage for the cognates compared to the noncognates in the LOW APRENDA group.

Students in the HIGH APRENDA category did better across the board on all the words, and the degree that they acquired the cognates was 36% more compared to their peers in the LOW APRENDA category. These findings indicate that all students, irrespective of APRENDA category, are likely bootstrapping onto their cognate knowledge as seen through the findings in the exposure condition.

Even though there is a significant cognate advantage for both groups in the exposure condition, it appears that as intensity of instruction increased, the "cognate advantage" seen in the exposed condition dissipated. With the instruction coming in to play, students are learning noncognates to a similar degree as the cognates, with students in the HIGH APRENDA category presenting with substantially greater gains across cognates and noncognates. However, both groups of students were differentially acquiring the cognates in the exposed only condition compared to noncognates. Students in the HIGH APRENDA group acquired more than 1/3 of the cognates compared to their peers in the LOW APRENDA category. This corroborates prior findings from studies such as Deacon (2017), Dressler (2000), and Genesee and Dressler (2006). These studies show stronger first language proficiency aids the natural acquisition of cognates.

The Aprenda 3 finding in this study also substantiates prior research that suggests that cognate knowledge impacts English word learning (Carlo, et al., 2004) and that bilingual learners more readily acquire the cognates (Nation, 2001). Given the limitation that robust vocabulary instruction cannot target large swaths of words due to time constraints, the bulk of vocabulary learning for students will take place through conditions like the exposed condition in this study. As such, it is imperative for teachers to foster bilingual learners' cognate awareness. This study shows the degree to which students in the exposed condition more readily acquired the cognates, compared to noncognates.

Several studies have investigated the outcomes of cognate awareness instruction with bilingual learners and ELL students (i.e., August, et al., 2015; August et al., 2018; Carlo et al., 2004; Silverman et al., 2017; White & Horst, 2012). These studies documented that the attribute of first language proficiency is a facilitative feature for lexical acquisition of the second language. Accordingly, instruction should focus, in part, on helping bilingual students hone their cognate awareness skills through employing instructional practices such as those suggested in August et al. (2018), Carlo et al. (2004) and White and Horst (2012). These pedagogical practices include contrastive analysis activities, textual cognate searches, and explicit statements when targeted instructional words are cognates.

Scope and Limitations of the Study

I believe that this study is an important first step in the process of understanding the interactive word learning burden and the complexities of the interactions between level of instructional reinforcement, type of word, and student factors in second grade ELL students in transitional bilingual classrooms. However, this study does have several important limitations for which I offer suggestions for future work.

This study is generalizable to other second grade transitional bilingual programs where the language of instruction is Spanish. To increase generalizability of the study, this work should be replicated and executed with other populations, other regions of the U.S., and other program models, such as long-term developmental bilingual programs and English immersion programs. Effects related to type of word and instruction could be explored with a larger sample in future work. Because potential ceiling effects were found with respect to the cognates in the HIGH APRENDA category, future research should continue to investigate this topic.

Additionally, the study took place wholly within the context of fictional texts. To more closely align the pedagogical approach used in this research study with respect to current considerations within CCR standards, a mix of informational and narrative texts could be used for the shared interactive reading texts. Future research should explore the techniques presented in this study within a program that also draws on informational texts or a mix of both informational and narrative texts. For example, the “twin texts,” strategy is where a narrative text and informational text, aligned on a topic (i.e., *Stellaluna* by Janelle Cannon and *Bats!* by Scholastic), are used together in a thematic unit. This technique offers a promising approach in further studying some of the ideas presented in this research (see Camp, 2000 for a review of twin texts). Perhaps the narrative text provides some of the background to help students more fully engage with the informational text.

Although the actual classroom teachers implemented the instruction in the AVE program, the students in this study are all native Spanish speakers. However, this may impact the social validity of the work because ELL students in the U.S. are a heterogeneous population. Teachers in U.S. schools are often contending with more than one language background in their classrooms, and that reality will impact the way in which a teacher makes use of cognate awareness and cognate knowledge in vocabulary instruction. Future work could explore this study with other student populations to increase its generalizability.

This study does not include a between-subject's business-as-usual control group. Because the study was a within-subjects design it is not possible to examine the effects of the AVE program with respect to generalized literacy outcomes on standardized measures. However, the purpose of this proposed work is to better understand the interactions between condition, type of word, and student factors as a proof-of-concept study. Future work should examine this instructional intervention using between-subjects methods in order to assess general literacy gains.

Since an expressive measure of vocabulary knowledge was not administered with this intervention, depth of word knowledge is not able to be investigated with this work. Future work should extend this investigation by looking at various facets of word knowledge and investigate depth of word knowledge outcomes in relation to word-type, instructional methods, and learner variables. Similarly, this study did not explore the impact of vocabulary instruction with reading comprehension metrics. Future work should explore this instruction in connection to the impact the supra-lexical-level literacy skills.

Future research can investigate the word properties using more complex methods. Although this study explores the effect of two word properties—dichotomized by abstractness and cognate status—the study does not examine the influence of other word properties on the student outcomes. There are several other word property factors, such as polysemy and valence, that other research, primarily research in psycholinguistics, indicates might impact word acquisition and learning and therefore may be worth exploring.

Drawing on the work in psycholinguistics, future work could also investigate the two word-types presented in this dissertation in a more nuanced fashion. For example, future research might consider degrees of abstractness, concreteness, and, for cognate status, orthographic and phonological overlap when exploring the word properties of abstractness and cognate status.

Finally, with respect to student-level factors, this study employed a simplistic approach to categorizing the students by English vocabulary and Spanish language proficiency. I took a split half of the groups, where I found the median score and the students who scored above the median were in the top half of the group and those who scores below the median in the bottom half of the group. Future work could explore the student-level factors using other more complex procedures. Following the methods in Penno et al. (2002), research teams could use school-based metrics to determine composite ability scores. Other studies have grouped scores into three groups and compared the top and bottom third to the middle.

Although there are a number of limitations with this study, as an initial investigation centered on the interactive word learning burden, this examination fulfills the goal of a proof-of-concept exploration. This research study not only integrates two word learning theories into one theoretical model, the interactive word learning burden, it also sheds initial empirical light on the nature of word learning with respect to level of instructional reinforcement and the influencers of word and student factors.

Appendices

Appendix A: Curriculum Chart

If the activity is noted as the base instruction then words in the reinforcement and reinforcement plus condition were equally targeted. If the activity is not noted as base instruction then the number of encounters with the target word applied by condition.

