Published in final edited form as:

Lang Learn. 2018 March ; 68(1): 206–223. doi:10.1111/lang.12274.

Learning Words and Definitions in Two Languages: What Promotes Cross-Language Transfer?

Giang Pham^a, Danaee Donovan^b, Quynh Dam^{a,c}, and Amy Contant^d

^aSan Diego State University

^bFramingham Public School District

^cUniversity of California San Diego

^dHolyoke Medical Centre

Abstract

This study used a brief vocabulary training paradigm to examine two factors for cross-language transfer: how similar the first language (L1) is to the second language (L2) and L1-L2 proficiency levels. Fifty-four sequential bilingual children (aged 6–8) with similar L2 English proficiency levels were assigned to three equal groups: a Vietnamese-English group with low L1 proficiency, a Spanish-English group with low L1 proficiency, and a Spanish-English group with high L1 proficiency. Individual training consisted of two mediated learning experiences conducted in the L1 targeting eight vocabulary items using narrative-based activities. Four of eight target words were cognates between Spanish and English. Pre- and posttesting measured definition quality in the L1 and L2. All groups showed improvement in the L1 following training, but only the Spanish-English group with high L1 proficiency improved in L2 English, revealing a degree of spontaneous cross-language transfer among children with high L1-L2 proficiency.

Keywords

vocabulary; sequential bilingualism; children; dynamic assessment; mediated learning experiences

Introduction

Children learn new words through frequent and meaningful interactions (e.g., Cameron-Faulkner, Lieven, & Tomasello, 2003). For children learning two languages, word learning opportunities are distributed across their first (L1) and their second (L2) languages in terms of the contexts in which words are learned (e.g., home vs. school) and the communicative purposes in which they are used (e.g., conversation vs. academic learning). Identifying which skills transfer between languages as well as anticipating the amount of transfer

This article has been awarded an Open Materials badge. Study materials are publicly accessible in the IRIS digital repository at <http://www.iris-database.org>. Learn more about the Open Practices badges from the Center for Open Science: <https://osf.io/tyyxz/wiki>.

Correspondence concerning this article should be addressed to Giang Pham, School of Speech, Language, and Hearing Sciences, San Diego State University, 5500 Campanile Drive, SLHS 238, San Diego, CA 92182. gpham@mail.sdsu.edu.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

possible have important theoretical and practical implications for facilitating vocabulary development among sequential bilingual children.

Children's L1 and L2 are interdependent (Cummins, 1979), and the development of each language relies on a common underlying proficiency (CUP; Cummins, 1980). Based on these theories, cross-language transfer occurs when skill development in the L1 facilitates the development of the L2. The CUP theory posits that cognitive/academic skills transfer between languages, while surface features in each language are learned separately. Consistent with the CUP theory, studies have found positive associations between L1 and L2 literacy (e.g., Cobo-Lewis, Eilers, Pearson, & Umbel, 2002), underscoring the transferability of academic skills across languages. Similarly, phonological awareness has been linked to literacy across languages, highlighting the connection between L1 metalinguistic skills and L2 reading (Melby-Lervåg & Lervåg, 2011).

Relatively fewer studies have examined cross-language transfer within oral language skills that provide the foundation for literacy. In oral language, the boundary between cognitive/academic skills and surface features is often blurred. To illustrate, vocabulary development at a basic level involves mapping phonological forms to semantic concepts. Establishing form–meaning associations can be considered a surface-level skill and, within the CUP theory, would not be readily transferable between languages. Support for this prediction is found in studies that have found minimal associations between L1 and L2 vocabulary knowledge when measured with standardized vocabulary tests (Branum-Martin et al., 2009). However, a single semantic concept can share phonological forms in two languages (i.e., cognates), such as the Spanish-English word pair *elefante–elephant*. Bilingual individuals have shown higher performance on standardized test items that were cognates versus noncognates (e.g., Kelley & Kohnert, 2012). The presence of a cognate advantage suggests that surface-level transfer occurs between languages that share structural similarities.

Within the CUP theory, Cummins (1979) posited the Dual Threshold Hypothesis that related children's language proficiency levels to cross-language transfer: Children who reached a minimal threshold of proficiency in both languages experienced neutral outcomes, while children who reached an upper threshold showed cognitive benefits from their bilingualism. In a large sample of over 18,000 English language learners in Grades 6, 7, and 8 in the United States, Ardasheva, Tretter, and Kinney (2012) found that bilingual children who had reached a high level in L2 English (and presumably in their L1, although not directly measured) had higher reading and math scores than bilingual children with lower levels of L2 English. Additionally, bilingual children with high English proficiency had higher academic scores than native English-speaking (monolingual) students. Researchers concluded that children who had reached the upper threshold in two languages demonstrated an academic advantage for their bilingualism.

Consistent with Cummins's theories, the present study examined two contributing factors for cross-language transfer in vocabulary knowledge: the degree of similarity between the L1 and L2 and L1-L2 proficiency levels. Participants were early school-age bilingual children who spoke either related languages (Spanish and English) or unrelated languages

(Vietnamese and English) of varying proficiency levels. We used a vocabulary training paradigm to examine which skills could transfer across languages and to what degree.

Background Literature

Cross-Language Associations in Word Definitions

The present study used word definition tasks to measure gains following vocabulary training. Children have been found to improve in their ability to define words over time in response to school instruction and academic demands (Snow, 1990). While definitional skills have been extensively studied in monolingual children (e.g., McGregor, Oleson, Bahnsen, & Duff, 2013), relatively few studies have examined bilingual performance on this task. Ordoñez, Carlo, Snow, and McLaughlin (2002) examined definition quality in a sample of 88 Spanish-English bilinguals in Grades 4 and 5 in the United States. Similar to previous studies on definitional skills (e.g., Snow, 1990), participants were asked to define six concrete nouns (e.g., What is a cat?). Definition quality was scored for three components: paradigmatic knowledge (PK), syntagmatic knowledge (SK), and communicative adequacy (CA). PK reflects how a word fits within a hierarchical taxonomy and has been indexed using the inclusion of a superordinate in the definition (e.g., a dog is an animal); SK has been indexed by the number of key characteristics produced (e.g., size, color, location); and CA, or the ability to convey word meaning, has been scored on how well a listener can identify the target word based on the information provided in the definition. Using correlation and regression analyses, Ordoñez and colleagues found that L1 PK accounted for the most variance in L2 PK, beyond what L1 and L2 vocabulary size could predict. To a lesser extent, L1 CA predicted L2 CA, but only for children with high L1 and L2 vocabulary skills. Consistent with the CUP theory (Cummins, 1980), metalinguistic or academic skills (i.e., PK) were more readily shared across languages than surface aspects (i.e., SK and CA). Consistent with the Dual Threshold Hypothesis (Cummins, 1979), children with high L1 and L2 proficiency showed cross-language transfer in cognitive/academic skills as well as in surface aspects of language (e.g., CA).

To our knowledge, there has been one study on word definition skills that involved an unrelated language pair. Kang (2013) examined definitional skills in Korean children, age 5 to 6 years, who were living in Korea and attending a L2 English immersion program. Children were from Korean-only speaking homes and had received at least 2 years of schooling in English. Definitions were scored in terms of superordinates, relative clauses, sentence structure complexity, and CA. Kang found strong associations among components within each language but minimal associations across languages.

Differences in the degree of cross-language transfer across studies may have been related to different home and school contexts for language use. In Kang (2013), Korean was the majority language and the children's L1; children were learning English as a L2 and were exposed to English solely in school. Using each language in separate contexts (e.g., home vs. school) and for highly distinct purposes may have limited cross-language transfer. Additionally, highly distinctive syntactic features between the two languages may have limited the transfer of definitional skills, which were scored heavily based on syntax. To illustrate, English canonical sentence structure is subject-verb-object, and relative pronouns

typically mark the beginning of relative clauses (e.g., the dog *that* ran down the street is tired). Korean sentence structure is subject-object-verb, and relative pronouns are not required to mark relative clauses. Korean children in the process of learning English may possess limited L2 syntactic knowledge, which may have inhibited the transfer of definitional skills across languages. The present study included a Vietnamese-English bilingual group to contribute to the study of definitional skills in the case of two unrelated languages. Similarly to Korean, Vietnamese does not require a relative pronoun to mark relative clauses (Miller, 1976). Vietnamese-English bilinguals from the present study might thus provide additional empirical evidence on whether children who speak two unrelated languages show cross-language transfer in definition quality.

