INTERNATIONAL LITERACY ASSOCIATION READING RESEARCH QUARTERLY

English Learners' Acquisition of Academic Vocabulary: Instruction Matters, but So Do Word Characteristics

Diane August
Center for Applied Linguistics, Washington, DC, USA

Paola Uccelli
Harvard Graduate School of Education, Cambridge, Massachusetts, USA

Lauren Artzi
American Institutes for Research, Washington, DC, USA

Christopher Barr
Rice University, Houston, Texas, USA

David J. Francis
University of Houston, Texas, USA

ABSTRACT

In this cluster randomized control trial study, the authors explored the efficacy of an intervention designed to improve second-grade English learners' knowledge of challenging, high-utility English vocabulary. The authors also examined whether the intervention had a differential effect on content words that differ on two attributes (cognate status and abstractness designation) and whether the effects lasted across time. Within schools, teachers were randomly assigned to treatment or control conditions, resulting in 12 treatment classrooms and 10 control classrooms. Teachers in the treatment condition developed vocabulary using six methods: interactive shared reading, direct instruction of individual words, instruction in cognate use, activities to foster word consciousness and reinforce instructed words, and use of visual and linguistic supports. Teachers in the control condition read the same books with the target vocabulary inserted into the stories. Findings on the curriculum-based and standardized measure of vocabulary indicated that the intervention implemented for four 50-minute lessons per week for 18 weeks was effective in helping English learners acquire challenging high-utility school-relevant vocabulary (Cohen's d = 1.88 for content words; d = 0.41 for connecting words; d = 0.47 for the Test of Oral Language Development). The effects differed by word characteristics, with higher effect sizes for words that are noncognates (d = 1.57 for both concrete and abstract noncognates; d = 1.02 concrete cognates; d = 0.81 for abstract cognates). Ten months after the intervention, treatment students still outperformed control students on content vocabulary learned during the intervention (d = 1.31). This study helps validate a multifaceted approach to vocabulary intervention and research.

In this study, we explored the immediate and sustained effects of an interactive shared reading intervention, Acquiring Vocabulary in English, on the learning of high-utility English vocabulary in second-grade Spanish-speaking English learners (ELs). The intervention is innovative in three specific ways. First, it targets two types of words: content words, which are meaning-carrying words, including nouns, verbs, adjectives, and most adverbs, and connecting words,[1] which are function words that signal relations between ideas in text (e.g., *because*, *therefore*). (Connecting words are sometimes referred to as connectives, but we use the term *connecting words* in this manuscript to distinguish these words from content words.) Most vocabulary interventions to date have focused exclusively on content words, overlooking other types of words. Second, we examined whether the impact of the intervention varied by word type (i.e., content vs. connecting words), by content word characteristics, including cognate status (i.e., whether words are cognates or not), and by

INTERNATIONAL LITERACY ASSOCIATION

Reading Research Quarterly, 0(0)
pp. 1–24 | doi:10.1002/rrq.323
© 2020 International Literacy Association.

abstractness designation (i.e., the extent to which a word's meaning is imageable and perceptible). These attributes have rarely been explored in vocabulary intervention research with school-age students. Third, the structured interactive shared reading was intended to support the development of oral language in general, not solely the high-frequency target words, as it provided opportunities for teachers and students to engage in extended discourse characterized by multiturn conversations, student-initiated conversations, and teacher elicitations and extensions, all features associated with students' language growth (Cabell, Justice, McGinty, DeCoster, & Forston, 2015).

In the following sections, we discuss the theoretical framework and empirical research that drives our focus on vocabulary as a crucial area to support ELs' comprehension, review lessons learned about intervention design from prior research, and present a research-based rationale for the need to move beyond content words in vocabulary instruction.

Expanding Vocabulary Knowledge to Support Comprehension

From Gough and Tunmer's (1986) simple view of reading to Kintsch's (1991) construction–integration model and Perfetti and Stafura's (2014) reading systems framework, psychological theories of reading have long and consistently considered vocabulary knowledge a crucial component of reading comprehension. Further elaborating on the role of vocabulary in reading, Perfetti's lexical quality hypothesis posits that readers' vocabulary knowledge serves a determining double role: It supports both word recognition (i.e., knowing a word makes it easier to recognize it in print) and text comprehension (i.e., knowing a word enables readers' access to the text's meaning; (Perfetti, Landi, & Oakhill, 2005; Perfetti & Stafura, 2014). According to this widely accepted theory, having high-quality lexical representations entails being skilled in accessing orthographic, phonological, and semantic representations of words in a text (Perfetti & Hart, 2002).

Extensive empirical research has also identified vocabulary knowledge as a crucial source of variation for word recognition and comprehension in English-proficient readers (Beck, McKeown, & Kucan, 2002; Foorman et al., 2016; Freebody & Anderson, 1983; Goswami, 2002; Perfetti et al., 2005) and ELs learning in English (Lesaux & Geva, 2006; Mancilla-Martinez & Lesaux, 2010; National Academies of Sciences, Engineering, and Medicine [NASEM], 2017). Vocabulary knowledge is an essential enabler for text comprehension in that readers are estimated to need to know 98% of the words in a text for unassisted comprehension (Schmitt, Jiang, & Grabe, 2011). Even though these are approximate estimates that would vary according to the combination of text, reader, and activity (RAND Reading Study Group, 2002), these percentages are instructive in light of current U.S. English language arts standards that require students to comprehend more complex texts at earlier grades to be college and career ready (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). Given these more demanding standards, it is imperative to find effective methods to promote ELs' learning of school-relevant vocabulary early on that will support their comprehension of challenging texts read aloud to them or read by them.

Helping students develop their vocabulary in the lower elementary grades is important also because vocabulary knowledge is highly sensitive to instructional and other contextual factors, as well as cumulative in that learning from context is hampered by limited vocabulary (Biemiller & Slonim, 2001; Mancilla-Martinez & Lesaux, 2011; Stanovich, 1986). Informed by this research, the Acquiring Vocabulary in English intervention was intended to improve ELs' school-relevant high-utility vocabulary early in their education to make them more autonomous learners.

Designing a Vocabulary Intervention: Lessons From Prior Research

The development of the intervention presented in this study was informed by findings from prior studies focused on developing English vocabulary in both English-proficient learners (Beck et al., 2002; Graves, 2016; Language and Reading Research Consortium, Jiang, & Davis, 2017) and ELs (August, Artzi, Barr, & Francis, 2018; Crevecoeur, Coyne, & McCoach, 2014; Lesaux, Kieffer, Kelley, & Harris, 2014; Proctor et al., 2011; Proctor, Silverman, Harring, Jones, & Hartranft, 2020; Stahl & Nagy, 2006; Vadasy, Sanders, & Nelson, 2015). As a result of these intervention studies, the field has learned much about how to support word learning through pedagogy (Graves, 2015; NASEM, 2017). Research has pointed to a multifaceted approach that includes immersion in language-rich environments, opportunities for direct instruction of individual words, instruction in word-learning strategies, the development of students' interest in and awareness of words, and visual and linguistic supports for students learning in a second language.

Interactive shared book reading, in which adults read narrative and informational texts aloud to children, pausing to discuss the meanings of a relatively small number of individual words, asking questions about the book, and following up and elaborating on students' responses, constitutes a common method of immersing students in

language-rich environments (Beck & McKeown, 2001, 2007; Biemiller & Boote, 2006; Wiseman, 2011). Research conducted with primary school students has highlighted the importance of using particular types of questions, those that "develop students' inferential language—such as predicting, problem-solving, hypothesizing, or contrasting—with conversations before, during, and after read-alouds" (Foorman et al., 2016, p. 8) to build students' academic language. Studies conducted with ELs have used similar techniques but used "right there" and inferential questions to support ELs' comprehension (e.g., August, Artzi, & Barr, 2016). Research using these techniques has found positive results for vocabulary outcomes for ELs (August et al., 2016, 2018; Biemiller & Boote, 2006; Carlo et al., 2004; Foorman et al., 2016; Roberts, 2008; Silverman, 2007).

Direct teaching of content and connecting words is supported by a large body of research with English-proficient students (Foorman et al., 2016; Graves, 2016) and ELs (e.g., August et al., 2016; Crosson & Lesaux, 2013; Lawrence, Capotosto, Branum-Martin, White, & Snow, 2012; Mancilla-Martinez, 2010; Silverman, 2007). Direct teaching of content and connecting words constitutes defining vocabulary terms for students prior to and/or during instruction, as differentiated from acquisition through exposure alone. For both word types, research has indicated that successful strategies for elementary school students include the direct teaching of these words in meaningful and engaging contexts; use of picture cards that support understanding; paraphrasing, synonyms, and generation of multiple examples to describe and illustrate meaning; use of students' home languages as a resource; and engaging students in experiences that promote deep processing and recurrent production of word meanings, such as songs and Readers Theatre (August et al., 2016; Carlo et al., 2004; Crevecoeur et al., 2014; Graves, 2016; Graves, August, & Mancilla-Martinez, 2013). Additionally, for connecting words, playful, kinesthetically based demonstrations of the function of connecting words and sentence combining through writing have also been used to support acquisition (Crosson & Lesaux, 2013). Word-learning strategies have also been proven effective across various studies. For English-proficient learners, studies have included teaching students to use (a) context to infer the meanings of unknown words, (b) word parts to glean word meanings, and (c) the dictionary (Graves, 2016; Stahl & Nagy, 2006). The second-language research literature is more limited, but methods that teach students about word parts (Lesaux, Kieffer, Faller, & Kelley, 2010) and cognates (August et al., 2016) have been effective in supporting ELs' vocabulary development.

Given the magnitude of the vocabulary-learning task, fostering students' awareness of and interest in words and word learning is indispensable. In a series of studies associated with the Vocabulary Innovations in Education (VINE) intervention, students in multilingual classrooms were provided with an enriched focus on words in their reading, writing, and discussion. Findings indicated that all students, including the ELs, acquired enough generative knowledge of words and word-learning strategies through the enriched focus to significantly outperform a comparison group of students on a vocabulary test comprised of words that were not explicitly taught (Scott, 2015; Scott, Miller, & Flinspach, 2012; Scott & Nagy, 2004). Findings from other studies that included building word consciousness as one component of a multifaceted approach also produced positive results in ELs (Carlo et al., 2004; Lesaux et al., 2010).

Research also has pointed to the importance of providing additional visual and linguistic supports for ELs who are learning in a second language. These supports enhance ELs' comprehension and engagement (August et al., 2016; Carlo et al., 2004; Dressler, Carlo, Snow, August, & White, 2011; Fillmore, 2014; Goldenberg, 2013). Examples of visual supports include gestures, pictures, realia, and multimedia (i.e., a combination of text, sound, and/or video). Examples of linguistic supports include capitalizing on ELs' home languages and literacies, modeling language and tasks, elaborating on students' comments, recasting incorrect production, and providing additional exposure to words, concepts, and skills.

Integrating the research discussed above, the Acquiring Vocabulary in English intervention was designed to strategically combine interactive read-louds, direct vocabulary instruction, teaching of word-learning strategies, promotion of word awareness, and visual and linguistic supports.

Expanding the Intervention Beyond Content Words to Include Connecting Words and Word-Level Factors

Most intervention studies to date have targeted content words almost exclusively. The Acquiring Vocabulary in English intervention included two types of English words (i.e., content, connecting), as well as content words that differ by two characteristics: cognate status and abstractness designation (i.e., the extent to which a word's meaning is imageable and perceptible).

Connecting Words

Besides understanding the content words in a text (i.e., those that carry the semantic content, including nouns, main verbs, adjectives, and most adverbs), reading comprehension is also supported by knowledge of connecting words, namely, words that mark relations between ideas (e.g., *therefore*, *meanwhile*). Knowledge of connecting

words positively contributes to reading comprehension for all readers by providing them with cues into how ideas are linked in a text. Connecting words function as glue words that signal temporal or logical relations, such as additive, contrastive, or causal relations, which serve as helpful guides for readers, including elementary readers (Crosson & Lesaux, 2013). Moreover, textual analysis research has shown that connecting words are highly frequent in late elementary U.S. school texts (Nair, 2007).