	Activity	**Purpose**
Prereading	Word Flash Game	Reviews previous day's vocabulary through Active Response Techniques
	Vocabulary Instruction with Picture Card (base instruction)	Initial instruction of two to three target words for extended instruction
Interactive Reading on Rug	Picture Walk Review of Book (base instruction)	Reviews the reading selection from prior day
	Interactive Reading (base instruction)	Reading and discussion conducted in intervals with reinforcement for word consciousness
	Interactive Reading Closure (base instruction)	Partner discussion about a prompt related to the text
Postreading	Vocabulary Game 1	Reinforces target vocabulary
	Vocabulary Game 2	Reinforces target vocabulary
	Vocabulary Song 1	Song connected to unit theme with the target vocabulary words
	Vocabulary Song 2	Song connected to unit theme with the target vocabulary words

	Glossary Work with Vocabulary Words	Reinforces target vocabulary words with writing and illustrating
	Cognate Treasure Hunt	Helps students practice finding cognates in text
	Glossary Work with Cognates	Activity developed for cognate awareness
	Scaffolded Writing with Vocabulary	Cloze passage with the target vocabulary.
	Progress Monitoring Matching Assessment	Matching assessment with the target vocabulary at the end of the unit.
	Vocabulary Review Game	Reviews prior units' words in the context of the current unit theme. For unit 1, a cognate awareness activity replaced this activity.
	Review Writing Activity	Reviews prior units' words in the context of the current unit theme. For unit 1 a cognate awareness activity replaced this.

Appendix B: Intervention Words by Book

Examples of Reinforcement Plus Words		Examples of Reinforcement Words	
Abstract Noncognates Remarkable (Unit 1) Dreadful (Unit 1) Wealth (Unit 2) Throughout (Unit 3) Daring (Unit 3) Standard (Unit 4) Aware (Unit 5) Pride (Unit 5) Proper (Unit 6) Keen (Unit 7) Fairly (Unit 7) Profit (Unit 8)	**Abstract Cognates** Confidence (Unit 1) Tremendous (Unit 2) Affects (Unit 2) Concern (Unit 3) Attitude (Unit 4) Ideal (Unit 4) Enthusiasm (Unit 5) Survived (Unit 6) Complicated (Unit 6) Impression (Unit 7) Considering (Unit 8) Satisfaction (Unit 8)	**Abstract Noncognates** Mood (Unit 1) Realize (Unit 2) Threat (Unit 2) Overcome (Unit 3) Quality (Unit 4) Suitable (Unit 4) Scarcely (Unit 5) Judgment (Unit 6) Lately (Unit 6) Actually (Unit 7) Successful (Unit 8) Poisonous (Unit 8)	**Abstract Cognates** Especially (Unit 1) Atmosphere (Unit 1) Treatment (Unit 2) Common (Unit 3) Informed (Unit 3) Severe (Unit 4) Preferred (Unit 5) Extremely (Unit 5) Opinions (Unit 6) Particularly (Unit 7) Responsibility (Unit 7) Predict (Unit 8)
Concrete Noncognates Pace (Unit 1) Environment (Unit 2) Motionless (Unit 2) Opposite (Unit 3) Boundary (Unit 4) Display (Unit 4) Relief (Unit 5) Clumsy (Unit 6) Increases (Unit 6) Rage (Unit 7) Wound (Unit 8) Fierce (Unit 8)	**Concrete Cognates** Reaction (Unit 1) Producing (Unit 1) Calm (Unit 2) Assembly (Unit 3) Previous (Unit 3) Various (Unit 4) Event (Unit 5) Reduced (Unit 5) Delicate (Unit 6) Mysterious (Unit 7) Occurs (Unit 7) Singular (Unit 8)	**Concrete Noncognates** Hastily (Unit 1) Dusk (Unit 1) Weary (Unit 2) Household (Unit 3) Sturdy (Unit 3) Corridor (Unit 4) Sigh (Unit 5) Ease (Unit 5) Clung (Unit 6) Illness (Unit 7) Lack (Unit 7) Ridge (Unit 8)	**Concrete Cognates** Impatiently (Unit 1) Central (Unit 2) Exposed (Unit 2) Quantity (Unit 3) Presence (Unit 4) Evidence (Unit 4) Applied (Unit 5) Graceful (Unit 6) Interior (Unit 6) Observed (Unit 7) Obtained (Unit 8) Orderly (Unit 8)

Examples of Exposure Words	
Abstract Noncognates Sake (Unit 1) Attempt (Unit 1) Wonders (Unit 2) Pity (Unit 3) Bound (Unit 3) Treasures (Unit 4) Deadly (Unit 5) Truly (Unit 5) Weird (Unit 6) Likely (Unit 7) Seems (Unit 7) Thoroughly (Unit 8)	**Abstract Cognates** Demand (Unit 1) Native (Unit 2) Rare (Unit 2) Mission (Unit 3) National (Unit 4) Custom (Unit 4) Assured (Unit 5) Opportunity (Unit 6) Confusion (Unit 6) Conscious (Unit 7) Current (Unit 8) Expedition (Unit 8)
Concrete Noncognates Glimpse (Unit 1) Cast (Unit 2) Landscape (Unit 2) Gear (Unit 3) Link (Unit 4) Fled (Unit 4) Sharply (Unit 5) Prey (Unit 6) Harsh (Unit 6) Gathering (Unit 7) Layers (Unit 8) Request (Unit 8)	**Concrete Cognates** Portion (Unit 1) Occasion (Unit 1) Magnificent (Unit 2) Ruins (Unit 3) Entry (Unit 3) Respond (Unit 4) Absolutely (Unit 5) Directed (Unit 5) Rescue (Unit 6) Expand (Unit 7) Horror (Unit 7) Territory (Unit 8)

Bibliography

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

Arteagoitia I. & Howard E. R. (2015). The role of native language in the literacy development of Latino students in the United States. In J. Cenoz & D. Gorter (Eds.), *Multilingual Education: Between Language Learning and Translanguaging*. Cambridge, UK: Cambridge University Press.

Artzi, L., Uccelli, P., & Massoud, L. (2015, April). *The Connectives Intervention.* Paper presented at the 2015 annual meeting of AERA, Chicago, IL. Abstract retrieved from www.aera.org.

Artzi, L., August, D., & Barr, C. (2014, April). *Investigating long term effects of a supplemental vocabulary intervention.* Paper presented at the 2014 annual meeting of AERA, Philadelphia, PA. Abstract retrieved from www.aera.org.

Artzi, L., Massoud, L., August, D., & Barr, C. (2012, April). *Assessing and instructing English vocabulary in Spanish-speaking English language learners.* Paper presented at the 2012 Language Testing Research Colloquium (LTRC), Princeton, NJ. Abstract retrieved from www.cal.org.

August, D. & Barr, C. (2011). *Developing literacy in Spanish-speaking children: Acquisition of vocabulary in English* (Technical Report 2 Submitted to the

National Institute of Child Health and Human Development). Washington, DC: Center for Applied Linguistics.

August, D. & Barr, C. (2010). *Developing literacy in Spanish-speaking children: Acquisition of vocabulary in English* (Technical Report 2 Submitted to the National Institute of Child Health and Human Development). Washington, DC: Center for Applied Linguistics.

August, D., Artzi, L., & Barr, C. (2013, April). *The Acquisition of Vocabulary in English Project*. Paper presented at the 2013 annual meeting of AERA, Vancouver, CA. Abstract retrieved from aera.org.