Cognate Advantage

The present study included a subset of target words that were cognates between Spanish and English. The ability to recognize cognates in text (i.e., orthographic cognates) has been related to better performance in reading comprehension (e.g., Nagy, Garcia, Durguno lu, & Hancin-Bhatt, 1993). However, fewer studies have focused on cognates in oral language (i.e., phonological cognates). Early studies on phonological cognates did not find a cognate advantage. In receptive vocabulary tests, no differences in performance on cognates versus noncognates emerged for a sample of 105 Spanish-English bilingual first graders (Umbel, Pearson, Fernandez, & Oller, 1992) and 102 Spanish-English first, third, and sixth graders (Umbel & Oller, 1994) in the United States. However, recent studies have noted that these analyses did not take into account item difficulty: Items classified as cognates may have been more difficult (e.g., lower in frequency or acquired later) than the items that were classified as noncognates, thus canceling out any potential cognate advantage (Kelley & Kohnert, 2012). Additionally, cognate status (i.e., the degree of phonological overlap) was not systematically examined in those studies.

To build on these earlier studies, Kelley and Kohnert (2012) examined the responses of Spanish-English bilingual children in English expressive and receptive vocabulary tests. Translation equivalents were scored using the Cross-Linguistic Overlap Scale for Phonology (COSP; Kohnert, Windsor, & Miller, 2004) to estimate the degree of phonological overlap in initial sound, number of syllables, vowels, and consonants. Items with scores of 0 to 5 were considered noncognates (e.g., *dog/perro*), and items with scores of 6 to 10 were considered cognates (e.g., *elephant/elefante*). Item difficulty was also separated into easy, medium, and hard based on basal and ceiling calculations for individual participants. With the finer-grained analyses of phonological similarity and item difficulty, Kelley and Kohnert found that children responded correctly to a larger proportion of cognates versus noncognates.

Cognate awareness has been related to language exposure and proficiency. In a study of 89 Spanish-English bilinguals in kindergarten and first grade who completed a receptive vocabulary test in English, Méndez-Pérez, Peña, and Bedore (2013) found that children with high exposure to Spanish in the United States performed better on English cognates of Spanish words than children with balanced exposure or children with high exposure to English. In a sample of 80 Spanish-English preschoolers in the United States, Simpson Baird, Palacios, and Kibler (2016) found a relation between vocabulary size and cognate

awareness. Items from standardized tests of expressive and receptive vocabulary in Spanish and English were classified as cognates using the COSP (Kohnert et al., 2004). Main findings included increased performance on cognates in both languages (Spanish and English) and modalities (expressive and receptive) and a positive association between Spanish cognate recognition and Spanish expressive vocabulary scores.

Finally, the role of cognates in vocabulary development has been examined in children who speak two related versus unrelated languages. Sheng, Lam, Cruz, and Fulton (2016) created picture-naming tasks of cognates and noncognates between Spanish and English that controlled for word frequency, length, age of acquisition, and COSP scores. They found that the two bilingual groups in the United States (Spanish-English and Mandarin-English) named pictures in English at comparable levels. However, Spanish-English bilinguals named a greater proportion of cognates than noncognates. In contrast, the Mandarin-English group showed no difference in performance based on cognate status because these two languages did not share cognates. This study showed that L2 performance may be similar across bilingual groups, while the path to L2 proficiency may differ depending on whether children speak two related or unrelated languages. Overall, findings from the literature indicate that bilingual children are sensitive to phonological cognates, and the presence of a cognate advantage can depend on various factors, including the degree of phonological overlap between the L1 and L2, control for item difficulty, children's overall language proficiency, and exposure to the L1.

Issues of Study Design

There has been a growing body of literature on cross-language transfer at the word level. Recent studies on vocabulary development have employed sophisticated statistical analyses, such as hierarchical linear modeling (e.g., Branum-Martin et al., 2009; Pham & Kohnert, 2014). Similarly, recent studies on cognate facilitation in bilingual children have made substantial strides in controlling for item difficulty and phonological properties (e.g., Sheng et al., 2016). These studies have provided a solid basis for identifying the presence of cross-language associations. However, the vast majority of studies have measured performance at a single time point. The limitation of analyzing data from a single time point is the inability to make causal inferences between variables. One way to extend the study of cross-language transfer is to examine change over time. Longitudinal studies that directly measure the L1 and L2 at multiple time points provide the opportunity to examine the directionality of transfer. A broad finding from longitudinal studies with school-age sequential bilinguals is that transfer occurs from a minority L1 to the majority L2, but not necessarily in reverse (e.g., Pham, 2016; Uccelli & Pérez, 2007).

Another way to extend the study of transfer has been through the use of experimental designs. Experimental designs have the advantage of manipulating variables that contribute to L1 and L2 growth. Consistent with findings from longitudinal studies (e.g., Pham, 2016), studies comparing bilingual and L2-only conditions have suggested unidirectional transfer from the L1 to the L2 in cases where the L1 is a minority language. In a meta-analysis of 17 studies on program effectiveness, Rolstad, Mahoney, and Glass (2005) found that bilingual instruction was just as effective as L2-only instruction in attaining L2 academic outcomes

and had the additional benefit of promoting L1 maintenance. Similarly, Ebert and Kohnert (2016) conducted a systematic review of treatment studies for bilingual children with language impairment and found that children in bilingual conditions made gains in the L1 and L2, while children in L2-only conditions made gains solely in the L2.

The bilingual conditions reported in previous studies, by definition, included instruction in two languages. That is, bilingual gains from bilingual instruction could have been attributed to training in the L1, the L2, or a combination of both languages. Therefore, it would be difficult to identify direct connections between previous L1 skills and gains in the L2. In order to extend the existing literature, the present study employed an experimental design in which vocabulary items were trained solely in the L1, and outcomes were measured bilingually. The advantage to this design was the ability to isolate skills in the L1 and test for spontaneous (i.e., untrained) transfer to the L2.

The Present Study

For bilingual children, the transference of skills from one language to the other can facilitate overall vocabulary development. The present study examined the influence of L1-L2 similarity and proficiency level on cross-language transfer using a vocabulary training paradigm. Participants spoke two languages that were either related (Spanish and English) or unrelated (Vietnamese and English). They were matched in L2 English proficiency and differed in L1 proficiency level (high vs. low) for a total of three groups: Spanish-English bilinguals with high L1 proficiency, Spanish-English bilinguals with low L1 proficiency, and Vietnamese-English bilinguals with low L1 proficiency. We conducted training solely in children's L1 and measured outcomes bilingually in order to address three sets of research questions related to L1 learning, transfer to the L2, and cognates:

1. Do bilingual children show improvements in L1 vocabulary knowledge after completing two brief training sessions? How do the three groups in this study compare in their L1 learning?
2. Do bilingual children transfer newly learned skills from their L1 to the (untrained) L2? How do the three groups compare in the transfer of skills to the L2?
3. Do children who speak two related languages (Spanish-English) show better performance on cognates than on noncognates? How do Spanish-English bilingual children with high versus low L1 proficiency compare?

We anticipated that all children would show improvement on the target words and definitions in the L1 following training. Typically developing bilingual children have been shown to make language gains even with minimal instruction (Peña, glesias, & Lidz, 2001). We included a set of control words in the L1 and L2 that were measured before and after training but were not targeted during training. Pre-to posttraining improvement on the target words and no change on the control words would verify the training effect. Consistent with the Dual Threshold Hypothesis (Cummins, 1979), we predicted that the Spanish-English group with high L1 proficiency (i.e., high proficiency in two languages) would show the greatest amount of L1 learning. We predicted cross-language transfer to occur among

children who spoke two related languages (Spanish and English) as measured by gains on the (untrained) English translations of the target words. In contrast, children who spoke two unrelated languages (Vietnamese and English) would show minimal to no transfer to English. If Vietnamese-English bilinguals did show a degree of cross-language transfer, it would occur on the components of definition quality related to metalinguistic ability (Cummins, 1980). We anticipated that all Spanish-English bilingual children would show better performance on cognates versus noncognates, demonstrating a sensitivity to cognate status (e.g., Sheng et al., 2016).