Research has highlighted that knowledge of connecting words varies considerably across individuals and is positively associated with reading comprehension in first and second languages, even after controlling for vocabulary knowledge and word recognition skills (Barr, Uccelli, & Phillips Galloway, 2019; Crosson & Lesaux, 2013; Pretorius, 1996, 2006; Uccelli, et al., 2015; Uccelli, Deng, Phillips Galloway, & Qin, 2019; Welie, Schoonen, Kuiken, & van den Bergh, 2017). Prior research has demonstrated that by grade 4, ELs were significantly behind their English-proficient peers in knowledge of connecting words even when instructed in English and Spanish (Crosson & Lesaux, 2013; Crosson, Lesaux, & Martiniello, 2008). Readers benefit from texts with more relevant causal connecting words in their first and second languages (Degand & Sanders, 2002; O'Reilly & McNamara, 2007; Sánchez & García, 2009), but these markers will be helpful only if readers understand them. As such, connecting words represent an underexplored yet especially promising vocabulary knowledge domain to leverage ELs' reading comprehension (Crosson et al., 2008; Crosson & Lesaux, 2013).

Word-Level Factors: Abstractness Designation and Cognate Status

Beyond the broad classification of words into content and function words, content words can be further grouped into types of words based on specific word characteristics that may be associated with the ease of vocabulary learning. Prior research has suggested that a variety of word characteristics influence vocabulary acquisition: opaqueness, morphological family size, level of abstractness, and cognate status (Brysbaert, Warriner, & Kuperman, 2014; Goodwin & Cho, 2016; Nagy, Anderson, & Herman, 1987; Nagy & Townsend, 2012; Proctor & Mo, 2009; Salsbury, Crossley, & McNamara, 2011; Tonzar, Lotto, & Job, 2009). In the current study, we focused on two of these word characteristics that significantly predict vocabulary acquisition for ELs: level of abstractness and cognate status.

With regard to level of abstractness, the shape, individuation, concreteness, and imageability continuum argues that regardless of grammatical form, students learn words earlier whose meanings are perceptually salient, clearly distinguishable from others, concrete, and likely to yield a mental image (Maguire, Hirsh-Pasek, & Golinkoff, 2006). Second-language learners have also been shown to acquire these words earlier than words that are more abstract (Salsbury et al., 2011) and with more ease (Ellis & Beaton, 1993) and to be more likely to retain these words (De Groot & Keijzer, 2000). Guided by this developmental research, we anticipated that more abstract word meanings would be more challenging for ELs to learn.

ELs' learning of individual English words is facilitated by their home-language vocabulary knowledge, especially when pairs of words across languages are phonologically transparent cognates with orthographic and semantic overlap (e.g., Spanish/English *inteligente*/*intelligent*; De Groot & Keijzer, 2000; Dressler et al., 2011; Tonzar et al., 2009). Students able to read in both Spanish and English who have cognate awareness draw on this knowledge when reading in English (Nagy, García, Durgunoğlu, & Hancin-Bhatt, 1993). For example, in one study, bilingual students identified as successful English readers used three strategies that capitalized on their first-language knowledge: actively transferring information across languages, translating from one language to the other, and accessing cognate knowledge when reading in their less dominant language (Jiménez, García, & Pearson, 1996). Supporting ELs in using cognate knowledge to unpack the meanings of English cognates is especially important because Spanish cognate knowledge in Spanish–English bilingual students has also been associated with higher scores on measures of English reading comprehension (Proctor, Carlo, August, & Snow, 2005; Proctor & Mo, 2009).

The Current Study

This study expands our knowledge base related to vocabulary development in ELs in multiple ways. The school-relevant vocabulary knowledge targeted in the intervention included not only content words but also connecting words. Connecting words (e.g., *therefore*, *because*) were a focus of the intervention because they have high utility for text comprehension and yet are unknown in their English form to most young Spanish-speaking ELs. This is, to our knowledge, the first intervention study that has attempted to teach connecting words to such a young sample of bilingual learners.

Besides measuring the impact of the intervention on the learning of these two types of words separately, we also paid attention to the role of content word–level characteristics by examining whether the intervention functioned differently across four types of content words: concrete cognates, abstract cognates, concrete noncognates, and abstract noncognates. Finally, we explored the immediate effect and sustained effects of this multifaceted vocabulary intervention on vocabulary that was directly instructed and on general vocabulary knowledge as indexed by a standardized vocabulary measure.

This study is driven by the following research questions (RQs):

RQ1a: What are the effects of the intervention on ELs' high-utility school-relevant English vocabulary knowledge, including content and connecting words?

RQ1b: Does the effect of the intervention on ELs' knowledge of English content words differ by word characteristics (i.e., cognate/noncognate, concrete/abstract)?

RQ2: Does the effect of the intervention generalize beyond the words being taught, as measured by a standardized vocabulary assessment?

RQ3: Do the effects last 10 months after the end of the intervention?

Guided by prior research, we anticipated that students in treatment classrooms would make greater gains on instructed words than students in control classrooms would. Because of the robust instruction in conjunction with reinforcement practices, we expected treatment students to make greater gains on a standardized assessment of vocabulary relative to control students and expected vocabulary gains to be maintained at delayed posttest. As for word type, we hypothesized that students might display lower gains in connecting words because they are considered more challenging to learn than content words; also, in contrast to the instruction of content words, this aspect of the intervention was much more innovative, with only scarce research to draw on on how to teach connecting words to young bilingual learners. Finally, for the impact of word characteristics, we expected that abstract and noncognate words would be more difficult to learn without instruction, and thus, perhaps the intervention would have a stronger impact on learning words with those characteristics.

Method

Design

We conducted a stratified cluster randomized control trial, randomizing teachers within schools to either the treatment or control condition. In schools with even numbers of teachers, the same number of teachers were assigned to treatment and control. Schools with odd numbers of teachers were randomly paired, and then one school was randomly selected to have two treatment teachers and a control teacher and the other school to have two control teachers and one treatment teacher. In two schools, a single teacher served as the English as a second language (ESL) instructor, with one of these schools having three ESL classrooms and the other school having two. Due to their similarity in terms of instruction taking place in the designated ESL classrooms, these two schools were paired, and the school with three classrooms was randomly selected to the treatment condition and the other school to the control condition. This entire process yielded 12 treatment classrooms and 10 control classrooms.

Setting

The study took place in eight elementary schools in a school district in the Southwestern United States. In the school year in which the study took place, on average, Hispanic/Latinx students accounted for the majority of students in the eight schools (mean of 75.7% across the eight schools), followed by African American students (mean of 19.7%) and White students (mean of 1.2%). In these schools, on average, 94% of the students were economically disadvantaged.

Sample

The sample consisted of 424 second-grade Spanish-speaking ELs. Students were identified as ELs by the district using the state's English Language Proficiency Assessment System, given annually to students in the district identified as speaking a home language other than English. The study sample consisted of emergent bilingual learners; students were less proficient in English in second grade than most EL samples, as students in the sample had been primarily instructed in Spanish at the time of the intervention and were not yet reading in English.

Survey findings indicated that 92% of families earned less than $40,000 per year; the median education level for both mothers and fathers was completion of junior high (grades 7–9), and home language use was all or mostly Spanish for both mothers and fathers. Appendix A provides additional information about the family survey and family characteristics. To examine whether treatment and control groups differed on demographic variables, we ran a series of one-way analyses of variance (ANOVAs) with income, parental education, or home language use being predicted by group assignment. We found group assignment to be statistically nonsignificant in all models.

Findings from a survey administered to teachers at the start of the intervention indicated that they were predominantly female (n = 18; 90%). Most teachers identified as Hispanic/Latinx (n = 17), with one control teacher reporting as White, one control teacher reporting as Black, and one control teacher reporting as multiracial. Treatment and control teachers were similar in tenure (mean = 13 years for treatment teachers and 15.4 years for control teachers) and years of teaching ESL (mean = 13.3 years for treatment teachers and 13.6 years for control teachers).

Eighteen teachers had bilingual teaching certification (10 treatment and eight control). Of these 18, eight also had an ESL certification (four treatment and four control),

and 10 were certified in regular classroom instruction (three treatment and seven control). In the treatment group, all four teachers with ESL certification also had bilingual certification, but in the control group, only two of the four teachers with ESL certification had bilingual certification.

On the survey, teachers rated their language proficiency in both English and Spanish on four domains—listening, speaking, reading, and writing—on a scale of 1 (*very limited*) to 4 (*high*). Treatment teachers' average Spanish proficiency self-ratings were slightly higher than those of control teachers (for treatment teachers, 4.0 for speaking, listening, and reading and 3.9 for writing; for control teachers, 3.6 in speaking and reading, 3.5 in listening, and 3.4 in writing). However, treatment teachers' average English proficiency self-ratings were slightly lower than those of control teachers (for treatment teachers, 3.8 in listening, 3.6 in speaking and reading, and 3.4 in writing; for control teachers, 3.9 for speaking and listening and 3.8 for reading and writing).

The intervention mentors rated treatment teachers' levels of oral proficiency in English and Spanish (fluency, vocabulary, pronunciation, and grammar) on a 5-point scale (1 = *very limited*; 5 = *high*). Findings indicated that teacher self-ratings and mentor ratings on English were strongly correlated at around 0.6 on average across all the dimensions assessed. With regard to Spanish, there was virtually no variability, with self- and mentor ratings being uniformly rated at the maximum level on the scale.

The Acquisition of Vocabulary in English Intervention

Overview

The Acquisition of Vocabulary in English intervention was implemented for 18 weeks in students' 50-minute designated ESL class. For the remainder of the day, students were instructed mostly in Spanish because they participated in a transitional bilingual education program. As reported by the study teachers, Spanish was used, on average, for 66% of the day in treatment classrooms and 63% of the day in control classrooms.

The intervention consisted of nine 2-week units; each unit was composed of eight 50-minute lessons designed to be administered four days per week. The last unit was a general review of the vocabulary taught during the first eight units. Each intervention unit was based on a high-quality children's trade book and aligned with state English language arts content standards. Books selected were award-winning children's literature with multicultural themes in a Big Book format.

Materials for treatment teachers included teaching guides for each unit that supported teachers in implementing the intervention activities; picture cards for all content and connecting words; student workbooks aligned with the teacher guides, with activities to reinforce student learning of content and connecting words (e.g., songs with questions, plays, crossword puzzles, weekly quizzes); and student home connection packets with engaging activities that students could complete at home to reinforce word meanings (e.g., clip magazine pictures that represent target words, select a favorite word from a list of unit target words and explain why).

Teachers in the control condition also implemented Acquiring Vocabulary in English control activities for 18 weeks during students' designated ESL class. For the remainder of the day, students were instructed mostly in Spanish, as was the case for treatment students, because they participated in a transitional bilingual education program. Teachers in the control condition were given the same trade books that teachers in the treatment condition used and teaching guides for each unit that included the storybook script with target words inserted into it.

Control teachers were asked to read the books to students using the same methods usually used by the teachers. Any other activities that took place during the designated ESL class were at the discretion of the control teachers. In this sense, the control group offered a counterfactual scenario that was stronger than just a business-as-usual control. Even if unlikely, one plausible scenario was to find that exposure only was sufficient for students to learn the new words.