August, D., McCardle, P., & Shanahan, T. (2014). Developing literacy in English language learners: Findings from a review of the experimental research. *School Psychology Review*, *43*(4), 490-498.

August, D., & Shanahan, T. (2006). *Executive summary*: *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth*. Mahwah, NJ: Lawrence Erlbaum Associates.

August, D., & Shanahan, T. (2009). English language learners: Developing literacy in second-language learners—Report of the National Literacy Panel on Language-Minority Children and Youth. *Journal of Literacy Research*, *41*(4), 432-452.

August, D., Artzi, L., & Barr, C. (2016). Helping ELLs meet standards in English language arts and science: An intervention focused on academic vocabulary. *Reading & Writing Quarterly*, *32*(4), 373-396.

August, D., Artzi, L., Barr, C., & Francis, D. (2018). The moderating influence of instructional intensity and word type on the acquisition of academic vocabulary in young English language learners. *Reading and Writing*, 31(4), 965-989.

August, D., Branum-Martin, L., Cardenas-Hagan, E., & Francis, D. J. (2009). The impact of an instructional intervention on the science and language learning of middle grade English language learners. *Journal of Research on Educational Effectiveness*, *2*(4), 345-376.

August, D., Branum-Martin, L., Cárdenas-Hagan, E., Francis, D. J., Powell, J., Moore, S., & Haynes, E. F. (2014). Helping ELLs meet the Common Core State Standards for literacy in science: The impact of an instructional intervention focused on academic language. *Journal of Research on Educational Effectiveness*, *7*(1), 54-82.

August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research & Practice*, *20*(1), 50-57.

August, D., McCardle, P., & Shanahan, T. (2014). Developing literacy in English language learners: Findings from a review of the experimental research. *School Psychology Review*, *43*(4), 490-498.

Bailey, A. L. (2007). *The language demands of school: Putting academic English to the test*. New Haven CT: Yale University Press.

Baker, S., Lesaux, N., Jayanthi, M., Dimino, J., Proctor, C. P., Morris, J., Gersten, R., Haymond, K., Kieffer, M. J., Linan-Thompson, S., & Newman-Gonchar, R. (2014). *Teaching academic content and literacy to English learners in elementary and middle school* (NCEE 2014-4012). Washington, DC: National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education. Retrieved from the NCEE website: http://ies.ed.gov/ncee/wwc/publications_reviews.aspx.

Batalova, J., & McHugh, M. (2010). *DREAM vs. reality: An analysis of potential DREAM Act beneficiaries*. Migration Policy Institute, National Center on Immigrant Integration Policy.

Baumann, J. F., Edwards, E. C., Boland, E. M., Olejnik, S., & Kame'enui, E. J. (2003). Vocabulary tricks: Effects of instruction in morphology and context on fifth-grade students' ability to derive and infer word meanings. *American Educational Research Journal*, *40*(2), 447-494.

Baumann, J. F., Kame'enui, E. J., & Ash, G. E. (2003). Research on vocabulary instruction: Voltaire redux. *Handbook of research on teaching the English language arts*, *2*, 752-785.

Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction*. New York, NY: Guilford Press.

Beck, I. L., & McKeown, M. G. (2007). Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *The Elementary School Journal*, *107*(3), 251-271.

Beck, I. L., McKeown, M. G., & McCaslin, E. S. (1983). Vocabulary development: All contexts are not created equal. *The Elementary School Journal*, *83*(3), 177-181.

Beck, I. L., Omanson, R. C., & McKeown, M. G. (1982). An instructional redesign of reading lessons: Effects on comprehension. *Reading Research Quarterly, 17,* 462-481.

Beck, I., & McKeown, M. (1991). Conditions of vocabulary acquisition. In R. Barr, M. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research*, Vol. 2. (pp. 789-814). New York, NY: Psychology Press.

Beck, Isabel L., McKeown, Margaret G., & Omanson, Richard C. (1987) The effects and uses of diverse vocabulary instructional techniques. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition*. (pp. 147-163). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Beck, I. L., Perfetti, C. A., & McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology*, *74*(4), 506.

Beck, I. L., McKeown, M. G., Omanson, R. C., & Pople, M. T. (1984). Improving the comprehensibility of stories: The effects of revisions that improve coherence. *Reading Research Quarterly*, 263-277.

Bialystok, E. (2007). Cognitive effects of bilingualism: How linguistic experience leads to cognitive change. *International Journal of Bilingual Education and Bilingualism*, *10*(3), 210-223.

Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology*, *98*(1), 44.

Biemiller, A., & Slonim, N. (2001). Estimating root word vocabulary growth in normative and advantaged populations: Evidence for a common sequence of vocabulary acquisition. *Journal of Educational Psychology*, *93*(3), 498.

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers*, *33*(1), 73-79.

Blachowicz, C. L., Fisher, P. J., Ogle, D., & Watts-Taffe, S. (2006). Vocabulary: Questions from the classroom. *Reading Research Quarterly*, *41*(4), 524-539.

Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The effects of morphological instruction on literacy skills a systematic review of the literature. *Review of Educational Research*, *80*(2), 144-179.

Bryant, D. P., Goodwin, M., Bryant, B. R., & Higgins, K. (2003). Vocabulary instruction for students with learning disabilities: A review of the research. *Learning Disability Quarterly*, *26*(2), 117-128.

Brysbaert, M., New, B., & Keuleers, E. (2012). Adding part-of-speech information to the SUBTLEX-US word frequencies. *Behavior Research Methods*, *44*(4), 991-997.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904-911.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977-990.

Burns, D. (2011). *Examining the effect of an overt transition intervention on the reading development of at-risk English language learners in first grade* (Unpublished doctoral dissertation). University of Oregon, Oregon.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, *96*(1), 31.

Camp, D. (2000). It takes two: Teaching with twin texts of fact and fiction. *The Reading Teacher*, *53*(5), 400-408.

Calderon M, August D, Slavin R, Duran D, Madden N, & Cheung A. (2005). Bringing words to life in classrooms with English-language learners. In: Hiebert EH, Kamil ML, editors. *Teaching and learning vocabulary: Bringing* *research to practice.* Lawrence Erlbaum Associates; Mahwah, NJ: 2005. pp. 115–136.

Carlo, M. S., August, D., McLaughlin, B., Snow, C. E., Dressler, C., Lippman, D. N.... & White, C. E. (2004). Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classrooms. *Reading Research Quarterly*, *39*(2), 188-215.

Carrier, S. J. (2013). Elementary preservice teachers' science vocabulary: Knowledge and application. *Journal of Science Teacher Education*, *24*(2), 405-425.

Catts, H. W., Adlof, S. M., & Ellis Weismer, S. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research*, *49*(2), 278-293.

Catts, H. W., Adlof, S. M., Hogan, T. P., & Ellis Weismer, S. (2005). Are specific language impairment and dyslexia distinct disorders? *Journal of Speech, Language, and Hearing Research*, *48*(6), 1378-1396.