Method

Participants

Participants were 54 typically developing bilingual children who spoke either Spanish or Vietnamese as the primary home language (L1). Participants were selected to be in three groups: children with low L1 proficiency in Spanish (LSE, $n = 18$) or in Vietnamese (LVE, $n = 18$) and children with high proficiency in Spanish (HSE, $n = 18$). Participants attended first or second grade and were recruited from schools or community centers from the northeastern region of the United States with prior human subjects' approval from our university and community organizations. The two Spanish-English groups (LSE and HSE) attended schools within the same school district. Participants scored within the normal range on the Test of Nonverbal Intelligence, 4th edition (TONI; Brown, Sherbenou, & Johnsen, 2010). There were no parent concerns for language or learning difficulties and no history of special education services.

LSE and LVE participants received school instruction solely in L2 English. In contrast, HSE participants attended a Spanish language immersion program in which children received instruction in both Spanish and English. The Spanish language immersion program comprised 50% English speakers and 50% Spanish speakers at home. All students were taught initial literacy skills in Spanish. The proportion of instruction in Spanish versus English started as a 90–10% split in kindergarten and decreased by 10% of Spanish instruction per grade level. First-grade students received 80% of instruction in Spanish, and second-grade students received 70% of instruction in Spanish. As shown in Table 1, independent variables were compared across groups using one-way analyses of variance. The three groups did not differ in age, level of maternal education, and home language use. Mean participant age was 7 years, with a similar proportion of boys and girls in each group. Based on written responses from a brief parent questionnaire, approximately 80% of each group reported the maternal education level to be high school or less. All parents reported speaking the L1 “most of the time” or “all of the time” at home.

Independent variables that differed across groups were birthplace, $F(2, 53) = 5.16, p = .01, \eta_p^2 = .17$, and TONI scores, $F(2, 53) = 8.42, p = .001, \eta_p^2 = .25$. All participants in the LSE group were born in the United States, compared to 61% of LVE participants and 83% of HSE participants. Post hoc comparisons revealed that the two Spanish-English groups were comparable in birthplace (HSE = LSE, $p = .18$); LVE participants were similar to HSE participants ($p = .07$); and more participants in the LSE group were born in the United States, compared to those in the LVE group ($p = .002, d = 1.10$). Although all groups

performed within normal limits on the nonverbal intelligence test, LVE participants had the highest TONI scores on average (LVE > HSE, $p = .02$, $d = 0.98$; LVE > LSE, $p < .001$, $d = 1.25$); and the two Spanish-English groups had comparable scores (HSE = LSE, $p = .11$). Differences in age of English onset approached statistical significance ($p = .06$, $\eta_p^2 = .11$): LSE participants began speaking English at a slightly younger age than participants in LVE and HSE (2 vs. 3.4 years).

In order to estimate L1 and L2 proficiency, participants completed measures of receptive and expressive vocabulary in each language using select items from the Spanish-English bilingual version of the Receptive and Expressive One-Word Picture Vocabulary Tests (ROW and EOW; Brownell, 2000a, 2000b). Items from each test were selected from the 5–11 age range (ROW = 48 items; EOW = 43 items) in order to include advanced vocabulary, such as *microscope* and *competitive*. Vietnamese items had been previously translated and validated for use with a separate sample of Vietnamese-English bilingual children (Pham & Kohnert, 2014). As shown in Table 1, HSE participants had higher proficiency in the L1 than the remaining two groups, as shown on measures of receptive vocabulary (HSE > LSE, $p < .001$, $d = 1.47$; HSE > LVE, $p = .01$, $d = 0.84$; LSE = LVE, $p = .18$) and expressive vocabulary (HSE > LSE, $p < .001$, $d = 1.52$; HSE > LVE, $p < .001$, $d = 1.12$; LSE = LVE, $p = .38$). HSE participants had similar levels of expressive vocabulary in the L1 and L2, indicating balanced performance across languages. In contrast, the LSE and LVE groups had stronger expressive vocabulary skills in English than in their L1 (see Table 1), with LVE outperforming LSE on L2 English expressive vocabulary (42% vs. 28%, $p = .002$, $d = 1.07$). All three groups had comparable levels of English receptive vocabulary, $F(2, 53) = 0.44$, $p = .65$.

Training Paradigm

Training comprised Mediated Learning Experiences (MLEs; Lidz, 1991). MLEs have been used within dynamic assessment procedures to identify children at risk for language impairment and to measure children's capacity for vocabulary learning (e.g., Peña et al., 2001). MLEs were used in the present study to promote active learning and skill generalization. Each training session included the four key features of MLE proposed by Lidz: intentionality, meaning, transcendence, and competence (see training scripts in Appendix S1 in the Supporting Information online). Intentionality consisted of an explicit statement of the goal (e.g., Today we are going to learn new words and how to define them). Meaning consisted of an explanation of the importance of the goal (e.g., Learning new words helps us understand what we hear and read). Transcendence was the explanation of how the new skill could be applied to other contexts (e.g., It is important to learn new words in Spanish to talk to your family members). Competence consisted of the reinforcement of one's learning (e.g., At first you did not know all these words. Now you are able to define...).

Training Procedure

Each participant group completed the study within a 3-week time period. Participants completed pretraining assessment measures during Week 1, training during Week 2, and posttraining measures during Week 3. Training was conducted solely in the participants' L1

(Spanish or Vietnamese). Training consisted of two sessions of 15–20 minutes each, and sessions were separated by 1 to 3 days. Participants worked individually with a trained research assistant in a quiet area of the school or community center. Research assistants were graduate students in speech-language pathology, fluent in the target language of Spanish or Vietnamese.

Training materials included a storybook, flash cards of the target words and definitions, and an elephant board game. We selected *Elmer* (McKee & Puncel, 1989) for the storybook because of its availability in Spanish and in Vietnamese. Based on the original story, we designed a story script that included five exemplars of each of the eight target words for a total of 40 exemplars. Within each training session, research assistants used a scripted protocol to highlight and define eight target words within the context of the storybook and interactive activities (cf. Hickman, Pollard-Durodola, & Vaughn, 2004). Training scripts for the two sessions and storybook scripts in Spanish and Vietnamese as well as the corresponding English translation are available in Appendixes S1 and S2 in the Supporting Information online.

Pre- and Posttraining Measures

Assessments were conducted in both languages before and after training in order to measure vocabulary learning in the L1 and the potential transfer of skills to English. The order of language administration was counterbalanced, and language tests were separated by at least 1 day. To prevent tester bias, research assistants did not conduct posttesting with the same participants for whom they provided training. Vocabulary knowledge was assessed using word definition tasks in the L1 and L2.

Target and Control Words—A total of eight words were targeted during training. As shown in Appendix S3 in the Supporting Information online, target words consisted of four nouns and four nonnouns (e.g., verbs and adjectives) selected from the storybook. Target words were Tier 2 vocabulary (Beck, McKeown, & Kucan, 2013) in order to avoid words typically found in everyday conversation (Tier 1) and in technical subject matter (Tier 3). Of the eight target words, four were cognates in Spanish and English, namely, words that shared meaning and form. Vietnamese and English do not share cognates. Target words were scored using the COSP (Kohnert et al., 2004) to estimate the degree of overlap between the target word and its English translation in initial sound, number of syllables, vowels, and consonants. Based on the COSP, words with a score of 6 or greater were considered cognates, and words with scores of 5 or less were considered noncognates (Kelley & Kohnert, 2012). As shown in Appendix S3, COSP scores for cognate items ranged from 6 to 9, and scores for noncognates ranged from 0 to 2. Definitions of the target words were adapted from definitions found in online dictionaries for each language. They were tailored by the research team to sound natural in the children's L1 and be comparable in length across languages (see Appendix S3). For experimental control, we measured performance on a set of six words that were measured at pre- and posttesting but not targeted during training. Control words consisted of concrete nouns, similar to the word definition tasks that have been previously used with school-age bilingual children (e.g., Kang, 2013; Ordoñez et al.,

2002). Control words were highly familiar objects across all three languages (i.e., dog, chair, house, clock, bed, and door).