Word Selection

Content Word Selection

The intervention was designed to teach content words that were purposefully selected to differ by cognate status and abstractness designation. Examples of the four types of content words are *event* (concrete cognate), *illness* (concrete noncognate), *responsibility* (abstract cognate), and *pride* (abstract noncognate). As a starting place, word forms that appeared at least 10 times per million words in school texts at grade levels 1–6 were drawn from *The Educator's Word Frequency Guide* (Zeno, Ivens, Millard, & Duvvuri, 1995) so findings would generalize back to a meaningful corpus. Because the guide consists of word forms that can have various meanings, meanings at the appropriate grade levels as indexed by the Living Word Vocabulary (Dale & O'Rourke, 1981) were selected and matched to the word forms that that been selected because they appeared frequently in grades 4–6 texts. The 1,335 content words that met the criteria for consideration were words that appeared frequently in grades 4–6 texts whose meanings had been acquired by most English-proficient students in grades 5 and 6 but not before to help ensure that students were not at ceiling at pretest. See August et al. (2018) for additional information about content word selection.

COGNATE STATUS. To code for cognate status, after initial training, two bilingual research assistants coded all 1,335 words, coding words that displayed orthographic and semantic overlap with a Spanish word (i.e., *event*) as cognates (1) and coding words that did not as noncognates (0). In the few instances of disagreement, words were discussed with a third researcher associated with the project who was bilingual, and consensus among the three coders was reached to determine cognate status. For the two raters, inter-rater reliability was indexed by a Kappa coefficient of .86.

ABSTRACTNESS DESIGNATION. Words selected as abstract were those whose referents were language based rather than situated within direct experience (Brysbaert et al., 2014) and less likely to stimulate a mental image (Paivio, Yuille, & Madigan, 1968). Informed by previous research (De Groot & Keijzer, 2000; Nagy et al., 1987; Nagy & Townsend, 2012), two research assistants used a 5-point scoring rubric (1 = *least abstract*; 5 = *most abstract*) to jointly code 15% of the file of words at three timepoints. Immediately after training, Kendall's coefficient of concordance for ordinal responses was .94, the midpoint was .90, and the end was .92. Coefficients above .7 are considered acceptable, and 1 would indicate perfect agreement. Appendix B displays raters' codes for the levels of abstractness. Because our goal was for subjects' scores to generalize back to *The Educator's Word Frequency Guide*, words rated as abstract had ratings of 4 and 5, words rated as concrete had ratings of 1 and 2, and words rated as 3 were omitted.

Words were sorted into four bins—concrete cognates, abstract cognates, concrete noncognates, and abstract noncognates—and randomly ordered within each bin. To select words to insert into the storybooks, we started with the word at the top of each bin and continued down the list until we could find a word that could be inserted into the storybook without changing the meaning of the text or the author's voice. For example, in the book *Too Many Tamales* by Gary Soto (1996), the word *remarkable* was selected, and two sentences, including one using the word *remarkable*, was added to the following target text: "Maria loved that ring. She loved how it sparkled, like their Christmas lights. *This is a remarkable ring, Maria thought. She had never seen anything so shiny*" (n.p.). (The added phrase is italicized here but not in the intervention materials.) The target content words by word characteristics inserted into the intervention storybooks are displayed in Appendix C.

Twenty-four words from each word group (abstract cognates, abstract noncognates, concrete cognates, and concrete noncognates) were inserted into the storybooks because pilot work indicated that this number was needed to show effects for the four word groups. Three words from each word group were inserted into each storybook. In all, 96 content words were inserted.

Connecting Word Selection

Using data from a previous pilot study (August et al., 2018) and word frequency lists, connecting words were selected using three criteria: (1) those documented to be highly frequent in school texts during the early grades (grades 1–3) and during the later grades (grades 4–6; Zeno et al., 1995), (2) those classified as level 4 or 6 of the Living Word Vocabulary (Dale & O'Rourke, 1981), and (3) those known in English by 15% or fewer third-grade ELs (August et al., 2018). Connecting words that met these criteria were stratified according to word frequency level (i.e., levels 4 and 6 of the Living Word Vocabulary) and conceptual relation (i.e., temporal/aspectual, causal, contrastive; Crosson et al., 2008). Words that were the focus of instruction were selected because they met these criteria and expressed relations that were core to the retelling of the main plot in the storybooks. Appendix D displays connecting words' conceptual relation designation and Living Word Vocabulary level.

Given that connecting words do not carry lexical meaning, signaling only temporal/aspectual or logical relations, we anticipated that understanding and learning these words would be more challenging than content words. Consequently, we focused on only a small set of connecting words and taught two connecting words per unit so children could better understand each connecting word in contrast to another (e.g., *because*, *meanwhile*). In all, 12 connecting words were selected for instruction, with two words taught per unit over the course of six units and then reinforced during three additional units.

Intervention Instructional Method

The intervention used an approach that included interactive shared reading, direct instruction of content and connecting words, teaching word-learning strategies, and activities to foster word consciousness. All activities were scaffolded to support the learning of emergent bilingual learners through visual supports and linguistic supports, including those linguistic supports that capitalized on students' home-language knowledge and skills. Prior to and throughout the intervention, treatment teachers were provided with training and support.

For the most part, there were no differences in how connecting and content words were taught. Both types of words were directly instructed with picture cards, were taught in the context of interactive shared reading, and were reinforced with a variety of playful activities. The differences were the following: Word-learning strategies were

focused on cognates; because connecting words do not carry lexical meaning on their own, to directly teach connecting words, teachers modeled combining sentences using a connecting word, called on individual students to combine sentences using the same connecting word, and then asked students to work in pairs to generate and write additional sentences using this connecting word. Appendix E lists activities by day to develop vocabulary for both word types. Appendix F displays examples of activities to develop intervention vocabulary described in the preceding section (Appendixes F–L are available as supporting information for the online version of this article).

Interactive Shared Reading

The weekly read-aloud book was a Big Book that had been carefully chosen for its high quality as a work of children's literature, an appropriate language level for second-grade ELs, and themes that were engaging and could be used to reinforce state content standards. Teachers used interactive shared reading techniques to read approximately one quarter of a book each day. After reading each page, the teacher summarized the text using paraphrased definitions of target words and questioned students about the meaning of words and connected text. For example, a teacher might model the meaning of the word *especially* (e.g., "Maria especially likes the diamond ring, and I especially like my pet cockapoo") and then ask a student an inferential question that requires using the word *especially* (e.g., "What is something you especially like?").

Direct Instruction of Content Words

Teachers introduced two content words during each lesson through direct instruction using picture cards. The picture cards included a picture of the target word on the front that illustrated the meaning of the word. Pictures were carefully selected for abstract words to ensure that they provided the context needed to support students' understanding of word meanings. On the back of each card was a script for teachers that described how the picture illustrated the word's meaning. For example, for the word *evidence*, an illustration of a boy with chocolate on his face appeared on the front of the card and the following script on the back: "The chocolate on the boy's face is evidence that he ate the candy bar." The script on the back of the card also defined the word in English and Spanish, noted if the word is a cognate, modeled the target word's use in a different context, and provided opportunities for students to work in pairs to create new sentences using the target word. Students then spelled the word aloud. A glossary of all words in each unit also was provided for students to support their learning during the unit.

Direct Instruction of Connecting Words

Instructing connecting words adjusted and expanded on the strategies used to teach content words. To teach connecting words, teachers also used picture cards and Spanish translations, but unlike the cards used for teaching content words, the pictures displayed two scenarios from the read-aloud book. The sentences were linked by the target connecting word (e.g., "Teresa is sick because she ate too many tamales"). The instructional routine involved the teacher reading an introductory sentence related to this scenario and combining additional example sentences that were not necessarily from the book using the target connecting word, individual students combining sentences using the same connecting word, and students pairs physically acting out the word's meaning and generating and writing additional sentences using the connecting word.

Instruction in Word-Learning Strategies

At the beginning of the intervention, a minilesson introduced students to cognates. As a class, students practiced identifying English words that are cognates and generating lists of cognates. When introducing content vocabulary words, teachers noted the words that are cognates.

Activities to Foster Word Consciousness

After target words were introduced, students were prompted to listen for the words during the interactive reading session and touch their ears when they heard them. As illustrated in Appendix F, students also engaged in peer talk, classroom discussion, and individual and partner writing. Many of the reinforcement activities described next also served to build word consciousness and alert students to the importance of learning new words.

Reinforcement

Much of the vocabulary was reinforced by defining words in context during interactive shared reading. For example, to reinforce the meaning of the word *especially*, teachers engaged students in the following exchange: "Maria explains that she liked wearing her mother's lipstick and apron, but she especially wanted to wear the ring. Remember, *especially* means very much." Each unit closed with a quiz on all target words.

Words previously instructed were also reinforced through multimodal and playful opportunities for oral and written language production, such as acting out scenarios, singing songs, role-playing, and games. Examples include the following:

- Students were given examples and nonexamples of content word meanings and had to indicate with a thumbs-up which were examples, as in the following exchange: "If I name an animal that is especially big, put your thumbs up and say 'especially big.' If not, put your thumbs down and say 'especially small.' Whale [pause], mouse [pause], elephant [pause], hummingbird [pause]."

- To reinforce the meaning of *because*, students pierced a balloon with a pin, and when it exploded, they described what happened: "The balloon pops because we stick a pin in it!"
- To reinforce the meanings of the content words *pace*, *reaction*, and *dreadful*, students listened to a recording of a song called "My Dog Steals Food" that includes the three words, sang the song several times, and then answered questions about the words.
- To reinforce the meaning of the connecting words *because* and *meanwhile*, students acted out a play (based on the *Too Many Tamales* book) called *The Cousin Says No* that includes the two words.
- Students played bingo and completed crossword puzzles that were composed of lesson content words.
- For each unit, there were student home connection packets with engaging activities that students could complete at home to reinforce word meanings (e.g., clip magazine pictures that represent target words, select a favorite word from a list of unit target words and explain why).

Supports for ELs

Visual and linguistic supports were used to support ELs' receptive and productive oral language. Examples of visual supports included pictures, realia, gestures, and actions. Examples of linguistic supports included the teacher defining words in context, providing examples, and modeling. Linguistic supports also included the use of Spanish to clarify and elaborate on word meanings presented in English. For example, the first person to use Spanish was the teacher, offering Spanish translations for all the newly taught English words. Although students were encouraged to use these words in English, Spanish was always welcome in the classroom to support EL engagement and participation.

Professional Development

Prior to the start of the intervention, treatment teachers received a one-day training to prepare them to implement the intervention. During the training session, teachers spent the first half of the day learning about the intervention. In the second half of the day, teachers practiced teaching the first lesson while mentors provided one-on-one instruction and support. At the close of the study, the control teachers were provided a set of all intervention materials and training on how to use them.

Coaches who were English–Spanish bilingual and master-level teachers with extensive experience in classroom mentoring met with the treatment teachers every two weeks to model lessons and observe and provide feedback on teachers' lesson implementation. During these observation periods, the master teachers completed a standardized observation protocol to provide feedback to the teachers.

Measures and Protocols

Three assessments of English vocabulary were given to participants as pretests and as posttests immediately following the intervention by trained research assistants. The assessments were one standardized assessment—the fourth edition of the Test of Oral Language Development–Primary (TOLD; Hammill & Newcomer, 2008)—and two researcher-developed curriculum-based measures: the English Learner's Vocabulary Assessment (ELVA) and the English Learner's Connectives Assessment (ELCA). The same random sample of ELs who had been assessed on the TOLD were also assessed on the ELVA and TOLD 10 months after the intervention was completed. Delayed ELVA posttests were composed of a random sample of 24 items stratified by content word type.

Standardized Language Assessment

The oral vocabulary subtest of the TOLD was used as a pretest, posttest, and delayed posttest to assess productive oral vocabulary. The test was individually administered to a randomly selected subset of 14 students in each class for a total of 296 students (treatment $n = 167$; control $n = 129$). During the test, teachers said a word and asked students to tell the test administrator what it means. Examiners recorded students' responses verbatim. Responses were scored according to the TOLD scoring rubric that allows examiners to evaluate each response to determine whether it is correct or incorrect (Hammill & Newcomer, 2008).

All subtests of the TOLD have been documented as reliable and valid (Spenciner, 2017). The average coefficient alpha for the composite is .90 or greater, suggesting that the assessment is internally consistent. Test–retest reliability indicates that the majority of correlation coefficients for a sample of students retested after two weeks falls between .80 and .90.