Catts, H. W., Herrera, S., Nielsen, D. C., & Bridges, M. S. (2015). Early prediction of reading comprehension within the simple view framework. *Reading and Writing*, *28*(9), 1407-1425.

Center on RTI (2011). *Developing an RTI Guidance Document.* Retrieved from http://www.rti4success.org/resource/developing-rti-guidance-document.

Chall, J. S., & Jacobs, V. A. (2003). The classic study on poor children's fourth-grade slump. *American Educator*, *27*(1), 14-15.

Chambers, B., Cheung, A. C., Madden, N. A., Slavin, R. E., & Gifford, R. (2006). Achievement effects of embedded multimedia in a success for all reading program. *Journal of Educational Psychology*, *98*(1), 232.

Chambers, B., Cheung, A., Gifford, R., Madden, N., & Slavin, R. E. (2004). Achievement effects of embedded multimedia in a Success for All reading program. Baltimore, MD: Success for All Foundation.

Coady, J., Magoto, J., Hubbard, P., Graney, J., & Mokhtari, K. (1993). High frequency vocabulary and reading proficiency in ESL readers. In T. Huckin, M. Haynes, J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 217-228). Norwood, NJ: Ablex Publishing.

Coyne, M. D., McCoach, D. B., & Kapp, S. (2007). Vocabulary intervention for kindergarten students: Comparing extended instruction to embedded instruction and incidental exposure. *Learning Disability Quarterly*, *30*(2), 74-88.

Coyne, M. D., Capozzoli, A., Ware, S., & Loftus, S. (2010). Beyond RTI for decoding: Supporting early vocabulary development within a multitier approach to instruction and intervention. *Perspectives on Language and Literacy*, 36(2), 18.

Crevecoeur, Y. C., Coyne, M. D., & McCoach, D. B. (2014). English language learners and English-only learners' response to direct vocabulary instruction. *Reading & Writing Quarterly*, *30*(1), 51-78.

Crossley, S. A., Salsbury, T., & McNamara, D. S. (2011). Predicting the proficiency level of language learners using lexical indices. *Language Testing*, *28*(4), 561-580.

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

Cunningham, J. W., & Moore, D. W. (1993). The contribution of understanding academic vocabulary to answering comprehension questions. *Journal of Literacy Research*, *25*(2), 171-180.

Dale, E., & O'Rourke, J. (1981). *The living word vocabulary: A national vocabulary inventory*. Chicago, IL: World Book/Childcraft International.

Davies, M. (2008). *Corpus of Contemporary American English*. Retrieved from http://corpus.byu.edu/coca/.

Davies, M. (2012). *Corpus of American Soap Operas*. Retrieved from http://corpus2.byu.edu/soap/overview.asp.

D'Angelo, N., Hipfner-Boucher, K., & Chen, X. (2017). Predicting growth in English and French vocabulary: The facilitating effects of morphological and cognate awareness. Developmental psychology, 53(7), 1242.

De Boeck, P., & Wilson, M. (2004). A framework for item response models. *Explanatory item response models* (pp. 3-41). New York, NY: Springer.

De Groot, A., & Keijzer, R. (2000). What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning*, *50*(1), 1-56.

Dressler, C. A. (2000). *The word-inferencing strategies of bilingual and monolingual fifth graders: A case study approach* (Unpublished doctoral dissertation). Harvard Graduate School of Education, Cambridge, MA.

Durso, F. T., & Coggins, K. A. (1991). Organized instruction for the improvement of word knowledge skills. *Journal of Educational Psychology*, *83*(1), 108.

Duursma, E., Romero-Contreras, S., Szuber, A., Proctor, P., Snow, C., August, D., & Calderón, M. (2007). The role of home literacy and language environment on bilinguals' English and Spanish vocabulary development. *Applied Psycholinguistics*, *28*(01), 171-190.

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), 1-44.

Elley, W. B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly*, *24*(2) 174-187.

Ellis, N. C. (2008). The associative learning of constructions, learned attention, and the limited L2 endstate. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition*, 372-405. New York, NY: Routledge.

Ellis, N. C., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning*, *43*(4), 559-617.

Filippini, A. L. (2007). *Effects of a vocabulary-added instructional intervention for at-risk English learners: Is efficient reading instruction more effective?* (Unpublished doctoral dissertation). University of California, Santa Barbara, CA.

Filippini, A. L., Gerber, M. M., & Leafstedt, J. M. (2012). A vocabulary-added reading intervention for English learners at-risk of reading difficulties. *International Journal of Special Education*, *27*(3), 14-26.

Fowler, A. E. (1991). How early phonological development might set the stage for phoneme awareness. In S. A. Brady & D. P. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman*. (pp. 97-117). Mahwah, NJ: Lawrence Erlbaum Associated, Inc.

García, J. R., & Cain, K. (2013). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research*, 0034654313499616.

Gee, J. P. (2009). *New digital media and learning as an emerging area and "worked examples" as one way forward*. Cambridge, MA: MIT Press.

Genesee, F. (2006). *Educating English language learners: A synthesis of research evidence*. New York, NY: Cambridge University Press.

Genesee, F., Lindholm-Leary, K., Saunders, W., & Christian, D. (2006). Educating English language learners: A synthesis of empirical evidence. *Journal of Education for Students Placed at Risk, 10*(4), 363of Ed

Gersten, R., Baker, S. K., Shanahan, T., Linan-Thompson, S., Collins, P., & Scarcella, R. (2007). Effective Literacy and English Language Instruction for English Learners in the Elementary Grades. IES Practice Guide. NCEE 2007-4011. *What Works Clearinghouse*.

Gersten, R., Newman-Gonchar, R. A., Haymond, K. S., & Dimino, J. (2017). What is the evidence base to support reading interventions for improving student outcomes in grades 1–3? (REL 2017–271). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast. Retrieved from http://ies.ed.gov/ncee/edlabs.

Geva, E., & Genesee, F. (2006). First-language oral proficiency and second-language literacy. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth*, 185-196. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Goodwin, A. P., & Ahn, S. (2010). A meta-analysis of morphological interventions: Effects on literacy achievement of children with literacy difficulties. *Annals of Dyslexia*, *60*(2), 183-208.

Gottardo, A., & Mueller, J. (2009). Are first- and second-language factors related in predicting second-language reading comprehension? A study of Spanish-speaking children acquiring English as a second language from first to second grade. *Journal of Educational Psychology*, *101*(2), 330.

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

Graves, M. F., August, D., & Mancilla-Martinez, J. (2012). *Teaching vocabulary to English language learners*. New York, NY: Teachers College Press.

Graves, M. F., Baumann, J. F., Blachowicz, C. L., Manyak, P., Bates, A., Cieply, C.... & Von Gunten, H. (2014). Words, words everywhere, but which ones do we teach? *The Reading Teacher*, *67*(5), 333-346.