Picture Identification Tasks—Comprehension of the eight target words was measured using picture identification tasks (PicID) in the L1 and L2. Children were presented with a 2×2 array of photos and asked to point to the picture that corresponded to the target word. Picture stimuli consisted of color photographs gathered from freely available online resources. Items were embedded into the larger receptive vocabulary tasks that were used to measure general language proficiency, and the order of items differed across languages. Pictures were used only during pre-and posttesting; they were not shown during the training sessions, nor did they depict images related to the storybook used in training. Dependent measures were the number of words correctly identified (out of eight) in each language.

Word Definition Tasks—In the word definition tasks, the examiner said a word and asked the child to define it. The examiner provided up to three prompts: “What is (a), ----” “Tell me more about ----,” and “Anything else?” Participants completed two practice items, six control words, and then the eight target words in the same order for all languages. Definitions were audio recorded and transcribed using SALT software (Miller & Iglesias 2012). Consistent with how the tasks were administered, each of the 54 participants had two transcripts, one for each language, for a total of 108 transcripts. Each transcript consisted of the six control words and eight target words in one language. Consistent with the listening procedures outlined in Heilmann et al. (2008), a second trained research assistant fluent in the target language relistened to all audio recordings to verify transcription accuracy.

Definitions were scored for quality using an adapted version of Ordoñez et al. (2002), originally created for Spanish-English bilingual children, which consisted of three components: PK, SK, and CA. PK was initially scored on a 6-point scale consistent with Ordoñez et al.: 0 = no use of a superordinate; 1 = translation equivalent; 2 = synonym; 3 = empty superordinate (e.g., a boat is a *something* that...); 4 = not best superordinate (e.g., an envelope is a *paper*); 5 = adequate superordinate (e.g., a boat is a *vehicle*). However, unlike Spanish and English, canonical sentence structure in Vietnamese does not include relative clauses that are preceded by relative pronouns (Miller, 1976). While it is common for children to produce an empty superordinate in a relative clause in English (e.g., “a bird is something *that* flies”) and in Spanish (e.g., “un pájaro es algo *que* vuela”), this sentence structure is highly unusual and unnatural in Vietnamese (e.g., “con chim là cái gì *mà* bay”). Therefore, to equate the scoring system across languages, empty superordinates were excluded from the analysis, and PK scoring focused on the use of real superordinates (PKReal): Scores between 0 and 3 were collapsed to 0, and real superordinate scores of 4 or 5 were maintained. Dependent measures were total raw scores for PKReal in each language (out of 40).

For SK, definitions were given one point for each type of feature: size, quantity, shape, color, composition, parts, location, function/use, activities done by/with, comparisons, multiple meanings, metaphoric/simile, relating to personal experience, and belonging to (e.g., “a trunk is an *elephant’s* nose”). Consistent with the scoring system outlined in Ordoñez et al. (2002), each type of feature could have a maximum of 3 points each.

Dependent measures were total raw scores for SK in each language (unlimited maximum). CA was scored on a 4-point scale: 0 = no chance of identifying the intended item; 1 = item is identifiable, but definition includes misleading, general, or untrue information; 2 = item is identifiable, definition includes only true information, but is ambiguous, (i.e., could be 2–3 different items); 3 = adequate definition, that is, it is possible to identify the item solely on the basis of the provided information. Dependent measures were total raw scores for CA in each language (out of 24).

Reliability of Word Definition Scoring—Interrater reliability of the definition scoring system was conducted in a two-step process. First, we randomly selected 20% of the transcripts in each language to be scored independently by two trained research assistants. Using Cohen's kappa to estimate interrater agreement after correcting for chance, initial values for English were .82 for SK, .90 for PKReal, and .99 for CA. Initial values for Spanish were .58 for SK, .66 for PKReal, and .97 for CA. Initial values for Vietnamese were .42 for SK, .99 for PKReal, and .78 for CA. Discrepancies were verbally discussed to ensure scoring equivalence across languages, and then the full set of transcripts was reviewed a second time. In a second random sample of 20% of the transcripts, kappa values were over .90 for each language.

Training Fidelity

Half of Session 1 and half of Session 2 from each group were video recorded. Video recordings were scored by independent trained raters fluent in the target language for three measures of training fidelity: (a) adherence to the training script, (b) consistent use of the L1, and (c) training session length. Adherence to the training script was scored on a binary basis to rate whether each trainer had provided each specific instruction in the training script (1 = included, 0 = omitted or not administered correctly). Average point-by-point training fidelity across groups was 99% for Session 1 and 98% for Session 2. Second, consistent L1 use by the research assistant was scored using a time interval observation method (e.g., Olswang, Svensson, Coggins, Beilinson, & Donaldson, 2006), in which the first 15 minutes of a session were divided into 30-second intervals, and each interval was scored with the following rubric: 1 = used solely the L1; 0 = used both languages; NA = not applicable (e.g., silent interval). The average percentage of intervals in which the research assistant spoke solely in the L1 was 98% for Session 1 and 99% for Session 2. Finally, the start and end times for each section of the training script (e.g., introduction, book activity, matching game, ending) were recorded to calculate the length of each section and the total session length. The average session length across groups was 17 minutes per session.

Data Analysis

Regarding L1 learning (Research Question 1) and transfer to the L2 (Research Question 2), we conducted within-group analyses to measure absolute change from pre-to posttraining and between-group analyses to measure relative change across groups. For within-group analyses, we conducted repeated-measures multivariate analyses of variance (MANOVAs) for each group with three dependent measures (SK, PKReal, CA) and time as a within-subjects factor (pre-vs. posttraining). We conducted separate repeated-measures MANOVAs

for target and control words, in the L1 and in the L2. If the omnibus multivariate test was significant ($p < .05$), we then examined results from univariate repeated-measures ANOVAs for each dependent measure. For between-group analyses, we conducted separate univariate ANOVAs for each dependent measure at posttraining with group as the between-subjects factor (HSE, LSE, LVE) and the corresponding pretraining score as a covariate. Because the LVE group showed a higher average TONI score than the HSE and LSE groups (see Table 1), TONI was also included as a covariate in between-group analyses that included LVE. To illustrate, the ANOVA for SK had the posttraining score of SK as the dependent variable (SKPost), group as the between-subjects factor, and the pretraining score (SKPre) and TONI as covariates. Significant F ratios were followed by pairwise comparisons using LSD tests (equivalent to no adjustment for multiple comparisons).

To examine the role of cognates between Spanish and English (Research Question 3), we conducted separate repeated-measures ANOVAs for each group with two within-subjects factors of time (pre-vs. posttesting) and cognate status (cognate vs. noncognate). Significant models that included cognate status were then followed by between-group models to compare the HSE and LSE groups using univariate ANOVAs for the specific dependent measure in question and group as the between-subjects factor. Throughout the present study, we selected partial eta-squared coefficients (η_p^2) as the measure of effect size for its generalizability across within- and between-subjects analyses (Richardson, 2011). Consistent with Cohen (1988), η_p^2 values of .01, .06, and .14 were considered small, medium, and large effects, respectively.

Results

Tables 2 and 3 display participants' performance on the target and control words in their L1 and in L2 English, respectively. Consistent with reporting recommendations for educational research by Zientek and Thompson (2009), correlation matrixes have been provided for pre- and posttraining performance in the L1 and in English for each group in Appendix S4 in the Supporting Information online.