The ELVA

The ELVA is a researcher-designed, paper-and-pencil test that measures students' receptive knowledge of intervention content words (i.e., nouns, adjectives, verbs). The ELVA consists of four equivalent forms, each with 11 multiple-choice items. The division of the assessment into four equivalent forms enables it to be given over multiple days to avoid student fatigue. All students take all four forms.

Words selected for the ELVA were drawn from the 96 words used in the intervention. The words selected for

each item were balanced for Living Word Vocabulary level, frequency in terms of *U* value reported in *The Educator's Word Frequency Guide*, part of speech, and word type (e.g., abstract cognate). The word types were assigned to placement in the assessment so each word type appeared at equal intervals across assessment.

Each item in the student booklet of the ELVA consists of a picture that illustrates the meaning of a target word. The pictures were carefully chosen to represent the target word meaning as used in the intervention. For example, for the target word *responsibility* and meaning important job, there is a picture of two children bathing a dog. Three written words appear underneath the picture; one of the words is the target word. While students are looking at the item, the teacher describes the picture, provides a student-friendly definition of the target word, and asks the students to fill in the bubble for the word that matches the picture, sentence, and definition. To investigate reliability of the assessment, coefficient alphas were computed (pretest ELVA $\alpha = .78$; posttest ELVA $\alpha = .91$). Appendix G displays ELVA administrator instructions and sample items.

The ELCA

The ELCA is a researcher-designed, multiple-choice, paper-and-pencil test that measures students' knowledge of connecting words at the sentence level. The development of the ELCA was informed by the text cohesion task designed by Crosson et al. (2008) for fourth-grade students. In contrast to this, the ELCA was designed for second-grade students and is supported by pictures to increase engagement and support comprehension.

In the ELCA, students listen to two sentences and are asked to select, from among three options, the most appropriate connecting words to link two sentences. The test consists of 16 multiple-choice items, four per form. The forms are equivalent. All students take all four forms, each on a separate day, as was the case with the ELVA.

To investigate reliability of the assessment, coefficient alphas were computed (pretest ELCA $\alpha = .48$ for students in the treatment and control groups; posttest ELCA $\alpha = .66$). Although the pretest alphas are low, they are expected to be low because students are unlikely to know many of these connecting words prior to intervention. Reliabilities were higher at posttest for treatment students ($\alpha = .71$). Students' raw mean scores for the total assessment were used in modeling. Appendix H displays ELCA administrator instructions and sample items.

Observation Protocol

An observation protocol was used to collect data on fidelity and quality of implementation in treatment classrooms. Appendix I displays the observation protocol. Observations were conducted by the mentors associated with each teacher. The mentors were used to document fidelity because they were most familiar with the intervention curriculum and with the teachers. The curriculum was scripted, and the rating protocol was detailed to help ensure objectivity of raters.

Fidelity ratings were made separately for specified activities that occurred during each component of the intervention (content and connecting vocabulary instruction, interactive shared reading, and content and connecting vocabulary enrichment activities). For each specified activity, fidelity was rated on a 1–4 scale (1 = *no intervention activities*; 4 = *all intervention activities*). Activities not part of the Acquiring Vocabulary in English intervention were noted. In treatment classrooms, pacing, classroom management, and quality of teacher–student interactions were rated on a scale ranging from 1 (*poor*) to 3 (*excellent*). Data on fidelity of implementation were collected twice from each teacher during the intervention. The first observation took place during Unit 2 (weeks 3 and 4 of the intervention), and the second observation took place during Unit 7 (weeks 13 and 14).

Analytic Procedures

In all models, we fitted a random intercept with classrooms at the cluster level. For RQ1a, RQ1b, and RQ2, we analyzed data using SAS PROC MIXED (SAS version 9.4; SAS Institute, 2014). To investigate content vocabulary learning in RQ1a and RQ1b, we fitted a mixed-models ANOVA with performance on each type of content word (i.e., concrete cognates, concrete noncognates, abstract cognates, abstract noncognates) predicted by treatment, wave (wave 1 = pretest; wave 2 = posttest), cognate status, abstractness designation, and the interactions of these variables. To investigate connecting words in RQ1, we fitted a mixed-models ANOVA with overall connecting word performance predicted by treatment, wave, and treatment-by-wave interaction. The analyses for RQ2 were similar, with wave and treatment and the interaction predicting general vocabulary knowledge as measured by the TOLD. Finally, for RQ3, we fitted an explanatory item response theory model in SAS PROC GLMMIX with content vocabulary item performance predicted by treatment, wave, and word characteristics (cognate status and abstractness designation), plus the interactions of these variables. For all analyses, effect sizes are considered as follows: Cohen's $d = 0.2$ is a small effect size, 0.5 is a medium effect size, and 0.8 is a large effect size (Cohen, 1992).

Results

Fidelity of Implementation

For each specified activity, fidelity was rated on a 1–4 scale (1 = *no intervention activities*; 4 = *all intervention*

activities). Overall, fidelity was high, increasing numerically from the first observation to the second. The first and second overall fidelity ratings were 3.25 and 3.59. Fidelity by component and by time of observation were as follows for time 1: Content vocabulary instruction = 3.38, interactive shared reading = 3.32, content vocabulary enrichment = 3.26, connecting vocabulary instruction = 3.53, and connecting vocabulary enrichment = 2.69. For time 2, they were as follows: Content vocabulary instruction = 3.70, interactive shared reading = 3.39, content vocabulary enrichment = 3.61, connecting vocabulary instruction = 3.81, and connecting vocabulary enrichment = 3.45. Appendix J displays fidelity ratings by teacher and component for content and connecting word instruction.

In treatment classrooms, teachers were also rated on quality of implementation on a scale ranging from 1 (*poor*) to 3 (*excellent*). Average quality of implementation ratings were above 2.5 on all indicators at times 1 and 2, respectively in the following: Pacing = 2.53 and 2.73, classroom management = 2.63 and 2.68, and teacher–student interactions = 2.62 and 2.73. Appendix K displays the average quality of implementation ratings by component and time of observation. Analyses indicated that fidelity was not a statistically significant predictor of content and connecting vocabulary gains (time 1 $p = .788$; time 2 $p = .958$), likely caused by the high levels of fidelity with little between-teacher variability.

RQ1a: The Effect of the Intervention on ELs' English Content Word and Connecting Word Outcomes

To investigate the effects of the intervention on ELs' learning of English content and connecting words, we fitted random intercepts models with classroom performance at the between level, and scores on content and connecting words and the four content word groups, repeated within student. Word group scores were predicted by treatment status, wave, cognate status, and abstractness designation (abstract or not abstract). In examining our model results for content words, we also found that the interaction between treatment and wave was statistically significant. The full model results are provided in Table 1.

To follow up on this interaction, we examined the least square means for pretest and posttest for the treatment and control groups. These means can be seen in Table 2 and are plotted in Figure 1. As can be seen in both the table and the figure, the groups were fairly comparable at pretest, with the control group answering 46% of the content vocabulary items correctly and the treatment group answering 45% correctly. However, at posttest, the treatment group answered 81% of the items correctly, whereas the control group answered only 56% correctly. The 26 percentage-point differential gain between treatment and control groups was statistically significant and constituted a standardized mean difference of 1.88.

With regard to connecting words, the model results were similar, with a statistically significant wave-by-treatment interaction (see Table 1). Again, we examined the least square means (see Table 2) and found that the groups were comparable at pretest, with both the treatment and control groups answering 48% of the connecting word items correctly. At posttest, the treatment group answered

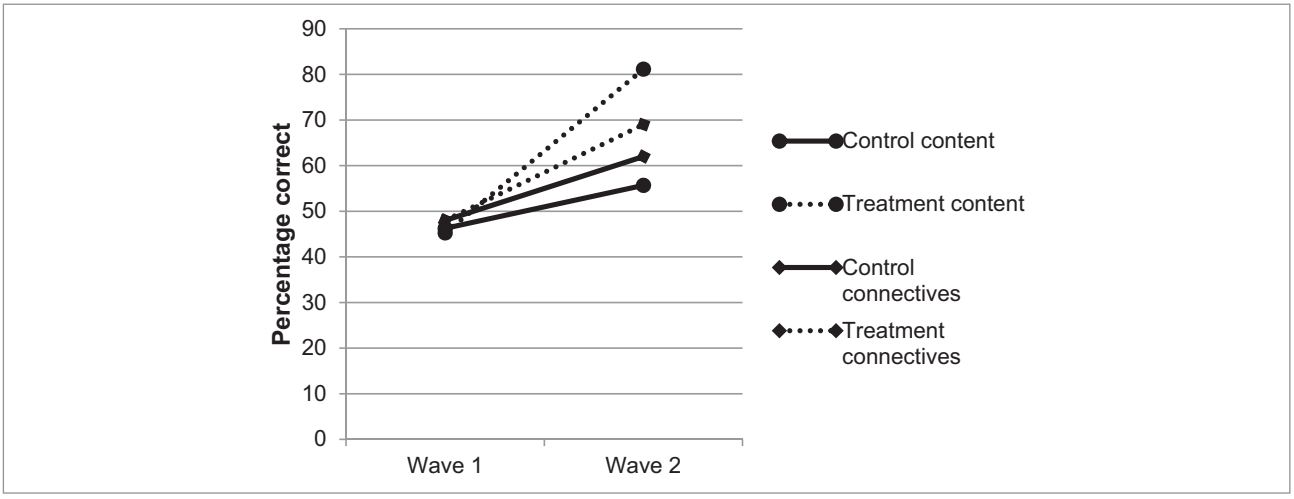
TABLE 2
Least Square Means for Treatment Group and Wave: Content and Connecting Vocabulary

	Content words		Connecting words	
Group	Wave 1	Wave 2	Wave 1	Wave 2
Control	0.46	0.56	0.48	0.62
Treatment	0.45	0.81	0.48	0.69

TABLE 1
Fixed Effects for the Multilevel Model of Content and Connecting Vocabulary Performance Predicted by Treatment, Wave, Cognate Status, and Complexity

		Degrees of freedom			
Outcome	Effect	Numerator	Denominator	*F*	*Pr > F*
Content words	Treatment	1	2,818	30.08	<.0001
	Wave	1	2,818	1,764.54	<.0001
	Cognate status	1	2,818	270.34	<.0001
	Complexity	1	2,818	102.59	<.0001
	Treatment × Wave	1	2,818	601.67	<.0001
Connecting words	Treatment	1	691	2.18	.140
	Wave	1	691	270.70	<.001
	Treatment × Wave	1	691	12.18	<.001

FIGURE 1
Pre—Post Gains for Content and Connecting Vocabulary by Treatment



69% of the connecting word items correctly, whereas the control group answered 62% correctly. The 7 percentage-point differential gain between treatment and control groups was statistically significant and constituted a standardized mean difference of 0.41. See Appendix L for additional information about the models used to investigate all of the research questions.

RQ1b: The Effect of the Intervention on ELs' English Content Vocabulary by Word Characteristics

To investigate whether the effect of the intervention differed across the four types of content words, we examined specific contrasts between the treatment and control groups' pre–post gains by content word type. These results are provided in Table 3 and indicate that there were statistically significant mean differences in the magnitude of the gains between the treatment and control groups for each content word group. The treatment group made gains relative to the control group that were greater by 34% for concrete noncognates (1.57 standardized mean difference) 33% for abstract noncognates (1.57 standardized mean difference), 23% for concrete cognates (1.02 standardized mean difference), and 17% for abstract cognates (0.81 standardized mean difference). The least square means tested in these contrasts are shown in Table 4 and plotted in Figures 2a–d. The table and figures show that for all word types, the treatment group made greater gains than the control group.