Gunn, B., Biglan, A., Smolkowski, K., & Ary, D. (2000). The efficacy of supplemental instruction in decoding skills for Hispanic and non-Hispanic students in early elementary school. *The Journal of Special Education*, *34*(2), 90-103.

Gunn, B., Smolkowski, K., Biglan, A., Black, C., & Blair, J. (2005). Fostering the development of reading skill through supplemental instruction results for Hispanic and non-Hispanic students. *The Journal of Special Education, 39*(2), 66-85.

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. New York: Routledge.

Hammill, D. D., & Newcomer, P. L. (2008). *Test of language development: Primary.* Alexandria, VA: Pearson Publishing.

Halliday, M. A., & Hasan, R. (1985). *Language, text and context.* Victoria, Australia: Deakin University Press.

Heath, S. B. (1983). *Ways with words: Language, life and work in communities and classrooms*. Cambridge, U.K.: Cambridge University Press.

Hiebert, E. H. (2011). Beyond single readability measures: using multiple sources of information in establishing text complexity. *Journal of Education*, *191*(2), 33-42.

Hiebert, E.H. (2012). The Common Core State Standards and text complexity. In M. Hougen & S. Smartt (Eds.), *Fundamentals of Literacy Instruction and Assessment, Pre-K–6*. (pp. 111-120). Baltimore, MD: aPaul Brookes Publishing.

Hiebert, E. H. (2005). The effects of text difficulty on second graders' fluency development. *Reading Psychology*, 26(2), 183-209.

Hipfner-Boucher, K., Milburn, T., Weitzman, E., Greenberg, J., Pelletier, J., & Girolametto, L. (2015). Narrative abilities in subgroups of English language learners and monolingual peers. International Journal of Bilingualism, 19(6), 677-692.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing*, *2*(2), 127-160.

Howard, E. & Arteagoitia, I. (2013). *The EVOCA Intervention.* Paper presented at the 2013 annual meeting of AERA, Vancouver, CA. Abstract retrieved from www.aera.org.

Hudson, P. T., & Bergman, M. W. (1985). Lexical knowledge in word recognition: Word length and word frequency in naming and lexical decision tasks. *Journal of Memory and Language*, *24*(1), 46-58.

Jenkins, J. R., & Dixon, R. (1983). Vocabulary learning. *Contemporary Educational Psychology*, *8*(3), 237-260.

Jitendra, A. K., Edwards, L. L., Sacks, G., & Jacobson, L. A. (2004). What research says about vocabulary instruction for students with learning disabilities. *Exceptional Children*, *70*(3), 299-322.

Joshi, R. M., & Aaron, P. G. (2000). The component model of reading: Simple view of reading made a little more complex. *Reading Psychology*, *21*(2), 85-97.

Kieffer, M. J., & Lesaux, N. K. (2012). Knowledge of words, knowledge about words: Dimensions of vocabulary in first and second language learners in sixth grade. *Reading and Writing*, *25*(2), 347-373.

Kirby, J. R., Deacon, S. H., Bowers, P. N., Izenberg, L., Wade-Woolley, L., & Parrila, R. (2012). Children’s morphological awareness and reading ability. *Reading and Writing*, *25*(2), 389-410.

Krashen, S. (2013). Reading and vocabulary acquisition: supporting evidence and some objections. *Iranian Journal of Language Teaching Research*, *1*(1), 27-43.

Laufer, B. (1998). What's in a word that makes it hard or easy? Intralexical factors affecting the difficulty of vocabulary acquisition. In B. Thomlinson (Ed.), *Materials Development in Language Teaching*, 140. Cambridge, U.K.: Cambridge University Press.

Laufer, B. (1998). The development of passive and active vocabulary in a second language: same or different? *Applied linguistics*, *19*(2), 255-271.

Laufer, B. (2001). Reading, word-focused activities and incidental vocabulary acquisition in a second language. *Prospect*, *16*(3), 44-54.

Lawrence, J. F., Capotosto, L., Branum-Martin, L., White, C., & Snow, C. E. (2012). Language proficiency, home-language status, and English vocabulary development: A longitudinal follow-up of the Word Generation program. *Bilingualism: Language and Cognition*, *15*(3), 437-451.

Lesaux, N. K., & Siegel, L. S. (2003). The development of reading in children who speak English as a second language. *Developmental Psychology*, *39*(6), 1005.

Lesaux, N. K., Crosson, A. C., Kieffer, M. J., & Pierce, M. (2010). Uneven profiles: Language minority learners' word reading, vocabulary, and reading comprehension skills. *Journal of Applied Developmental Psychology*, *31*(6), 475-483.

Lesaux, 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.

Lesaux, N. K., Kieffer, M. J., Kelley, J. G., & Harris, J. R. (2014). Effects of academic vocabulary instruction for linguistically diverse adolescents: Evidence from a randomized field trial. *American Educational Research Journal*, 51(6), 1159-1194.

Lonigan, C. J., Burgess, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: evidence from a latent-variable longitudinal study. *Developmental Psychology*, 36(5), 596.

Lotto, L., & De Groot, A. (1998). Effects of learning method and word-type on acquiring vocabulary in an unfamiliar language. *Language Learning*, *48*(1), 31-69.

Loveless, A. (2009). Thinking about creativity: developing ideas, making things happen. *Creativity in Primary Education. Exeter: Learning Matters*.

Mancilla-Martinez, J., & Lesaux, N. K. (2010). Predictors of reading comprehension for struggling readers: The case of Spanish-speaking language minority learners. *Journal of Educational Psychology*, *102*(3), 701.

Mancilla-Martinez, J., & Lesaux, N. K. (2011). The gap between Spanish speakers' word reading and word knowledge: A longitudinal study. *Child Development*, *82*(5), 1544-1560.

Marzano, R. J. (2004). *Building background knowledge for academic achievement: Research on what works in schools*. Alexandria, Virgina: ASCD.

McKeown, M. G. (1985). The acquisition of word meaning from context by children of high and low ability. *Reading Research Quarterly*, 482-496.

McKeown, M. G., & Beck, I. L. (1988). Learning Vocabulary Different Ways for Different Goals. *Remedial and Special Education*, *9*(1), 42-46.

McKeown, M. G., & Beck, I. L. (2004). Direct and rich vocabulary instruction. In J. F. Baumann & E. J. Kame (2004). *Direct vocabulary instruction: Research to practice* (pp. 13-27). New York: Guilford Press.

McKeown, M. G., Beck, I. L., Omanson, R. C., & Perfetti, C. A. (1983). The effects of long-term vocabulary instruction on reading comprehension: A replication. *Journal of Literacy Research*, *15*(1), 3-18.

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*, 522-535.

Meara, P. (2002). The rediscovery of vocabulary. *Second Language Research*, *18*(4), 393-407.

Metsala, J. L. (1997). An examination of word frequency and neighborhood density in the development of spoken-word recognition. *Memory & Cognition*, *25*(1), 47-56.

Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89-120). Hillsdale, NJ: Erlbaum.