L1 Learning

Absolute change in the L1 from pre- to posttraining was first measured within each group. For the target words, all omnibus multivariate tests were significant, indicating change in at least one dependent measure for each group: LVE, $F(4, 14) = 5.31$, $p = .008$, $\eta_p^2 = .60$; LSE, $F(4, 14) = 13.07$, $p < .001$, $\eta_p^2 = .79$; and HSE, $F(4, 14) = 42.03$, $p < .001$, $\eta_p^2 = .92$. Follow-up ANOVAs for each group confirmed changes from pre- to posttraining that corresponded to large effect sizes for all three measures (see Table 2). For LVE, participants were able to identify the words in novel pictures (PicID): $F(1, 17) = 8.77$, $p = .01$, $\eta_p^2 = .34$; they produced more key features (SK): $F(1, 17) = 6.42$, $p = .02$, $\eta_p^2 = .27$, and superordinate labels (PKReal): $F(1, 17) = 7.11$, $p = .02$, $\eta_p^2 = .30$, in their definitions; and communicated word meaning more effectively (CA): $F(1, 17) = 8.26$, $p = .01$, $\eta_p^2 = .33$. Similarly, LSE participants improved in PicID, $F(1, 17) = 21.60$, $p < .001$, $\eta_p^2 = .56$; SK, $F(1, 17) = 11.02$, $p = .004$, $\eta_p^2 = .39$; PKReal, $F(1, 17) = 18.26$, $p = .001$, $\eta_p^2 = .52$; and CA, $F(1, 17) = 4.89$, $p = .04$, $\eta_p^2 = .22$. Finally, HSE participants improved in PicID, $F(1, 17) = 22.67$, $p < .001$, =

$\eta_p^2 = .57$; SK, $F(1, 17) = 129.57, p < .001, \eta_p^2 = .88$; PKReal, $F(1, 17) = 34.86, p < .001, \eta_p^2 = .67$; and CA, $F(1, 17) = 41.88, p < .001, \eta_p^2 = .71$.

Between-group comparisons confirmed that HSE showed the largest amount of L1 learning for the three scores of definition quality. For PicID, there was no main effect for group, $F(2, 53) = 2.48, p = .09$. However, for SK, there was a main effect for group, $F(2, 53) = 4.80, p = .01, \eta_p^2 = .16$. Pairwise comparisons revealed that HSE showed the greatest gains for SK (HSE > LSE, $p = .01, d = 1.43$; HSE > LVE, $p = .01, d = 1.03$) and that the two low L1 proficiency groups showed equal gains on this measure (LSE = LVE, $p = .99$). For PKReal, the main effect for group was significant, $F(2, 53) = 3.61, p = .03, \eta_p^2 = .13$. Pairwise comparisons revealed equal performance between the two Spanish-English groups (HSE = LSE, $p = .46$), equal performance between the two low L1 proficiency groups (LSE = LVE, $p = .09$), and greater performance for HSE than LVE ($p = .01, d = 0.75$). For CA, the main effect for group was significant, $F(2, 53) = 6.78, p = .003, \eta_p^2 = .22$. Pairwise comparisons revealed that HSE showed the greatest gains (HSE > LSE, $p = .001, d = 1.77$; HSE > LVE, $p = .047, d = 0.95$) and that the two low L1 proficiency groups showed equal gains on this measure (LSE = LVE, $p = .19$).

Omnibus multivariate tests for the control words were not significant, indicating no change from pre-to posttesting for any group (see Table 2): LVE, $F(3, 15) = 1.12, p = .37$; LSE, $F(3, 15) = 2.30, p = .12$; and HSE, $F(3, 15) = 0.16, p = .93$. Improvement solely on the target words confirmed the direct relationship between training and L1 gains.

Transfer to the L2

Absolute change in the (untrained) L2 from pre-to posttraining was first measured within each group (see Table 3). Omnibus multivariate tests were not significant for LVE, $F(4, 14) = 0.81, p = .54$, and LSE, $F(4, 14) = 0.71, p = .60$. For these two groups, there were no changes from pre-to posttesting on the English translations of the target words. In contrast, the omnibus multivariate test for HSE was significant, $F(4, 14) = 5.51, p = .01, \eta_p^2 = .61$. Follow-up ANOVAs for HSE showed significant differences between pre- and posttesting for two of three measures, with large effect sizes: PKReal, $F(1, 17) = 6.40, p = .02, \eta_p^2 = .27$, and CA, $F(1, 17) = 15.70, p = .001, \eta_p^2 = .48$. PicID and SK did not show change in English from pre-to posttraining: PicID $F(1, 17) = 3.24, p = .10$, and SK, $F(1, 17) = 2.79, p = .11$.

There were no changes in the English translations of the control words for LSE, $F(3, 15) = 2.15, p = .14$, nor HSE, $F(3, 15) = 0.43, p = .74$. For LVE, changes in the English translations of the control words approached significance, $F(3, 15) = 3.26, p = .051$; however, the effect was in the opposite direction, with poorer performance at posttesting. No positive change for the English translations of the control words provided additional evidence that the transfer of skills to the L2 was directly related to training in the L1.

Between-group comparisons showed that HSE showed greater L2 transfer than the other groups for at least one measure, namely, CA: $F(2, 53) = 4.16, p = .02, \eta_p^2 = .15$. As shown in Figure 1, pairwise comparisons revealed HSE showed the greatest gains (HSE > LSE, $p = .01, d = 0.93$; HSE > LVE, $p = .05, d = 0.51$) and that the two low L1 proficiency groups showed equal gains on this measure (LSE = LVE, $p = .59$). There was no main effect for

group in analyses of PicID, $F(2, 53) = 0.50, p = .61$, of SK, $F(2, 53) = 0.04, p = .96$, and of PKReal, $F(2, 53) = 1.28, p = .29$.

Cognate Status

Analyses of cognates focused on CA, the measure of definition quality for which there were between-group differences in English performance following training. We first divided target words into cognates and noncognates and conducted repeated-measures ANOVAs for each group. For LVE, there was no main effect for time, $F(1, 17) = 1.95, p = .18$, indicating no change from pre-to posttesting, and no time $\times$ cognate interaction, $F(1, 17) = 3.03, p = .10$. There was a main effect for cognate, $F(1, 17) = 13.28, p = .002, \eta_p^2 = .44$; however, the effect was in the opposite direction in that performance on noncognates was better than performance on cognates. No cognate effects were expected for LVE as there were no cognate pairs between Vietnamese and English. For LSE, there was no main effect for time, $F(1, 17) = 0.35, p = .56$, indicating no change from pre-to posttesting. Similarly, there was no main effect for cognate, $F(1, 17) = 1.60, p = .22$, and no time $\times$ cognate interaction, $F(1, 17) = 0.09, p = .77$. For HSE, there was a main effect for time, $F(1, 17) = 15.70, p = .001, \eta_p^2 = .48$; however, there was no main effect for cognate, $F(1, 17) = 4.17, p = .06$, and no time $\times$ cognate interaction, $F(1, 17) = 0.35, p = .56$.

One potential reason for the lack of a cognate effect in Spanish-English groups (HSE or LSE) may have been related to differences in item difficulty. As a gross measure of difficulty level, we referenced an English lexical database for age of acquisition information (Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012). As shown in Appendix S3, the age of acquisition was comparable for the cognate and noncognate target words that were nouns. However, the age of acquisition of the nonnouns was greater for the cognates than the noncognates, suggesting a difference in difficulty level among the nonnoun target words. To address this concern, we reran the same analyses with scores from the nouns only (two cognates vs. two noncognates). For LVE, there continued to be no main effect for time, $F(1, 17) = 0.11, p = .74$, no main effect for cognate, $F(1, 17) = 2.14, p = .16$, and no two-way interaction, $F(1, 17) = 0.86, p = .37$. For LSE, there continued to be no main effect for time, $F(1, 17) = 2.04, p = .17$, and no two-way interaction, $F(1, 17) = 3.44, p = .08$. However, the main for cognate was significant with a large effect size, $F(1, 17) = 17.10, p < .001, \eta_p^2 = .50$. For HSE, there were main effects with large effect sizes for time, $F(1, 17) = 12.41, p = .003, \eta_p^2 = .42$, and cognate, $F(1, 17) = 28.00, p < .001, \eta_p^2 = .62$; and a significant two-way interaction, $F(1, 17) = 26.50, p < .001, \eta_p^2 = .61$. We then compared the HSE and LSE groups on cognate nouns and found a main effect for group with a medium effect size, $F(1, 35) = 4.89, p = .03, \eta_p^2 = .13$, indicating that HSE participants showed a larger cognate effect than LSE participants. As shown in Figure 2, these results indicated that both LSE and HSE groups showed higher performance for English cognate nouns versus noncognate nouns. However, HSE participants demonstrated improvements from pre-to posttesting for English cognate nouns and a larger cognate effect, compared to LSE participants.