Additionally, when excluding level of abstractness, we found a statistically significant Time × Treatment × Cognate Status interaction. Control students' cognate and noncognate gains were 12% and 6%, respectively, whereas treatment students' cognate and noncognate gains were 32% and 39%, respectively. The difference between cognate and noncognate gains is statistically significant for both groups but in different directions, control $t(172) = 4.35$, $p < .001$ in favor of cognates; treatment $t(208) = -5.56$, $p < .00$ in favor of noncognates.

RQ2: The Effect of the Intervention Beyond the Words Taught

To investigate whether the effects of the intervention generalized beyond words taught, we administered the oral

TABLE 3
Follow-Up Contrast for Content Word Types on the English Learner's Vocabulary Assessment

Label	Estimate	Error	Degrees of freedom	*t*-value	*Pr* > |*t*|	Hedges' *g*
Abstract cognates	0.1715	0.03951	1,354	4.34	<.0001	0.81
Abstract noncognates	0.3303	0.03951	1,354	8.36	<.0001	1.44
Concrete cognates	0.2257	0.03951	1,354	5.71	<.0001	1.02
Concrete noncognates	0.3383	0.03951	1,354	8.56	<.0001	1.57
Connecting words	0.07	0.03	325	2.15	.032	0.41

TABLE 4
Least Square Means for Treatment Group and Wave by Content Word Type on the English Learner's Vocabulary Assessment

Word type	Group	Wave 1	Wave 2
Content words			
• Concrete cognates	Control	0.4475	0.5835
	Treatment	0.4489	0.8103
• Concrete noncognates	Control	0.3784	0.4753
	Treatment	0.3867	0.8202
• Abstract cognates	Control	0.5679	0.6809
	Treatment	0.5476	0.8318
• Abstract noncognates	Control	0.4563	0.4877
	Treatment	0.4272	0.7839
Connecting words	Control	0.48	0.62
	Treatment	0.48	0.69

vocabulary subtest of the TOLD, our standardized oral language assessment. We fitted a random intercepts model with classroom performance at the between level, and raw scores from the TOLD oral vocabulary subtest, repeated within student, being predicted by group and wave.

The model parameters can be interpreted in the same fashion as in the previous subsection. As indicated in Table 5, the results for the model of performance on the TOLD oral vocabulary subtest predicted by group and wave revealed a statistically significant treatment-by-wave interaction, which corresponds to a standardized mean difference of 0.47. In addition, we examined the least square means (see Table 6 and Figure 3). The table indicates that despite randomization, students in the treatment group started out lower on the TOLD as compared with the control group (13.7 vs. 15.9 items correct). However, by the end of the intervention, the treatment group caught up to the control group on the TOLD (18.5 vs. 18.7 items correct). This differential pretest performance and growth by the treatment group can be seen in Figure 3, which indicates that the treatment group was below the control group at wave 1 but was effectively equal to the control group at wave 2.

RQ3: The Effect of the Intervention Over Time

To investigate whether students retained content words learned in the intervention, we fitted an explanatory item response theory mixed model with classroom at the between level, and item-level performance being predicted by wave and treatment. The explanatory item response theory model was similar to the mixed models presented earlier, except that instead of modeling the outcome as student means on each word type, we modeled the outcome at the level of the item because the delayed follow-up assessment used a subset of the items on the ELVA. As such, results are directly comparable across the waves when modeling common items at the item level. The model still nests students within classrooms, but instead of nesting word group scores within student as was done previously, items are nested within student, thus making this a logistic model.

Once again, the research question focused on whether the coefficient associated with the wave-by-treatment interaction was statistically significant. Because the delayed posttest occurred during a period when intervention instruction was not occurring, we did not expect a linear increase in vocabulary knowledge based on time. Therefore, we chose to treat wave as a class variable and model an effect for wave at each timepoint.

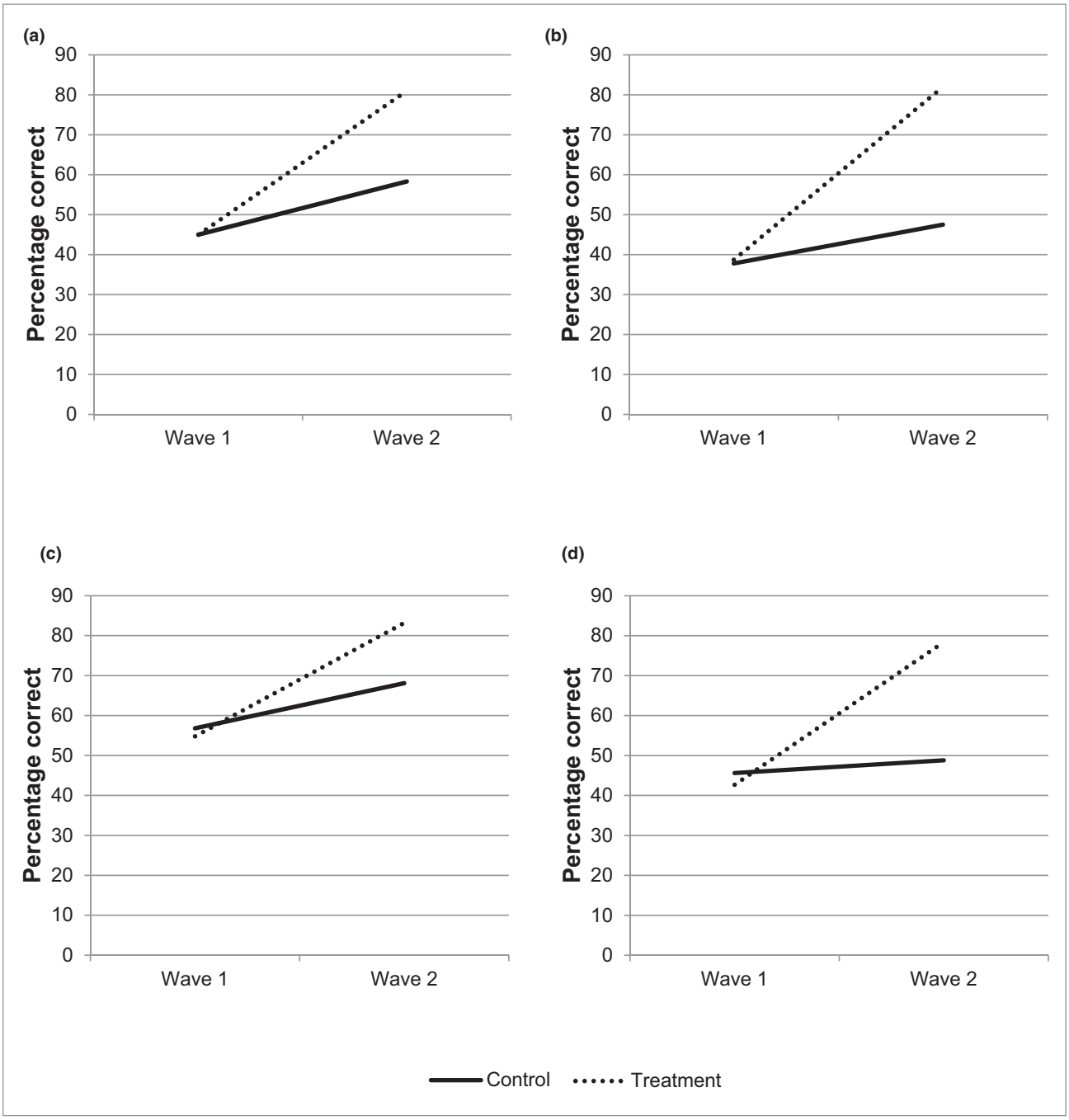
Full model results can be seen in Table 7, least square means for the percentage correct in Table 8, and the plot of these means in Figure 4. As displayed in the model results, there was a statistically significant wave-by-treatment interaction on the ELVA. In examining the least square means, after 10 months, the treatment group was not performing as well as they were at the conclusion of the intervention, but they were still performing substantially higher than the control group. We examined these two potential effects by comparing the treatment group's performance at wave 2 with their performance at wave 3, as well as by comparing the treatment and control groups at wave 3. We found both comparisons to be statistically significant ($t = 8.52$, $p < .001$, $d = 1.88$; and $t = 4.52$, $p < .001$, $d = 1.31$, respectively). Data indicate some drop-off as measured by the ELVA, our curriculum-based measure of content vocabulary, once the intervention ended, but treatment group students were still outperforming the control group students nearly 10 months after the end of the intervention.

We also examined delayed posttest results for the TOLD oral vocabulary test. Just as we did with the ELVA, we fitted a random intercepts model with classroom performance at the between level, and raw scores from the TOLD oral vocabulary subtest, repeated within student, being predicted by group and wave. Again, we examined whether the coefficient associated with the wave-by-treatment interaction was statistically significant. For the delayed posttest results, the treatment-by-wave interaction was statistically nonsignificant, $F(2, 626) = 1.65$, $p = .192$.

Discussion

In this study, we explored the efficacy of an intervention designed to improve second-grade Spanish-speaking ELs'

FIGURE 2
Pre—Post Gains by Treatment for (a) Concrete Cognates, (b) Concrete Noncognates (c) Abstract Cognates, and (d) Abstract Noncognates



knowledge of challenging, high-utility English vocabulary; whether the intervention had a differential effect on different types of words; and whether the effects of the intervention lasted across time. Developing the English oral proficiency of emergent bilingual students is very important because of the association among vocabulary, word reading, and comprehension. This study builds on findings from prior research syntheses (Baker et al., 2014; Foorman et al., 2016) and individual studies indicating that teaching a set of academic vocabulary words intensively using a variety of techniques is effective (August et al., 2016, 2018; Carlo et al., 2004; Crevecoeur et al., 2014; Lesaux et al., 2010, 2014; Proctor et al., 2011, 2020; Vadasy et al., 2015).

Our first question examined the effect of the intervention on ELs' knowledge of content and connecting words. Findings on proximal curriculum-based measures of learning content words (ELVA) and connecting words

TABLE 5
Fixed Effects for the Multilevel Model of the Test of Oral Language Development—Primary Oral Vocabulary Subtest Predicted by Treatment and Wave

	Degrees of freedom			
Effect	Numerator	Denominator	*F*	*Pr* > *F*
Group	1	499	0.42	.5148
Wave	1	499	66.63	<.0001
Group × Wave	1	499	4.05	.0447

TABLE 6
Least Square Means for Treatment Group and Wave: Test of Oral Language Development—Primary

Group	Wave 1	Wave 2
Control	15.8521	18.7301
Treatment	13.7135	18.4753

TABLE 7
Fixed Effects for the Multilevel Model of English Learner's Vocabulary Assessment Content Performance at Delayed Posttest Predicted by Treatment and Wave

	Degrees of freedom			
Effect	Numerator	Denominator	*F*	*Pr* > *F*
Group	1	34,915	31.3	<.0001
Wave	2	34,915	1,022.48	<.0001
Group × Wave	2	34,915	407.96	<.0001

TABLE 8
Least Square Means for Treatment Group and Wave for Content Vocabulary on the English Learner's Vocabulary Assessment

Group	Wave 1	Wave 2	Wave 3
Control	0.466	0.5555	0.6036
Treatment	0.4549	0.8594	0.7354

(ELCA) indicated that the intervention implemented for four 50-minute lessons per week for 18 weeks was effective in helping ELs acquire high-utility, school-relevant, challenging content and connecting words. Findings also indicated that there were smaller between-group differences on connecting words as compared with content words. One plausible explanation is that less variance in scores on the connecting words assessment (standard deviation [*SD*] = 0.17 for treatment students and 0.16 for control students) as compared with scores on the content words assessment (*SD* = 0.13 for treatment students and 0.11 for control students) resulted in smaller standardized mean differences. Another plausible explanation is that connecting words are harder to learn and that, in contrast to content word instruction, intervention research related to teaching connecting words to young students is just emerging. More research may lead to a better understanding of how best to support the learning of connecting words. The first question also examined the impact of content word characteristics on content word learning. Findings revealed that the treatment group significantly outperformed the control group in all four content word groups (concrete noncognate, abstract noncognate, concrete cognate, and abstract cognate) and that the standardized mean differences in favor of the treatment group were largest for noncognates (both concrete and abstract) as compared with cognates.