Mezynski, K. (1983). Issues concerning the acquisition of knowledge: Effects of vocabulary training on reading comprehension. *Review of educational research*, *53*(2), 253-279.

Migration Policy Institute. (2010). *DREAM vs. reality: An analysis of potential DREAM Act beneficiaries*. Washington, DC: Jeanne Batalova and Margie McHugh.

Moll, L. C., Amanti, C., Neff, D., & Gonzalez, N. (1992). Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory into practice*, *31*(2), 132-141.

Moll, L., Amanti, C., Neff, D., & Gonzalez, N. (2005). Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. In N. Gonzalez, L. C. Moll, & C. Amanti (Eds.), *Funds of knowledge: Theorizing practices in households, communities, and classrooms* (pp. 71-87). New York, NY: Routledge.

Morrow, L. M. (2008). Foreword. In C. C. Block, & S. R. Parris (Eds.), *Comprehension instruction: Research-based best practices*. New York, NY: Guilford Press.

Nagy, W. E., & Hiebert, E. H. (2011). Toward a theory of word selection. In M. L. Kamil, P. D. Pearson, E. B. Moje, & P. P. Afflerbach (Eds.), *Handbook of Reading Research* (Vol. 4; pp. 388-404). New York, NY: Longman.

Nagy, W. E., & Scott, J. A. (2000). Vocabulary processes. In M.L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of Reading Research* (Vol. 3; pp. 269-284. New York, NY: Routledge.

Nagy, W. E., García, G. E., Durgunoğlu, A. Y., & Hancin-Bhatt, B. (1993). Spanish-English bilingual students' use of cognates in English reading. *Journal of Literacy Research*, *25*(3), 241-259.

Nagy, W. E., & Herman, P. A. (1984). Limitations of vocabulary instruction. *Center for the Study of Reading Technical Report; no. 326*.

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

Nakamoto, J., Lindsey, K. A., & Manis, F. R. (2007). A longitudinal analysis of English language learners’ word decoding and reading comprehension. *Reading and Writing*, *20*(7), 691-719.

Nakamoto, J., Lindsey, K. A., & Manis, F. R. (2008). A cross-linguistic investigation of English language learners' reading comprehension in English and Spanish. *Scientific Studies of Reading*, *12*(4), 351-371.

National Center on Response to Intervention (March 2010). *Essential components of RTI – A closer look at response to intervention*. Washington, DC: U.S.

Department of Education, Office of Special Education Programs, National Center on Response to Intervention.

National Governors Association Center for Best Practices and the Council of Chief State School Officers (2010). *Common Core State Standards*. Retrieved from http://www.corestandards.org/.

National Reading Panel (US), National Institute of Child Health, & Human Development (US). (2000). *Report of the national reading panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. National Institute of Child Health and Human Development, National Institutes of Health.

National Academies of Sciences, Engineering, and Medicine. (2017). *Promoting the Educational Success of Children and Youth Learning English: Promising Futures.* Washington, DC: The National Academies Press. https://doi.org/10.17226/24677.

Nation, I. S. (2001). *Learning vocabulary in another language*. Cambridge, UK: Cambridge University Press.

Nation, K., & Snowling, M. J. (1998). Semantic processing and the development of word-recognition skills: Evidence from children with reading comprehension difficulties. *Journal of Memory and Language*, *39*(1), 85-101.

Nation, K., Clarke, P., Marshall, C. M., & Durand, M. (2004). Hidden language impairments in children: Parallels between poor reading comprehension and specific language impairment? *Journal of Speech, Language, and Hearing Research*, *47*(1), 199-211.

Nelson, J. R., Vadasy, P. F., & Sanders, E. A. (2011). Efficacy of a tier 2 supplemental root word vocabulary and decoding intervention with kindergarten Spanish-speaking English learners. *Journal of Literacy Research*, *43*(2), 184-211.

Neuman, S. B., & Koskinen, P. (1992). Captioned television as comprehensible input: Effects of incidental word learning from context for language minority students. *Reading Research Quarterly*, *27*(1), 95-106.

Neuman, S. B., & Wright, T. S. (2015). *All about words: Increasing vocabulary in the Common Core classroom, Pre K-2*. New York, NY: Teachers College Press.

Nelson, J. R., Vadasy, P. F., & Sanders, E. A. (2011). Efficacy of a tier 2 supplemental root word vocabulary and decoding intervention with kindergarten Spanish-speaking English learners. *Journal of Literacy Research*, 43(2), 184-211.

Nicholson, T., & Whyte, B. (1992). Matthew Effects in learning new words while listening to stories. In C. K. Kinzer & D. J. Leu (Eds.), *Literacy research, theory, and practice: Views from many perspectives. Forty-first yearbook of the National Reading Conference* (pp. 499-503). Chicago: National Reading Conference.

Ouellette, G., & Beers, A. (2010). A not-so-simple view of reading: How oral vocabulary and visual-word recognition complicate the story. *Reading and Writing*, *23*(2), 189-208.

Oller, D. K., Pearson, B. Z., & Cobo-Lewis, A. B. (2007). Profile effects in early bilingual language and literacy. *Applied Psycholinguistics*, *28*(02), 191-230.

Office of Elementary and Secondary Education (OESE) (2018, January). ED Data Express. Retrieved January 2018 from: https://eddataexpress.ed.gov/definitions.cfm.

Páez, M., Gómez, L., & Hunter, C. (2011). *The Kindergarten Language Study: Improving the language and literacy skills of Spanish-English bilingual kindergarteners.* Paper presented a biennial Society for Research in Child Development (SRCD) meeting.

Paivio, A. (2013). *Imagery and verbal processes*. New York: Psychology Press.

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, *76*(1, pt. 2), 1-25.

Paradis, J., & Kirova, A. (2014). English second-language learners in preschool: Profile effects in their English abilities and the role of home language environment. *International Journal of Behavioral Development*, *38*(4), 342-349.

Paris, D. (2012). Culturally sustaining pedagogy: A needed change in stance, terminology, and practice. *Educational Researcher*, *41*(3), 93-97.

Payán, R. M., & Nettles, M. T. (2008). Current state of English-language learners in the US K-12 student population. In *2008 English Language Learner Symposium, Princeton, NJ. Retrieved October*, Vol. 29 (pp. 2008).

Pearson, P. D., Hiebert, E. H., & Kamil, M. L. (2007). Vocabulary assessment: What we know and what we need to learn. *Reading research quarterly*, *42*(2), 282-296.

Pearson. (2005). *Aprenda: La prueba de logros en español.* Retrieved from https://www.pearsonassessments.com/learningassessments/products/100000585/aprenda-3-aprenda-la-prueba-de-logros-en-espanol-tercera-edicion.html.

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*(1), 22-37.

Petersen-Brown, S., & Burns, M. K. (2011). Adding a vocabulary component to incremental rehearsal to enhance retention and generalization. *School Psychology Quarterly*, *26*(3), 245.