Discussion

This study focused on performance by three bilingual groups following brief vocabulary training in children's L1. All three groups had similar levels of L2 English proficiency, while two groups had low L1 proficiency (LVE and LSE), and one group had high L1 proficiency (HSE). We discuss the main findings in terms of the three sets of research questions regarding L1 learning, transfer to the L2, and cognate status.

L1 Learning

All three groups showed gains in their responses to the target words and definitions following training in the L1, with large effect sizes ranging from .22 to .88. It was notable that children in the LVE and LSE groups made gains in the L1, even though these groups had low L1 proficiency. This finding emphasizes children's capacity for language learning: Children can and do learn new skills, even in their relatively weaker language. Gains in the L1 following a minimal amount of instruction support the feasibility for longer-term educational interventions targeting L1 vocabulary development. For sequential bilingual children, the minority L1 plays an important role in maintaining family ties and developing a strong sense of identity (e.g., Oh & Fuligni, 2010). Programs that provide opportunities to learn vocabulary within meaningful social and academic contexts have been shown to be effective for L1 and L2 development (e.g., Cummins, 2005a; Hickman et al., 2004).

Comparing across groups, the HSE group showed the greatest amount of L1 learning, with medium effect sizes, ranging from .16 for SK to .22 for CA. This finding is consistent with the Matthew effect previously found among monolingual children (e.g., Penno, Wilkinson, & Moore, 2002): Children with richer vocabularies show greater gains following instruction than children with poorer vocabularies. While all groups were similar in L2 vocabulary, HSE participants had strong L1 vocabulary and presumably a larger overall vocabulary size, which contributed to larger gains following training.

Transfer to the L2

Consistent with our predictions, the LVE group did not show a transfer of skills from the L1 to the L2. Cross-language transfer may have been constrained by the minimal structural overlap between Vietnamese and English. This finding replicated the single previous study on word definitions in bilingual children who spoke two unrelated languages (Kang, 2013). It is also possible that proficiency level may have influenced these results. In addition to linguistic differences, children's overall low L1 proficiency may have inhibited their ability to transfer skills between languages.

The lack of cross-language transfer in the LSE group was unanticipated. We had predicted that LSE would show cross-language transfer given the structural similarities between Spanish and English. This finding differs from previous studies that found cross-language associations in Spanish-English bilinguals in word definition tasks (e.g., Ordoñez et al., 2002). Differences in findings may be related to the methodologies employed. The bulk of studies have examined transfer based on data from a single time point. The present study experimentally tested for transfer by training solely in the L1 and measured outcomes

bilingually. This paradigm is arguably a more stringent test of a causal connection between the L1 and L2. The lack of transfer to English in LSE children may have been due to their low L1 proficiency. LSE participants may not have had adequate skills in the L1 to capitalize on the L1-L2 structural similarities.

Only the HSE group showed improved skills in L2 English following training. These children showed gains in English, with large effect sizes ranging from .27 (PKReal) to .48 (CA). Across groups, HSE children showed higher performance than LSE and LVE participants on CA in the English definitions, with a medium effect size of .15. Consistent with Ordoñez et al. (2002), increased use of superordinates (PKReal) may reflect cognitive/academic language proficiency (Cummins, 1980), a skill that is transferable between languages. Improvements in CA complement the results reported by Ordoñez et al., who found minimal relations between languages for CA in the entire sample but relatively stronger cross-language relations among children with higher language proficiency. As we predicted, the HSE group showed the greatest amount of learning in the L1 as well as transfer to the L2. Within the Dual Threshold Hypothesis (Cummins, 1979), LVE and LSE participants may have met the lower proficiency threshold to continue to learn in each language. However, only HSE children may have reached the upper threshold of proficiency in the L1 and L2 to reap cognitive and academic benefits related to their advanced bilingualism (Ardasheva et al., 2012).

Cognate Status

The presence of cross-language transfer for HSE participants may have been driven by L1-L2 similarity and/or high proficiency in each language. Although this study cannot completely separate these factors, we can evaluate one aspect of L1-L2 similarity, namely, cognates. As anticipated, the LVE group did not show a cognate effect as there were no cognates between Vietnamese and English. Our initial analyses within the Spanish-English groups (LSE and HSE) did not find differences in performance between cognates and noncognates when the total set of target items was analyzed. However, further analysis on a reduced set of target items (e.g., nouns only) revealed a cognate effect in both the LSE and HSE groups. These results underscore the importance of controlling for difficulty level between cognate and noncognate stimuli in order to prevent the masking of a cognate advantage (Kelley & Kohnert, 2012; Sheng et al., 2016).

Although both Spanish-English groups showed higher performance for cognates, only HSE participants showed overall improvement from pre- to posttesting. For LSE participants, cross-language similarity alone was not enough to promote transfer between languages. While Spanish-English bilinguals may be sensitive to cognate status, only children with high L1 proficiency may be able to spontaneously apply this knowledge to the task of vocabulary learning. These findings are consistent with previous correlational studies that examined cognate awareness using standardized vocabulary tests (e.g., Méndez-Pérez et al., 2013). The present study extended this research to the use of a training paradigm administered in the L1 with bilingual pre- and posttraining measures.

It should be noted that the training paradigm did not explicitly focus on cognates. Although half of the target words were cognates, the similarities between the L1 and L2 were not

highlighted. Rather, training was conducted solely in the L1 to experimentally test whether skills would spontaneously transfer to English. Findings suggest that children, particularly with low L1 proficiency, may need explicit instruction on cognates to tap into their cognate knowledge. In the literacy domain, instruction on orthographic cognates between Spanish and English has been found to boost English reading comprehension (e.g., Nagy et al., 1993). Interventions that promote the development of proficiency in both languages and include a component of explicit instruction targeting cognates may be more effective in promoting cross-language transfer than programs that solely focus on increasing proficiency (Cummins, 2005b).

Explicit cross-language instruction could also be expanded to syntactic similarities. Spanish and English share syntactic structures, such as relative clauses, that facilitate the production of definitions and the expansion of sentences in spoken and written language (e.g., Ordoñez et al., 2002). For two unrelated languages, explicit instruction on how syntax differs across languages can promote metalinguistic awareness skills. For example, the lack of transfer in the LVE group may have been due to low L1 proficiency as well as the differences between syntactic structures used in definitions (Kang, 2013). Explicit instruction on cross-language similarities and differences can be tailored to specific language pairs.

Limitations

There were at least three limitations to this study. One limitation was the relatively small sample size in each group. The effects found in this study ranged in size from medium to large. However, small sample size may have limited our ability to detect more subtle changes from pre-to posttraining within each group and differences between groups. With a fuller sample, L1 proficiency could be also examined as a continuous variable rather than binary (high vs. low). A wider range of proficiency levels could allow for a more nuanced evaluation of a threshold effect. A second limitation was the lack of a Vietnamese-English group with high L1 proficiency. Future studies with bilingual children who speak two unrelated languages with high proficiency are needed to further distinguish between the roles of L1-L2 similarity and varying proficiency levels. In the present study, the lack of cross-language transfer in LVE participants could have been due to speaking two unrelated languages, low L1 proficiency, or a combination of both factors. Similarly, L2 transfer found in HSE participants could have been due to high L1 proficiency, speaking two related languages, or both factors. Comparisons between high L1 proficiency groups that speak two related versus unrelated languages would help to distinguish between the relative importance of L1-L2 similarity for cross-language transfer.