FIGURE 3
Pre—Post Gains for the Test of Oral Language Development—Primary by Treatment

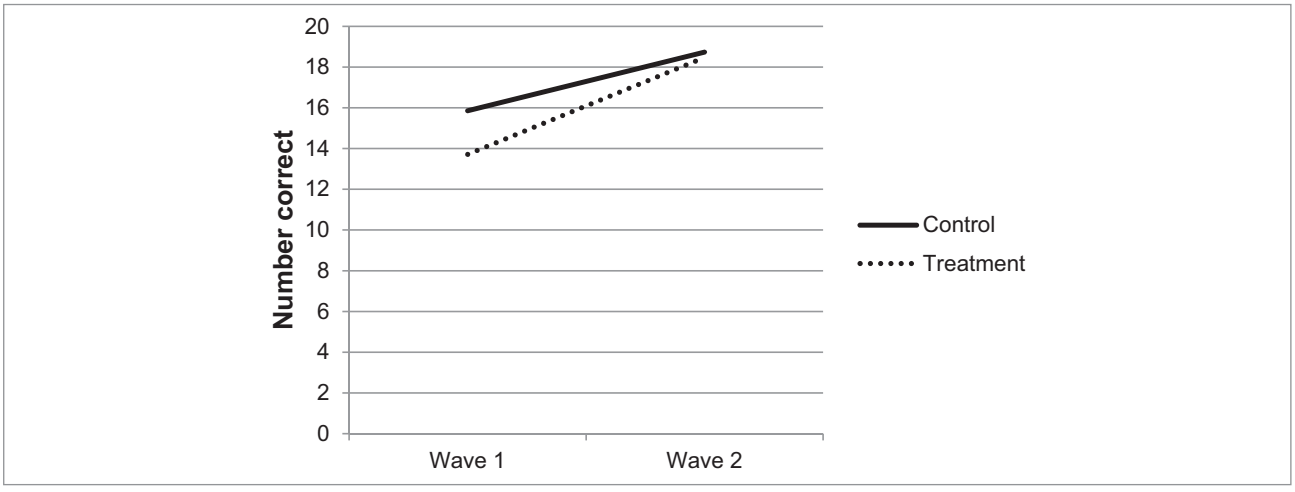
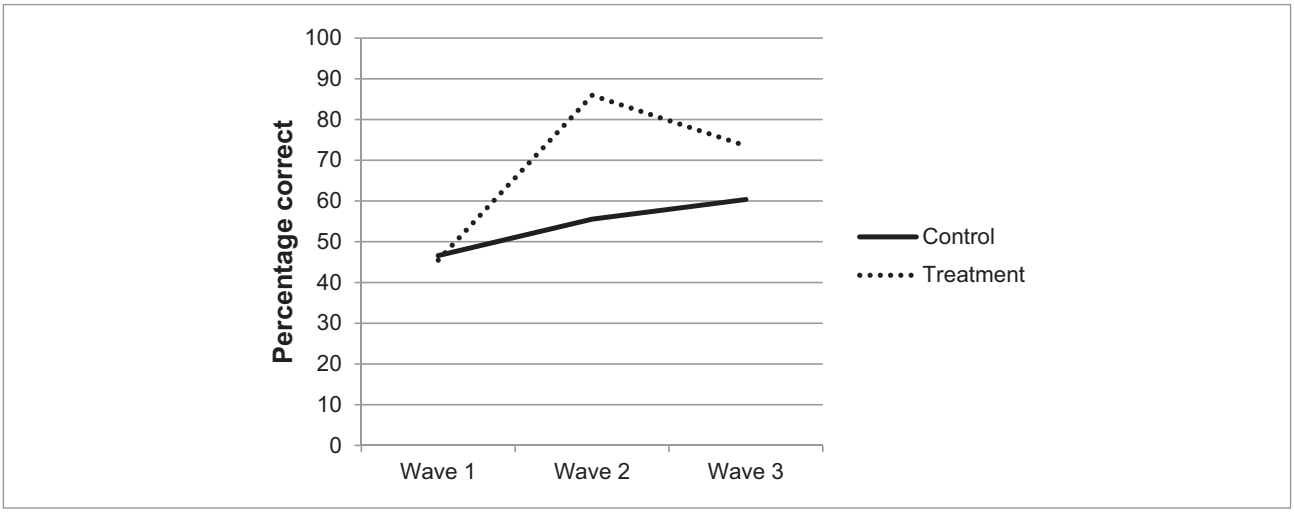


FIGURE 4
Scores by Wave and Treatment for Content Words



The second question investigated whether the effect of the intervention generalized beyond the words taught. Findings indicated that gains on a standardized measure of oral vocabulary (TOLD) were significantly greater for students in the treatment group as compared with students in the control group, as indicated by a standardized mean difference of 0.47.

Finally, the third question investigated the effects of the intervention over time. Findings on a curriculum-based assessment indicated that after 10 months, the treatment group did not perform as well on the delayed posttest as they did on the assessment taken immediately after the intervention, but they still performed better than the control group ($t = 4.52$, $p < .001$).

With regard to the TOLD, the standardized oral vocabulary assessment, at the beginning of the intervention, treatment students were performing significantly worse than control students on the TOLD. At the end of the intervention, treatment students' significantly greater gains on the TOLD, as compared with control students' gains, caught them up. Ten months after the intervention, there were no treatment control differences on the TOLD, $F(2, 626) = 1.65$, $p = .192$, indicating that treatment students had not lost ground in general vocabulary knowledge as compared with control students.

Results from this study help validate a multifaceted approach to vocabulary instruction that includes interactive shared reading, direct instruction of individual content and connecting words, instruction in word-learning strategies, development of word consciousness, reinforcement through multimodal and playful language production opportunities, and scaffolding through visual and linguistic supports. The intervention was implemented with a high level of fidelity that we attribute to the intervention curriculum and mentors in the school who provided ongoing support to help ensure that the intervention was used properly.

Vocabulary instruction: A Focus on Challenging, High-Utility Words

Content words selected for instruction were of high utility for understanding texts. They appear frequently in texts that students were reading or listening to at their particular grade level or higher grade levels, were unknown to most students at this particular grade level, and were less likely to be acquired without instruction. Findings indicate that this can indeed be an effective starting place as early as second grade if coupled with the methods used in this intervention. As highlighted by Hiebert, Scott, Castaneda, and Spichtig (2019), elementary school core reading programs and, to some extent, vocabulary instruction more broadly "lack an underlying framework for word selection" (p. 2) and, relatedly, tend to focus on easy-to-learn words that do not sufficiently prepare students for the challenges of late elementary reading, especially those students who depend heavily on school to become literate in English.

Feasibility of the Approach for Vocabulary Learning

We interpret the findings of this intervention study as generating practical implications at least at two levels. First, the Acquiring Vocabulary in English intervention could be implemented as designed during designated English-language development instruction. At a more general level, these findings can inspire other intervention researchers to pursue the goals of this study, not by following these units exactly but instead by refining and enhancing them in multiple ways (e.g., applying some of

the methods to language arts or other content area curricular materials already in use in classrooms, adding family supports, developing similar initiatives for other languages besides Spanish), or by using the instructional design principles to develop additional approaches.

Understanding Vocabulary Learning: The Role of Content Word Characteristics

At a general level, the findings demonstrate that through the strategic selection of unfamiliar high-utility content words, this research-based intervention significantly contributed to vocabulary learning that would not occur through mere exposure for young ELs. Consistent with prior research findings, in this study, we found that cognate status (Nagy et al., 1993) and word meaning abstractness (Hiebert et al., 2019) related to word difficulty for ELs. At pretest, both treatment and control students performed 9% better on words that were cognates as compared with words that were not cognates, $F(1, 1{,}584) = 164.82$, $p < .001$. The same was the case for words that were concrete. At pretest, treatment and control students performed 8% better on concrete words as compared with words that were abstract, $F(1, 1{,}584) = 139.54$, $p < .001$.

Additionally, as described in the Results section, when excluding level of abstractness, there was a statistically significant Time × Treatment × Cognate Status interaction. Control students' cognate and noncognate gains were 12% and 6%, respectively. The pattern for control students fits the theory that emergent bilingual students make greater gains on cognates than noncognates likely because of these students' ability to draw on their Spanish word knowledge (Dressler et al., 2011; Jiménez et al., 1996; Nagy et al., 1993; NASEM, 2017). However, the pattern of gains is the opposite for the treatment students, whose cognate and noncognate gains were also statistically significant but in the opposite direction: 32% and 39%, respectively. One obvious reason for this is that the intervention supported treatment students in learning both types of words; additionally, treatment students were much closer to ceiling on cognates at posttest (mean = 0.80, $SD = 0.2$), whereas control students were not (mean = 0.64, $SD = 0.20$). Findings from recent systematic analyses of the relation between word characteristics and students' vocabulary development (August et al., 2018; Hiebert et al., 2019) and the present study strongly support the further examination of word characteristics as a promising research path to continue to inform the selection of words for instruction and the tailoring of instructional methods to word type.

The Importance of Teaching Connecting Words

An especially innovative feature of this intervention was that it targeted not only content words but also connecting words. Even in recent comprehensive efforts to measure vocabulary knowledge to inform evidence-based instruction, such as the innovative InSight and VASE (Vocabulary Assessment Study in Education) vocabulary assessments (for more information, see Hiebert et al., 2019), the focus on vocabulary instruction and assessment continues to be on content words (i.e., nouns, verbs, adjectives, adverbs). The results of this intervention are promising because they reveal a positive effect of the intervention on knowledge of connecting words, even for students as young as second graders. Results are particularly encouraging given the documented gap in knowledge of connecting words for elementary school Spanish-speaking ELs, the contribution of knowledge of connecting words to reading comprehension above and beyond word recognition skills and vocabulary knowledge, and the high frequency of connecting words documented in elementary school texts (Crosson et al., 2008; Crosson & Lesaux, 2013; Nair, 2007). Combined with the fact that they represent a relatively closed class of words (i.e., a limited set that could plausibly be taught in its entirety), all of these factors make connecting words a promising, high-leverage domain for vocabulary instruction that might lead to advances in text comprehension.

The challenge of teaching connecting words stems from their lack of lexical meaning, which requires that their instruction incorporates not only previously recommended strategies, such as paraphrasing or synonyms, use of multiple examples, and sentence combining through writing (Crosson & Lesaux, 2013; Proctor, et al., 2011), but also additional strategies. The intervention teachers adjusted the routines used to teach content words to support the learning of connecting words. For example, for each connecting word, the routine involved the teacher displaying two pictures that represented scenarios from the books while concurrently reading an introductory sentence related to this scenario that used the connecting word and engaged students in playful, kinesthetically based demonstrations of the function of the connecting word.

This study, however, represents only an initial and encouraging attempt at teaching connecting words. Additional research on how learning connecting words can be best supported via intervention across grades for monolingual and multilingual learners, as well as how their use in increasingly complex texts across disciplines affects these learners' understanding, can inform evidence-based and strategic vocabulary instruction. Similar to the analysis by content word type and informed by prior research (Crosson et al., 2008), future research needs to explore whether the impact of an intervention focused on connecting words may vary by word characteristics (e.g., word frequency level, conceptual relation), as well as its potential impact on participants' reading comprehension.

Study Strengths, Limitations, and Future Directions

As is the case for other intervention studies that we conducted (August et al., 2016, 2018), strengths of the study include implementation in settings typical of many classrooms in which ELs are educated: schools with high percentages of students from low-income families, teacher–student ratios that exceed 20 students per teacher, and classroom teachers (rather than researchers) implementing the intervention. An additional strength is that intervention methods are aligned with college and career readiness standards (e.g., National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). Teachers engaged students in discussion of grade-level texts through answering sequences of text-dependent questions and acquiring vocabulary in context.