Peercy, M. M., Artzi, L., Martin-Beltran, M., & Silverman, R. D. (2015). Meeting the demands of the ELA Common Core for English language learners: Developing vocabulary and reading comprehension skills in a language-rich classroom environment. In L. C. de Oliveira, M. Klassen, & M. Maune, (Eds.), *The Common Core State Standards in English Language Arts for English Language Learners: Grades 6-12.* Alexandria, VA: TESOL.

Penno, J. F., Wilkinson, I. A., & Moore, D. W. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew effect? *Journal of Educational Psychology*, *94*(1), 23.

Perfetti, C. A., Wlotko, E. W., & Hart, L. A. (2005). Word learning and individual differences in word learning reflected in event-related potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(6), 1281.

Perfetti, C., McKeown, M. C., & Kucan, L. (2010). Decoding, vocabulary and comprehension: The golden triangle of reading skill. In M.C., McKeown and L. Kucan (Eds.), *Bringing reading research to life* (pp. 291-303). New York, NY: Guiliford Press.

Pressley, M., Kamil, M., L., Mosenthal, P. B, Pearson, P. D, Barr, R., (2000). *Handbook of reading research*, Vol. III. Mahwah, NJ: Lawrence Erlbaum Associates.

Proctor, C. P., & Mo, E. (2009). The relationship between cognate awareness and English comprehension among Spanish-English bilingual fourth grade students. *TESOL Quarterly*, *43*(1), 126-136.

Proctor, C. P., August, D., Snow, C., & Barr, C. D. (2010). The interdependence continuum: A perspective on the nature of Spanish–English bilingual reading comprehension. *Bilingual Research Journal*, *33*(1), 5-20.

Proctor, C. P., Carlo, M., August, D., & Snow, C. (2005). Native Spanish-speaking children reading in English: Toward a model of comprehension. *Journal of Educational Psychology*, *97*(2), 246.

Proctor, C. P., Dalton, B., Uccelli, P., Biancarosa, G., Mo, E., Snow, C., & Neugebauer, S. (2011). Improving comprehension online: Effects of deep vocabulary instruction with bilingual and monolingual fifth graders. *Reading and Writing*, *24*(5), 517-544.

Purcell-Gates, V. (2002). Authentic literacy in class yields increase in literacy practices. *Literacy Update*, *11*(7), 1-9.

Pullen, P. C., Tuckwiller, E. D., Konold, T. R., Maynard, K. L., & Coyne, M. D. (2010). A tiered intervention model for early vocabulary instruction: The effects of tiered instruction for young students at risk for reading disability. *Learning Disabilities Research & Practice*, 25(3), 110-123.

Ramirez, G., Chen, X., & Pasquarella, A. (2013). Cross-linguistic transfer of morphological awareness in Spanish-speaking English language learners: The facilitating effect of cognate knowledge. *Topics in Language disorders*, *33*(1), 73-92.

Read, J. (2000). *Assessing vocabulary*. Cambridge, UK: Cambridge University Press.

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. *Law and Human Behavior*, *29*(5), 615.

Roberts, T. A. (2011). Preschool foundations for reading and writing success. In R. E. O'Connor and P.F. Vadasy (Eds.), *Handbook of reading interventions* (pp. 27-62). New York, NY: The Guilford Press. Roberts, T. A. (2009). *No limits to literacy for preschool English learners*. Thousand Oaks, CA: Corwin Press.

Rojas, R., & Iglesias, A. (2013). The language growth of Spanish-speaking English language learners. *Child Development*, *84*(2), 630-646.

Ross, S. G., & Begeny, J. C. (2015). An Examination of treatment intensity with an oral reading fluency intervention: Do intervention duration and student-teacher instructional ratios impact intervention effectiveness? *Journal of Behavioral Education*, *24*(1), 11-32.

Ross, S. G., & Begeny, J. C. (2011). Improving Latino, English language learners' reading fluency: The effects of small-group and one-on-one intervention. *Psychology in the Schools*, *48*(6), 604-618.

Rupley, W. H., & Nichols, W. D. (2005). Vocabulary instruction for the struggling reader. *Reading & Writing Quarterly*, *21*(3), 239-260.

Rupley, W. H., Blair, T. R., & Nichols, W. D. (2009). Effective reading instruction for struggling readers: The role of direct/explicit teaching. *Reading & Writing Quarterly*, 25(2-3), 125-138.

Santoro, L. E., Jitendra, A. K., Starosta, K., & Sacks, G. (2006). Reading well with Read Well: Enhancing the reading performance of English language learners. *Remedial and Special Education*, *27*(2), 105-115.

Saunders, W., & Goldenberg, C. (2010). Research to guide English language development instruction. *Improving education for English learners: Research-based approaches* (pp. 21-81). Sacramento, CA: State Department of Education.

SAS Institute. (2014). *SAS/STAT 9.4 User's Guide the Reg Procedure:(*Book Excerpt). SAS Institute.

Saville-Troike, M. (1991). Teaching and testing for academic achievement: The role of language development. *FOCUS Occasional Papers in Bilingual Education, 4.* , Retrieved from http://www.ncela.gwu.edu/pubs/focus/focus4.htm

Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective*. Mahwah, NJ: Lawrence Erlbaum Associates.

Schwanenflugel, P. J., Akin, C., & Luh, W. M. (1992). Context availability and the recall of abstract and concrete words. *Memory & Cognition*, *20*(1), 96-10.

Schwanenflugel, P. J. (1991). Contextual constraint and lexical processing. *Advances in Psychology*, *77*, 23-45.

Shanahan, T., & Beck, I. L. (2006). Effective literacy teaching for English-language learners. In August, D., & Shanahan, T. (Eds), *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth. Executive Summary*. Mahwah, NJ: Lawrence Erlbaum Associates.

Silverman, R., & Hines, S. (2009). The effects of multimedia-enhanced instruction on the vocabulary of English-language learners and non-English-language learners in pre-kindergarten through second grade. *Journal of Educational Psychology*, *101*(2), 305.

Silverman, R.D., Martin-Beltran, M., Peercy, M.M., Hantraft, A.M., McNeish, D.M., Artzi, L., & Nunn, S.G. (2017). Effects of a cross-age peer learning program on the vocabulary and comprehension of ELs and Non-ELs in elementary school. *Elementary School Journal, 117*(3), 485-512.

Silverman, R. D. (2007). Vocabulary development of English-language and English-only learners in kindergarten. *The Elementary School Journal*, *107*(4), 365-383.

Snow, C. E., Lawrence, J. F., & White, C. (2009). Generating knowledge of academic language among urban middle school students. *Journal of Research on Educational Effectiveness*, *2*(4), 325-344.

Stahl, S. (1983). Differential word knowledge and reading comprehension. *Journal of Literacy Research*, *15*(4), 33-50.

Stahl, S. A. (2003). Words are learned incrementally over multiple exposures. *American Educator*, *27*(1), 18-22.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 22*, 360-407.