Third, for the two Spanish-English groups (HSE and LSE), L1 proficiency level could not be separated from school programming: The HSE group received dual-language instruction in school, while the LSE group received instruction in English only. Future studies could examine children with high L1 proficiency who do not receive dual-language instruction to further examine specific factors that lead to cross-language transfer in highly proficient bilingual children. It is possible that HSE children held other advantages over LSE children in addition to their high proficiency level. For example, children who attend a Spanish immersion program may have come from families who place a relatively higher value on L1

maintenance. Additionally, the HSE group was accustomed to academic instruction in the L1 (unlike LSE or LVE participants) and may have responded more readily to vocabulary training in the L1, compared to the other two groups. However, LSE and HSE participants were from the same school district and were comparable in ethnic background. Both groups had a similar level of maternal education, a gross measure of socioeconomic status, and had similar overall L2 English proficiency, as measured by expressive and receptive vocabulary tasks, and similar performance on the English control words. Nonetheless, individual differences and external factors beyond what were measured in this study may have contributed to the transfer effects found in HSE (e.g., de Jong, 2002).

Conclusion

The present study underscores bilingual children's capacity to learn new words and definitions even in their weaker language. However, low proficiency in the L1 may prevent children from transferring skills to their L2. Children in this study who had strong L1 and L2 skills showed the greatest amount of L1 learning and cross-language transfer to the L2. Educational support that develops L1 proficiency and makes explicit connections between the L1 and L2 can promote advanced bilingualism and its related cognitive and academic benefits.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported in part by the University of Massachusetts Amherst Faculty Research Grant and Public Service Endowment Grant awarded to G. Pham. Portions of this research were presented at the 2014 and 2015 annual meetings of the American Speech-Language-Hearing Association. We acknowledge community partners at Barbieri Elementary School, Brophy Elementary School, Framingham Public School District, and the Vietnamese American Initiative for Development; research assistants Mircea Cimisliu, Quynh Do, Jennifer Vo, Amanda Dembowski, Jihyun Jung, Anastasia Chobany, Margaux Frank, Sarah Brugman, Abigail Asselin, Ann Nguyen, Tina Nguyen, and Diane Guerrero; and the participating children and their families. We thank Catherine Snow and Kerry Ebert for helpful comments on earlier versions of this manuscript.

References

- Ardasheva Y, Tretter TR, Kinny M. English language learners and academic achievement: Revisiting the threshold hypothesis. *Language Learning*. 2012; 62:769–812. DOI: 10.1111/j.1467-9922.2011.00652.x
- Beck I, McKeown M, Kucan L. *Bringing words to life: Robust vocabulary instruction*. New York, NY: Guilford Press; 2013.
- Branum-Martin L, Mehta PD, Francis DJ, Foorman BR, Cirino PT, Miller JF. Pictures and words: Spanish and English vocabulary in classrooms. *Journal of Educational Psychology*. 2009; 101:897–911. DOI: 10.1037/a0015817
- Brown L, Sherbenou RJ, Johnsen S. *TONI-4: Test of Nonverbal Intelligence*. 4th. Austin, TX: Pro-Ed; 2010.
- Brownell R. *Expressive one-word picture vocabulary test: Spanish-bilingual edition*. Novato, CA: Academic Therapy Publications; 2000a.
- Brownell R. *Receptive one-word picture vocabulary test: Spanish-bilingual edition*. Novato, CA: Academic Therapy Publications; 2000b.

- Cameron-Faulkner T, Lieven E, Tomasello M. A construction based analysis of child directed speech. *Cognitive Science*. 2003; 27:843–873. DOI: 10.1207/s15516709cog2706_2
- Cobo-Lewis A, Eilers R, Pearson BZ, Umbel V. Interdependence of Spanish and English knowledge in language and literacy among bilingual children. In: Eilers R, Ollers K, editors *Language and literacy in bilingual children*. Clevedon, UK: Multilingual Matters; 2002. 118–134.
- Cohen J. *Statistical power analysis for the behavioural sciences*. 2nd. Hillsdale, NJ: Erlbaum; 1988.
- Cummins J. Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*. 1979; 49:222–251. DOI: 10.3102/00346543049002222
- Cummins J. The construct of language proficiency in bilingual education. In: Atlatis J, editor *Current issues in bilingual education*. Washington, DC: Georgetown University Press; 1980. 81–103.
- Cummins J. A proposal for action: Strategies for recognizing heritage language competence as a learning resource within the mainstream classroom. *The Modern Language Journal*. 2005a; 89:585–592.
- Cummins J. TESOL Symposium on Dual Language Education: Teaching and Learning Two Languages in the EFL Setting. Alexandria, VA: TESOL; 2005b. Teaching for cross-language transfer in dual language education: Possibilities and pitfalls; 1–16.
- de Jong EJ. Effective bilingual education: From theory to academic achievement in a two-way bilingual program. *Bilingual Research Journal*. 2002; 26:65–84. DOI: 10.1080/15235882.2002.10668699
- Ebert KD, Kohnert K. Language learning impairment in sequential bilingual children. *Language Teaching*. 2016; 49:301–338. DOI: 10.1017/S0261444816000070
- Heilmann J, Miller JF, Iglesias A, Fabiano-Smith L, Nockerts A, Andriacchi KD. Narrative transcription accuracy and reliability in two languages. *Topics in Language Disorders*. 2008; 28:178–188. DOI: 10.1097/01.TLD.0000318937.39301.76
- Hickman P, Pollard-Durodola S, Vaughn S. Storybook reading: Improving vocabulary and comprehension for English-language learners. *The Reading Teacher*. 2004; 57:720–730.
- Kang JY. Decontextualized language production in two languages: An investigation of children's word definition skills in Korean and English. *Applied Psycholinguistics*. 2013; 34:211–231. DOI: 10.1017/S0142716411000671
- Kelley A, Kohnert K. Is there a cognate advantage for typically developing Spanish-speaking English-language learners? *Language, Speech, and Hearing Services in Schools*. 2012; 43:191–204. DOI: 10.1044/0161-1461(2011/10-0022)
- Kohnert K, Windsor J, Miller R. Crossing borders: Recognition of Spanish words by English-speaking children with and without language impairment. *Applied Psycholinguistics*. 2004; 25:543–564. DOI: 10.1017/S0142716404001262
- Kuperman V, Stadthagen-Gonzalez H, Brysbaert M. Age-of-acquisition ratings for 30,000 English words. *Behaviour Research Methods*. 2012; 44:978–990. DOI: 10.3758/s13428-012-0210-4
- Lidz CS. MLE components and their roots in theory and research. In: Feuerstein R, Klein PS, Tannenbaum AJ, editors *Mediated learning experience (MLE): Theoretical, psychosocial and learning implications*. London: Freund; 1991. 271–291.
- McGregor KK, Oleson J, Bahnsen A, Duff D. Children with developmental language impairment have vocabulary deficits characterized by limited breadth and depth. *International Journal of Language & Communication Disorders*. 2013; 48:307–319. DOI: 10.1111/1460-6984.12008 [PubMed: 23650887]
- McKee D, Puncel M. Elmer. London: Andersen Press; 1989.
- Melby-Lervåg M, Lervåg A. Cross-linguistic transfer of oral language, decoding, phonological awareness and reading comprehension: A meta-analysis of the correlational evidence. *Journal of Research in Reading*. 2011; 34:114–135. DOI: 10.1111/j.1467-9817.2010.01477.x
- Miller C. Structural ambiguity in the Vietnamese relative clause. *Mon-Khmer Studies*. 1976; 5:233–267.
- Miller J, Iglesias A. SALT: Systematic analysis of language transcripts [Research version]. Middleton, WI: SALT Software; 2012.