As was the case for many studies focused on vocabulary (NASEM, 2017), whereas the methods of instruction used in the present study resulted in statistically significant word learning, the volume of words learned may be too limited to support ELs' understanding of grade-level texts. In the case of this intervention, students were exposed to 96 content words and 12 connecting words over 18 weeks. Given the number of words that ELs need to acquire to understand and read grade-level text, it will be important to identify and assess effectiveness of other methods to increase the word learning of ELs.

One strategy is to support teachers in using similar techniques throughout the day that provide opportunities to engage in practices that target specific words and practices that promote oral language more generally, such as engaging students in multiturn conversations, providing opportunities for student-initiated conversations, and teacher use of elicitations and extensions (Cabell et al., 2015). Allowing for student practice in using words learned in context, cumulative review, and spaced retrieval would also be important. Another method might be to match the intensity of method to word type; words that are more easily acquired (e.g., those that are more concrete, cognates for students with primary languages other than English) might be defined in context rather than instructed using pictures and more elaborate routines. A third method might be to teach students word-learning strategies to make them more autonomous vocabulary learners. Such strategies include capitalizing on morphology and cognate knowledge and using context and digital dictionaries (e.g., August et al., 2016, 2018; Lesaux et al., 2014). It would be important for these investigations to consider the ability of teachers to enact these practices, as well as the interaction of these approaches with word types and student characteristics, such as students' levels of English or Spanish proficiency.

It would be important to collect both qualitative and quantitative data on how well these methods work to develop oral language proficiency in ELs. For example, studies might explore how teachers adapt the intervention to their particular context and how well they were able to implement it on their own following some amount of initial training and mentoring. Also, future studies might include additional curriculum-based measures of vocabulary to assess students' productive and receptive vocabulary knowledge and their depth of word knowledge for content words. Assessing students' Spanish word knowledge of instructed English words prior to and following an intervention would also help inform findings related to cross-language transfer for young students. Language sample analysis of students' oral responses to comprehension questions and oral retellings of stories that have not been taught might be used to determine whether multifaceted methods such as those used in the current study improve other components of students' oral language besides vocabulary, such as morphology, syntax, and narrative structure (Miller, Andriacchi, & Nockerts, 2016). An advantage of using language sample analysis of oral story retells is that students' responses can be coded for responses in Spanish, English, or Spanish and English concurrently.

In summary, this study makes a valuable contribution to the field by validating the efficacy of a multifaceted approach, implemented in a real-life context, on the acquisition of word knowledge in young Spanish-speaking ELs. The routines that compose the lessons (e.g., direct instruction of content and connecting words through picture cards and definitions in context, interactive shared reading, playful multimodal reinforcement) can easily be adapted to other teaching contexts because they are consistent with methods that teachers already use in many classrooms (e.g., storybook reading, vocabulary instruction), but are more strategic and instructionally intense. Future studies should build on this research by exploring the efficacy of methods that support the language learning of young ELs that are more efficient and by using additional measures of ELs' oral language.

NOTES

This work was supported by a grant (HD039530) from the Eunice Kennedy Shriver National Center for Child Health and Human Development at the National Institutes for Health. We acknowledge the support of Iris Gutierrez and Edna Navarro, research assistants at the University of Houston; Lindsey Massoud, a research assistant at the Center for Applied Linguistics; Franne Rosenthal, a codesigner of the connective intervention at Harvard University; and Donna Christian, the president of the Center for Applied Linguistics at the time of this study who provided research guidance and review.

REFERENCES

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. https://doi.org/10.1080/10573569.2015.1039738

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. https://doi.org/10.1007/s11145-018-9821-1

Baker, S., Lesaux, N., Jayanthi, M., Dimino, J., Proctor, C.P., Morris, J., … 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, Institute of Education Sciences, U.S. Department of Education.

Barr, C.D., Uccelli, P., & Phillips Galloway, E. (2019). Specifying the academic language skills that support text understanding in the middle grades: The design and validation of the core academic language skills construct and instrument. *Language Learning*, *69*(4), 978–1021. https://doi.org/10.1111/lang.12365

Beck, I.L., & McKeown, M.G. (2001). Text talk: Capturing the benefits of read-aloud experiences for young children. *The Reading Teacher*, *55*(1), 10–20.

Beck, I.L., & McKeown, M.G. (2007). Increasing young children's oral vocabulary repertoires through rich and focused instruction. *The Elementary School Journal*, *107*(3), 251–271. https://doi.org/10.1086/511706

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

Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology*, *98*(1), 44–62. https://doi.org/10.1037/0022-0663.98.1.44

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), 498–520. https://doi.org/10.1037/0022-0663.93.3.498

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. https://doi.org/10.3758/s13428-013-0403-5

Cabell, S.Q., Justice, L.M., McGinty, A.S., DeCoster, J., & Forston, L.D. (2015). Teacher–child conversations in preschool classrooms: Contributions to children's vocabulary development. *Early Childhood Research Quarterly*, *30*(1, Pt. A), 80–92. https://doi.org/10.1016/j.ecresq.2014.09.004

Carlo, M., August, D., McLaughlin, B., Snow, C., Dressler, C., Lipman, D., … White, C. (2004). Closing the gap: Addressing the vocabulary needs of English language learners in bilingual and mainstream classrooms. *Reading Research Quarterly*, *39*(2), 188–215. https://doi.org/10.1598/RRQ.39.2.3

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

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. https://doi.org/10.1080/10573569.2013.758943

Crosson, A.C., & Lesaux, N.K. (2013). Does knowledge of connectives play a unique role in the reading comprehension of English learners and English-only students? *Journal of Research in Reading*, *36*(3), 241–260. https://doi.org/10.1111/j.1467-9817.2011.01501.x

Crosson, A.C., Lesaux, N.K., & Martiniello, M. (2008). Factors that influence comprehension of connectives among language minority children from Spanish-speaking backgrounds. *Applied Psycholinguistics*, *29*(4), 603–625. https://doi.org/10.1017/S0142716408080260

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

Degand, L., & Sanders, T. (2002). The impact of relational markers on expository text comprehension in L1 and L2. *Reading and Writing*, *15*(7/8), 739–757. https://doi.org/10.1023/A:1020932715838

De Groot, A.M.B., & 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. https://doi.org/10.1111/0023-8333.00110

Dressler, C., Carlo, M.S., Snow, C.E., August, D., & White, C.E. (2011). Spanish-speaking students' use of cognate knowledge to infer the meaning of English words. *Bilingualism: Language and Cognition*, *14*(2), 243–255. https://doi.org/10.1017/S1366728910000519

Ellis, N.C., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning*, *43*(4), 559–617. https://doi.org/10.1111/j.1467-1770.1993.tb00627.x

Fillmore, L.W. (2014). English language learners at the crossroads of educational reform. *TESOL Quarterly*, *48*(3), 624–632. https://doi.org/10.1002/tesq.174

Foorman, B., Beyler, N., Borradaile, K., Coyne, M., Denton, C.A., Dimino, J., … Wissel, S. (2016). *Foundational skills to support reading for understanding in kindergarten through 3rd grade* (NCEE 2016-4008). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Freebody, P., & Anderson, R.C. (1983). Effects of vocabulary difficulty, text cohesion, and schema availability on reading comprehension. *Reading Research Quarterly*, *18*(3), 277–294. https://doi.org/10.2307/747389

Goldenberg, C. (2013). Unlocking the research about English learners: What we know—and don't yet know—about effective instruction. *American Educator*, *37*(2), 4–11.

Goodwin, A.P., & Cho, S.-J. (2016). Unraveling vocabulary learning: Readers and item-level predictors of vocabulary learning within comprehension for fifth and sixth graders. *Scientific Studies of Reading*, *20*(6), 490–514. https://doi.org/10.1080/10888438.2016.1245734

Goswami, U. (2002). Early phonological development and the acquisition of literacy. In S.B. Neuman & D.K. Dickinson (Eds.), *Handbook of early literacy research* (Vol. 1, pp. 111–125). New York, NY: Guilford.

Gough, P.B., & Tunmer, W.E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7*(1), 6–10. https://doi.org/10.1177/074193258600700104

Graves, M.F. (2015). Building a vocabulary program that really could make a significant contribution to students becoming college and career ready. In P.D. Pearson & E.H. Hiebert (Eds.), *Research-based practices for teaching Common Core literacy* (pp. 123–142). New York, NY: Teachers College Press.

Graves, M.F. (2016). *The vocabulary book: Learning and instruction* (2nd ed.). New York, NY: Teachers College Press.

Graves, M.F., August, D., & Mancilla-Martinez, J. (2013). *Teaching vocabulary to English language learners*. New York, NY: Teachers College Press.

Hammill, D.D., & Newcomer, P.L. (2008). *Test of Language Development–primary* (4th ed.). Austin, TX: PRO-ED.

Hiebert, E.H., Scott, J.A., Castaneda, R., & Spichtig, A. (2019). An analysis of the features of words that influence vocabulary difficulty. *Education Sciences*, *9*(1), article 8. https://doi.org/10.3390/educsci9010008

Jiménez, R.T., García, G.E., & Pearson, P.D. (1996). The reading strategies of bilingual Latina/o students who are successful English readers: Opportunities and obstacles. *Reading Research Quarterly*, *31*(1), 90–112. https://doi.org/10.1598/RRQ.31.1.5

Kintsch, W. (1991). The role of knowledge in discourse comprehension: A construction-integration model. *Advances in Psychology*, *79*, 107–153. https://doi.org/10.1016/S0166-4115(08)61551-4

Language and Reading Research Consortium, Jiang, H., & Davis, D. (2017). Let's know! Proximal impacts on prekindergarten through grade 3 students' comprehension-related skills. *The Elementary School Journal*, *118*(2), 177–206. https://doi.org/10.1086/694220

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. https://doi.org/10.1017/S1366728911000393

Lesaux, N., & Geva, E. (2006). Synthesis: Development of literacy in language-minority students. In D.L. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth* (pp. 53–74). Mahwah, NJ: Erlbaum.

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. https://doi.org/10.1598/RRQ.45.2.3

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. https://doi.org/10.3102/0002831214532165

Maguire, M.J., Hirsh-Pasek, K., & Golinkoff, R.M. (2006). A unified theory of word learning: Putting verb acquisition in context. In K. Hirsh-Pasek & R.M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 364–391). Oxford, UK: Oxford University Press.

Mancilla-Martinez, J. (2010). Word meanings matter: Cultivating vocabulary knowledge in fifth-grade Spanish-speaking language minority learners. *TESOL Quarterly*, *44*, 669–699. https://doi.org/10.5054/tq.2010.213782

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), 701–711. https://doi.org/10.1037/a0019135

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. https://doi.org/10.1111/j.1467-8624.2011.01633.x

Miller, J.F., Andriacchi, K., & Nockerts, A. (2016). *Assessing language production using SALT software: A clinician's guide to language sample analysis*. Middleton, WI: SALT Software.

Nagy, W.E., Anderson, R.C., & Herman, P.A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal*, *24*(2), 237–270. https://doi.org/10.3102/00028312024002237

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 Reading Behavior*, *25*(3), 241–259. https://doi.org/10.1080/10862969009547816

Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly*, *47*(1), 91–108. https://doi.org/10.1002/RRQ.011

Nair, M. (2007). *An analysis of the words appearing in middle school textbooks* (Unpublished doctoral dissertation). Harvard University, Cambridge, MA.

National Academies of Sciences, Engineering, and Medicine. (2017). *Promoting the educational success of children and youth learning English: Promising futures* (R. Takanishi & S. Le Menestrel Eds.). Washington, DC: The National Academies Press.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects*. Washington, DC: Authors.