Stanovich, K. E., Nathan, R. G., & Vala-Rossi, M. (1986). Developmental changes in the cognitive correlates of reading ability and the developmental lag hypothesis. *Reading research quarterly*, 267-283.

Stevens, R. J., Liu, X., Baker, D. P., Ray, M., Eckert, S., & Gamson, D. (2015). Assessing the cognitive demands of a century of reading curricula: An analysis of reading text and comprehension tasks from 1910 to 2000. *American Educational Research Journal*, *52*(3), 582-617.

Swart, N. M., Muijselaar, M. M., Steenbeek-Planting, E. G., Droop, M., de Jong, P. F., & Verhoeven, L. (2017). Differential lexical predictors of reading comprehension in fourth graders. *Reading and writing*, *30*(3), 489-507.

Taboada, A., & Rutherford, V. (2011). Developing reading comprehension and academic vocabulary for English language learners through science content: A formative experiment. *Reading Psychology*, *32*(2), 113-157.

Thomas, W. P., & Collier, V. P. (2002). *A national study of school effectiveness for language minority students' long-term academic achievement.* Santa Cruz, CA: Center for Research on Education, Diversity and Excellence, University of California-Santa Cruz.

Thorndike, E. L., & Lorge, I. (1944). *The Teacher's Wordbook of 30,000 Words.* New York: Columbia University, Teachers College.

Tonzar, C., Lotto, L., & Job, R. (2009). L2 vocabulary acquisition in children: Effects of learning method and cognate status. *Language Learning*, *59*(3), 623-646.

Townsend, D., & Collins, P. (2009). Academic vocabulary and middle school English learners: An intervention study. *Reading and Writing*, *22*(9), 993-1019.

Townsend, D., Filippini, A., Collins, P., & Biancarosa, G. (2012). Evidence for the importance of academic word knowledge for the academic achievement of diverse middle school students. *The Elementary School Journal*, *112*(3), 497-518.

Turner, H. M., & Bernard, R. M. (2006). Calculating and synthesizing effect sizes. *Contemporary issues in communication science and disorders*, 33(1), 42-55.

U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance (NCEE). (2014). Teaching academic content and literacy to English learners in elementary and middle school (Report No. NCEE 2014-4012). Retrieved from http://ies.ed.gov/ncee/wwc/publications_reviews.aspx

U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. (2007). *Effective literacy and English language instruction for English learners in the elementary grades: IES practice guide* (Report No. NCEE 2007-4011). Retrieved from https://ies.ed.gov/ncee/wwc/pdf/practice_guides/20074011.pdf

U.S. National Institutes of Health, National Institute of Child Health & Human Development, National Reading Panel (2000). *Report of the national reading panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.

Uccelli, P., Barr, C. D., Dobbs, C. L., Galloway, E. P., Meneses, A., & Sanchez, E. (2015). Core academic language skills: An expanded operational construct and a novel instrument to chart school-relevant language proficiency in preadolescent and adolescent learners. *Applied Psycholinguistics*, *36*(05), 1077-1109.

Uccelli, P., Dobbs, C. L., & Scott, J. (2013). Mastering academic language organization and stance in the persuasive writing of high school students. *Written Communication*, *30*(1), 36-62.

Uccelli, P., Rosenthall, F., August, D., Artzi, L., & Barr, C. (2012, April). *Promising connections: Play, multimodality, and bilingualism in learning connectives at school*. In S. Marinova-Todd & Y. Uchikoshi (Chairs), *Words, palabras, 詞/*词*: Effect of vocabulary on literacy of ELLs from pre-K through grade two*. Symposium conducted at the annual meeting of AERA, Vancouver, Canada.

Uchikoshi, Y. (2014). Development of vocabulary in Spanish-speaking and Cantonese-speaking English language learners. *Applied Psycholinguistics*, *35*(01), 119-153.

Uchikoshi, Y. (2005). Narrative development in bilingual kindergarteners: can Arthur help? *Developmental Psychology*, *41*(3), 464.

Vadasy, P. F., Nelson, J. R., & Sanders, E. A. (2013). Longer term effects of a tier 2 kindergarten vocabulary intervention for English learners. *Remedial and Special Education*, *34*(2), 91-101.

Vadasy, P. F., Sanders, E. A., & Nelson, J. R. (2015). Effectiveness of supplemental kindergarten vocabulary instruction for English learners: A randomized study of immediate and longer-term effects of two approaches. *Journal of Research on Educational Effectiveness*, 8(4), 490-529.

Vaughn, S., Cirino, P. T., Linan-Thompson, S., Mathes, P. G., Carlson, C. D., Hagan, E. C., & Francis, D. J. (2006). Effectiveness of a Spanish intervention and an English intervention for English-language learners at risk for reading problems. *American Educational Research Journal*, *43*(3), 449-487.

Vaughn, S., Martinez, L. R., Linan-Thompson, S., Reutebuch, C. K., Carlson, C. D., & Francis, D. J. (2009). Enhancing social studies vocabulary and comprehension for seventh-grade English language learners: Findings from two experimental studies. *Journal of Research on Educational Effectiveness*, *2*(4), 297-324.

Verhoeven, L., van Leeuwe, J., & Vermeer, A. (2011). Vocabulary growth and reading development across the elementary school years. *Scientific Studies of Reading*, *15*(1), 8-25.

Verhoeven, L., & Perfetti, C. A. (2011). Vocabulary growth and reading skill. *Scientific Studies of Reading*, *15*(1), 1-7.

Verhoeven, L., van Leeuwe, J., & Vermeer, A. (2011). Vocabulary growth and reading development across the elementary school years. *Scientific Studies of Reading*, *15*(1), 8-25.

Walley, A. C., Metsala, J. L., & Garlock, V. M. (2003). Spoken vocabulary growth: Its role in the development of phoneme awareness and early reading ability. *Reading and Writing*, *16*(1-2), 5-20.

White, T. G., Graves, M. F., & Slater, W. H. (1990). Growth of reading vocabulary in diverse elementary schools: Decoding and word meaning. *Journal of Educational Psychology*, *82*(2), 281.

Williamson, G. L., Fitzgerald, J., & Stenner, A. J. (2013). The Common Core State standards' quantitative text complexity trajectory figuring out how much complexity is enough. *Educational Researcher*, *42*(2), 59-69.

Williamson, G. L., Fitzgerald, J., & Stenner, A. J. (2014). Student reading growth illuminates the Common Core text-complexity standard. *The Elementary School Journal*, *115*(2), 230-254.

Wijekumar, K. K., Meyer, B. J., & Lei, P. (2016). Improving content area reading comprehension with 4-6th grade Spanish ELLs using web-based structure strategy instruction. In *Competencies in teaching, learning and educational leadership in the digital age* (pp. 123-135). Springer International Publishing.

White, J. L., & Horst, M. (2012). Cognate awareness-raising in late childhood: teachable and useful. *Language Awareness*, *21*(1-2), 181-196.

Zeno, S. M., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide*. Brewster, NY: Touchstone Applied Science Associates, Inc.