- Nagy WE, García GE, Durguno u AY, Hancin-Bhatt B. Spanish-English bilingual students' use of cognates in English reading. *Journal of Literacy Research*. 1993; 25:241–259. DOI: 10.1080/10862969009547816
- Oh JS, Fuligni AJ. The role of heritage language development in the ethnic identity and family relationships of adolescents from immigrant backgrounds. *Social Development*. 2010; 19:202–220. DOI: 10.1111/j.1467-9507.2008.00530.x
- Olswang L, Svensson L, Coggins TE, Beilinson JS, Donaldson AL. Reliability issues and solutions for coding social communication performance in classroom settings. *Journal of Speech, Language, and Hearing Research*. 2006; 49:1058–1071. DOI: 10.1044/1092-4388(2006/075)
- Ordoñez C, Carlo M, Snow C, McLaughlin B. Depth and breadth of vocabulary in two languages: Which vocabulary skills transfer? *Journal of Educational Psychology*. 2002; 94:719–728. DOI: 10.1037//0022-0663.94.4.719
- Peña E, Iglesias A, Lidz CS. Reducing test bias through dynamic assessment of children's word learning ability. *American Journal of Speech-Language Pathology*. 2001; 10:138–154. DOI: 10.1044/1058-0360(2001/014)
- Penno J, Wilkinson I, Moore D. Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew effect? *Journal of Educational Psychology*. 2002; 94:23–33. DOI: 10.1037/0022-0663.94.1.23
- Pérez AM, Peña ED, Bedore LM. Cognates facilitate word recognition in young Spanish-English bilinguals' test performance. *Early Childhood Services*. 2010; 4:55–67. [PubMed: 23565068]
- Pham G. Pathways for learning two languages: Lexical and grammatical associations within and across languages in sequential bilingual children. *Bilingualism: Language & Cognition*. 2016; 19:928–938. DOI: 10.1017/S1366728915000899
- Pham G, Kohnert K. A longitudinal study of lexical development in children learning Vietnamese and English. *Child Development*. 2014; 85:767–782. DOI: 10.1111/cdev.12137 [PubMed: 23869741]
- Richardson J. Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*. 2011; 6:135–147. DOI: 10.1016/j.edurev.2010.12.001
- Rolstad K, Mahoney K, Glass GV. The big picture: A meta-analysis of program effectiveness research on English language learners. *Educational Policy*. 2005; 19:572–594. DOI: 10.1177/0895904805278067
- Sheng L, Lam BPW, Cruz D, Fulton A. A robust demonstration of the cognate facilitation effect in first-language and second-language naming. *Journal of Experimental Child Psychology*. 2016; 141:229–238. DOI: 10.1016/j.jecp.2015.09.007 [PubMed: 26463463]
- Simpson Baird A, Palacios N, Kibler A. The cognate and false cognate knowledge of young emergent bilinguals. *Language Learning*. 2016; 6:448–470. DOI: 10.1111/lang.12160
- Snow CE. The development of definitional skill. *Journal of Child Language*. 1990; 17:697–710. DOI: 10.1017/S0305000900010953 [PubMed: 2269704]
- Uccelli P, Pérez MM. Narrative and vocabulary development of bilingual children from kindergarten to first grade: Developmental changes and associations among English and Spanish skills. *Language, Speech, and Hearing Services in Schools*. 2007; 38:225–236. DOI: 10.1044/0161-1461(2007/024)
- Umbel V, Oller D. Developmental changes in receptive vocabulary in Hispanic bilingual school children. *Language Learning*. 1994; 44:221–242. DOI: 10.1111/j.1467-1770.1994.tb01101.x
- Umbel V, Pearson B, Fernandez M, Oller D. Measuring bilingual children's receptive vocabularies. *Child Development*. 1992; 63:1012–1020. [PubMed: 1505238]
- Zientek LR, Thompson B. Matrix summaries improve research reports: Secondary analyses using published literature. *Educational Researcher*. 2009; 38:343–352. DOI: 10.3102/0013189X09339056

Figure 1. Children's performance on English translations of the target words (communicative adequacy) across groups from pre-to posttraining. Error bars represent ± 1 standard error.

Figure 2. Children's performance on English cognate nouns (communicative adequacy) across groups from pre-to posttraining. Error bars represent ± 1 standard error.

Table 1

Descriptive statistics for participant groups

Variable	LVE		LSE		HSE	
	M	SD	M	SD	M	SD
Composition (boys/girls)	18 (9/9)		18 (9/9)		18 (8/10)	
Age (years)	7.3	0.7	7.3	0.6	7.3	0.5
Born in the United States (%)	61		100		83	
Maternal education (% high school)	82		83		77	
TONI standard score	105	8	94	10	98	7
L1 use at home (4-point scale; 1 = <i>all the time</i>)	1.4	0.5	1.3	0.5	1.2	0.4
Age of English onset (years)	3.6	1.7	2.1	2.4	3.4	1.6
L1 receptive vocabulary (% correct)	49.8	18.0	42.8	14.5	62.9	12.6
English receptive vocabulary (% correct)	53.5	12.1	49.9	12.8	52.0	9.2
L1 expressive vocabulary (% correct)	16.5	11.1	13.1	8.0	31.3	14.9
English expressive vocabulary (% correct)	41.6	12.4	27.8	13.5	31.0	12.4

Note. LVE = Vietnamese-English group with low L1 proficiency; LSE = Spanish-English group with low L1 proficiency; HSE = Spanish-English group with high L1 proficiency; TONI = Test of Nonverbal Intelligence, 4th edition (Brown et al., 2010).

Table 2

Pre- and posttraining measures for words in children's L1 by group

Measure	Pretraining						Posttraining					
	LVE			LSE			HSE			LVE		
	M	SD	M	M	SD	M	M	SD	M	M	SD	M
Target words												
PicID	5.8	1.4	3.7	1.9	4.3	1.7	7.1	**	0.8	5.9	***	1.7
SK	3.3	4.5	2.8	3.2	6.9	2.9	8.9	*	7.5	7.3	**	6.5
PKReal	1.3	2.1	0.0	0.0	1.7	3.8	3.9	*	5.3	4.7	***	4.7
CA	1.8	3.1	2.3	2.2	5.4	4.4	6.9	*	6.5	4.0	*	4.3
Control words												
SK	8.3	5.9	14.1	9.3	11.8	4.1	10.1	5.3	10.4	7.0	11.3	4.4
PKReal	1.7	3.4	1.1	2.1	3.2	4.8	3.2	6.6	1.7	2.4	3.7	3.9
CA	5.3	3.6	7.6	3.7	9.4	3.4	5.8	2.8	8.0	3.4	9.1	3.7

Note. PicID = Picture identification (maximum score of 8); SK = Syntagmatic knowledge (unlimited); PKReal = Paradigmatic knowledge of real superordinates (maximum score of 40); CA = Communicative adequacy (maximum score of 24).

* $p < .05$;

** $p < .01$;

*** $p < .001$.

Table 3
Pre- and posttraining measures for words translated into children’s L2 English by group

Measure	Pretraining						Posttraining					
	LVE			HSE			LVE			HSE		
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Target words												
PicID	5.5	1.2	5.1	1.3	5.8	1.5	6.2	1.0	5.4	1.6	6.2	1.8
SK	11.5	5.6	11.8	4.7	10.9	6.3	13.5	6.8	13.1	6.4	12.8	4.8
PKReal	3.4	4.6	2.1	3.0	2.5	3.1	3.5	4.4	2.7	2.5	4.3 [*]	4.4
CA	5.9	3.8	5.3	2.7	6.1	3.3	6.8	4.4	5.7	3.0	8.9 ^{***}	3.8
Control words												
SK	18.7	9.3	19.3	6.9	16.7	5.0	16.0	5.7	16.7	6.0	15.6	4.1
PKReal	3.9	3.9	0.8	1.9	2.9	3.3	3.4	4.7	1.1	2.1	3.7	5.5
CA	7.4	4.2	9.3	2.8	8.6	2.5	8.8	3.5	9.3	3.4	8.1	2.0

Note. PicID = Picture identification (maximum score of 8); SK = Syntagmatic knowledge (unlimited); PKReal = Paradigmatic knowledge of real superordinates (maximum score of 40); CA = Communicative adequacy (maximum score of 24).

* $p < .05$;

** $p < .01$;

*** $p < .001$.