O'Reilly, T., & McNamara, D.S. (2007). The impact of science knowledge, reading skill, and reading strategy knowledge on more traditional "high-stakes" measures of high school students' science achievement. *American Educational Research Journal*, *44*(1), 161–196. https://doi.org/10.3102/0002831206298171

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. https://doi.org/10.1037/h0025327

Perfetti, C.A., & Hart, L. (2002). The lexical quality hypothesis. In L. Vehoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189–213). Amsterdam, The Netherlands: John Benjamins. https://doi.org/10.1075/swll.11.14per

Perfetti, C.A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skills. In M.J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 227–247). Malden, MA: Blackwell.

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*(1), 22–37. https://doi.org/10.1080/10888438.2013.827687

Pretorius, E.J. (1996). A profile of causal development amongst ten-year-olds: Implications for reading and writing. *Reading and Writing*, *8*(5), 384–406.

Pretorius, E.J. (2006). The comprehension of logical relations in expository texts by students who study through the medium of ESL. *System*, *34*(3), 432–450. https://doi.org/10.1016/j.system.2006.02.003

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), 246–256. https://doi.org/10.1037/0022-0663.97.2.246

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. https://doi.org/10.1007/s11145-009-9218-2

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. https://doi.org/10.1002/j.1545-7249.2009.tb00232.x

Proctor, C.P., Silverman, R.D., Harring, J.R., Jones, R.L., & Hartranft, A.M. (2020). Teaching bilingual learners: Effects of a language-based reading intervention on academic language and reading comprehension in grades 4 and 5. *Reading Research Quarterly*, *55*(1), 95–122. https://doi.org/10.1002/rrq.258

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

Roberts, T.A. (2008). Home storybook reading in primary or second language with preschool children: Evidence of equal effectiveness for second-language vocabulary acquisition. *Reading Research Quarterly*, *43*(2), 103–130. https://doi.org/10.1598/RRQ.43.2.1

Salsbury, T., Crossley, S.A., & McNamara, D.S. (2011). Psycholinguistic word information in second language oral discourse. *Second Language Research*, *27*(3), 343–360. https://doi.org/10.1177/0267658310395851

Sánchez, E., & García, J.R. (2009). The relation of knowledge of textual integration devices to expository text comprehension under different assessment conditions. *Reading and Writing*, *22*, 1081–1108. https://doi.org/10.1007/s11145-008-9145-7

SAS Institute. (2014). SAS (Version 9.4) [Computer software]. Cary, NC: Author.

Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, *95*(1), 26–43. https://doi.org/10.1111/j.1540-4781.2011.01146.x

Scott, J.A. (2015). Essential, enjoyable and effective: The what, why and how of powerful vocabulary instruction. *Literacy Learning: The Middle Years*, *23*(1), 14–22.

Scott, J.A., Miller, T.F., & Flinspach, S.L. (2012). Developing word consciousness: Lessons from highly diverse fourth-grade classrooms. In E.J. Kame'enui & J.F. Baumann (Eds.), *Vocabulary instruction: Research to practice* (2nd ed., pp. 169–188). New York, NY: Guilford.

Scott, J.A., & Nagy, W.E. (2004). Developing word consciousness. In J.F. Baumann & E.J. Kame'enui (Eds.), *Vocabulary instruction: Research to practice* (pp. 201–217). New York, NY: Guilford.

Silverman, R.D. (2007). Vocabulary development in English-language and English-only learners in kindergarten. *The Elementary School Journal*, *107*(4), 365–383. https://doi.org/10.1086/516669

Spenciner, L.J. (2017). Test review of Test of Language Development–primary. In J.F. Carlson, K.F. Geisinger, & J.L. Jonson (Eds.), *The twentieth mental measurements yearbook*. Lincoln: Buros Center for Testing, University of Nebraska-Lincoln.

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

Stanovich, K.E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*(4), 360–407. https://doi.org/10.1598/RRQ.21.4.1

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. https://doi.org/10.1111/j.1467-9922.2009.00519.x

Uccelli, P., Barr, C.D., Dobbs, C.L., Phillips Galloway, E., Meneses, A., & Sánchez, E. (2015). Core academic language skills (CALS): An expanded operational construct and a novel instrument to chart school-relevant language proficiency in preadolescent and adolescent learners. *Applied Psycholinguistics*, *36*(5), 1077–1109. https://doi.org/10.1017/S014271641400006X

Uccelli, P., Deng, Z., Phillips Galloway, E., & Qin, W. (2019). The role of language skills in mid-adolescents' science summaries. *Journal of Literacy Research*, *51*(3), 357–380. https://doi.org/10.1177/1086296X19860206

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. https://doi.org/10.1080/19345747.2015.1033120

Welie, C., Schoonen, R., Kuiken, F., & van den Bergh, H. (2017). Expository text comprehension in secondary school: For which readers does knowledge of connectives contribute the most? *Journal of Research in Reading*, *40*(Suppl. 1), S42–S65. https://doi.org/10.1111/1467-9817.12090

Wiseman, A. (2011). Interactive read alouds: Teachers and students constructing knowledge and literacy together. *Early Childhood Education Journal*, *38*(6), 431–438. https://doi.org/10.1007/s10643-010-0426-9

Zeno, S.M., Ivens, S.H., Millard, R.T., & Duvvuri, R. (1995). *The educator's word frequency guide*. New York, NY: Touchstone Applied Sciences.

LITERATURE CITED

Soto, G. (1996). *Too many tamales*. New York, NY: Putnam & Grosset.

Submitted August 4, 2019
Final revision received May 4, 2020
Accepted May 7, 2020

DIANE AUGUST (corresponding author) was a senior research scientist at the Center for Applied Linguistics, Washington, DC, USA, at the time of this study and is now the principal educational consultant at D. August & Associates, Bethesda, Maryland, USA; email daugust@daugust-associates.com.

PAOLA UCCELLI is a professor at the Harvard Graduate School of Education, Cambridge, Massachusetts, USA; email paola_uccelli@gse.harvard.edu.

LAUREN ARTZI is a senior researcher at the American Institutes for Research, Washington, DC, USA; email lartzi@air.org.

CHRISTOPHER BARR is the director of assessment and evaluation of STEM programs in the Office of Research at Rice University, Houston, Texas, USA; email christopher.d.barr@rice.edu.

DAVID J. FRANCIS is a Hugh Roy and Lillie Cranz Cullen Distinguished University Chair and the director of the Texas Institute for Measurement, Evaluation, and Statistics at the University of Houston, Texas, USA; email dfrancis@uh.edu.

APPENDIX A

Demographic Family Information by Intervention Group

A short questionnaire was distributed to parents that included items related to the schooling history of the child, the language(s) spoken in the home, parental education, and family income. Home language use was measured on a 5-point scale (1 = *only Spanish*; 5 = *only English*). Family income was measured on a 9-point scale (1 = *less than $10,000*; 9 = *more than $80,000*) in increments of $10,000. Parental education was measured on a 15-point scale (0 = no *schooling*; 15 = *completed postgraduate degree*) in increments of schooling levels.

Variable	Level	Treatment (%)	Control (%)
Student gender	Female	48.37	42.78
	Male	51.63	57.22
Family income	Below $10,000	24.26	25.00
	$10,000–$19,999	30.69	38.41
	$20,000–$29,999	20.79	18.90
	$30,000–$39,999	14.36	11.59
	$40,000–$79,999	9.91	6.10
Mother's education	Finished elementary school	17.87	21.43
	Finished middle school	21.74	16.07
	Finished high school or GED	15.94	17.27
	Finished community college	2.42	0.60
	Finished university	2.90	1.79
	Finished graduate school	2.42	0.60
Father's education	Finished elementary school	17.71	22.38
	Finished middle school	13.14	11.19
	Finished high school or GED	17.14	13.99
	Vocational training	2.29	1.40
	Finished community college	0.57	1.40
	Finished university	2.29	0.70
	Finished graduate school	0.57	0
Mother's home language	Only Spanish	67.62	67.06
	Mainly in Spanish	18.57	17.06
	Equal	13.81	14.71
	Mainly in English	0	1.18
	Only English	0	0
Father's home language	Only Spanish	61.86	55.56
	Mainly in Spanish	19.59	16.99
	Equal	16.49	24.84
	Mainly in English	1.03	2.61
	Only English	1.03	0

Note. Due to rounding, some percentages do not add up to 100%.

APPENDIX B

Raters' Codes for Content Words' Level of Abstractness

The following table displays raters' codes for content words' level of abstractness. The rubric used to rate content words is a 5-point Likert-type scale (1 = *most concrete*; 5 = *most abstract*). Raters were instructed to rate the level of abstractness by (a) referring to the specific meaning of a word (e.g., direct as order, direct as manage), (b) imaging the word, and (c) rating the word on ease or difficulty of imaging (with ease/difficulty defined as the number of discrete mental pictures needed to image a word and, for each mental picture, the amount of context or other elements in the picture needed to image a word).

Target word	Living Word Vocabulary definition	Rating
row	To paddle a boat	1
aboard	On a ship	1
dinosaur	Animal no longer living	1
dining	Eating	2
abandoned	Gave up	3
direct	To control or manage	3
directed	Controlled or managed	3
spirit	Essential meaning or intent of a thing	4

APPENDIX C

Content Words by Word Characteristics

Abstract				Not abstract			
Not cognate		Cognate		Not cognate		Cognate	
actually	*profit*	*affects*	*impression*	*boundary*	*increases*	*applied*	*observed*
aware	*proper*	*atmosphere*	*informed*	*clumsy*	*lack*	*assembly*	*obtained*
daring	*quality*	*attitude*	*opinions*	*clung*	*motionless*	*calm*	*occurs*
dreadful	*realize*	*common*	*particularly*	*corridor*	*opposite*	*central*	*orderly*
fairly	*remarkable*	*complicated*	*predict*	*display*	*pace*	*delicate*	*presence*
judgment	*scarcely*	*concern*	*preferred*	*dusk*	*rage*	*event*	*previous*
keen	*standard*	*confidence*	*responsibility*	*ease*	*relief*	*evidence*	*producing*
lately	*successful*	*considering*	*satisfaction*	*environment*	*ridge*	*exposed*	*quantity*
mood	*suitable*	*enthusiasm*	*severe*	*fierce*	*sigh*	*graceful*	*reaction*
overcome	*threat*	*especially*	*survived*	*hastily*	*sturdy*	*impatiently*	*reduced*
poisonous	*throughout*	*extremely*	*treatment*	*household*	*weary*	*interior*	*singular*
pride	*wealth*	*ideal*	*tremendous*	*illness*	*wound*	*mysterious*	*various*

APPENDIX D

Connecting Words' Conceptual Relation and Living Word Vocabulary Level

Conceptual relation	Level 4	Level 6
Temporal/aspectual	*finally*[a] *until*	*meanwhile* *usually*[a]
Causal	*because* *if*	*otherwise* *therefore*
Contrastive	*but* *instead*	*although* *however*

[a]The words *usually* and *finally* are adverbs that convey temporal/aspectual meanings. They were used in Unit 1 as a transition to more abstract connecting words.

APPENDIX E

Activities to Develop Content Vocabulary by Day

Weekly Overview

Activity	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8
Picture card review flash		x	x	x	x	x	x	
Picture card introduction	x	x	x	x	x	x		
Interactive reading	x	x	x	x	x	x		
Picture walk review		x	x	x	x	x	x	
Picture card reinforcement	x	x	x	x	x	x		
Glossary	x	x	x	x	x	x	x	
Cognate preview activity	x							
Pocket chart sort		x		x		x		
Songs					x			
Bingo							x	
Crossword								x
Assessment								x

Supporting Information

Additional supporting information may be found in the online version of this article on the publisher's website:

- Appendix F: Methods to Teach Content and Connecting Words
- Appendix G: ELVA Administrator Instructions and Examples of Items
- Appendix H: ELCA Administrator Instructions and Examples of Items
- Appendix I: Acquisition of Vocabulary in English Fidelity and Quality of Intervention Protocol
- Appendix J: Fidelity Rating by Teacher and Component for Content and Connecting Words Instruction
- Appendix K: Average Quality of Implementation Ratings by Component and Time of Observation
- Appendix L: Models Used to Investigate Each Research Question