



Relations between oral language skills and English reading achievement among Spanish–English bilingual children: a quantile regression analysis

J. Marc Goodrich¹ · Lisa Fitton² · Lauren Thayer¹

Received: 30 August 2021 / Accepted: 4 May 2022 / Published online: 28 May 2022
© The Author(s), under exclusive licence to The International Dyslexia Association 2022

Abstract

Understanding factors that influence reading achievement among bilingual children is considerably more complex than it is for monolingual children. Research on dual language development indicates that bilingual children’s oral language abilities are often distributed across languages in varied ways, due to heterogeneity of dual language exposure and input. Consequently, there may be greater variability in the associations between oral language proficiency and reading achievement among bilingual children than there is for monolingual children. This study evaluated how vocabulary knowledge and morphosyntactic ability in Spanish and English were associated with English reading achievement among 117 bilingual kindergarten and first grade children in the USA using both OLS and quantile regression. Results indicated that although English vocabulary and morphosyntax were both significantly associated with reading achievement, English vocabulary knowledge was most strongly associated with reading at higher quantiles of reading achievement. Cross-language analyses indicated that both Spanish vocabulary and morphosyntax made significant contributions to predicting English reading achievement beyond the effects of English oral language. Spanish vocabulary was uniquely predictive of reading at high and low quantiles of English reading, whereas relations between Spanish morphosyntax and English reading did not differ across quantiles. These results were consistent with predictions derived from theoretical models such as the simple view of reading and suggest that Spanish vocabulary knowledge may provide more unique information about children’s underlying capacity for acquiring language and literacy skill than does morphosyntax.

Keywords Bilingual · Oral language · Quantile regression · Reading

Theoretical frameworks of reading development highlight the importance of oral language ability in skilled reading (e.g., Duke & Cartwright, 2021; Hoover & Gough, 1990), which

✉ J. Marc Goodrich
marcgoodrich@tamu.edu

¹ Department of Teaching, Learning, & Culture, Texas A&M University, College Station, TX, USA

² Communication Sciences & Disorders Department, University of South Carolina, Columbia, SC, USA

has been supported by decades of empirical evidence (e.g., Kendeou et al., 2009; Lonigan et al., 2018). As children decode words to read text, they must integrate various language skills, including knowledge of word meanings, morphological knowledge (e.g., how prefixes and suffixes change the meanings of words), knowledge of sentence structures, and knowledge of how text is constructed (e.g., stories have a beginning, middle, and end), to form a coherent representation of the text for comprehension. Although evidence indicates that the relations between oral language ability and reading achievement vary across languages and throughout development (Florit & Cain, 2011), there is consistent evidence that oral language proficiency is a key component of skilled reading. Even early oral language skills (i.e., oral language skills measured in preschool or kindergarten) are longitudinally predictive of children's future reading achievement (e.g., Lepola et al., 2016), including among Spanish–English bilingual children (Rojas et al., 2019).

There are several factors that complicate how measures of oral language ability may be used to predict risk for reading difficulty among bilingual children. First, and most evident, is the acquisition of two languages. If a monolingual child has low oral language proficiency, it is likely that they will have difficulty with reading (e.g., Snowling et al., 2020). However, for bilingual children, measuring only one of their languages is insufficient, as proficiency in each language depends on several factors, including relative amounts of input and use of each language. For example, a bilingual child may have high proficiency in their first language (L1) but low proficiency in their second language (L2). Or the child may have low proficiency in both their L1 and L2. If we wish to predict risk for reading difficulty from oral language proficiency, assessment of oral language ability in both the L1 and L2 is essential, as low L2 proficiency could be due to lack of exposure and/or opportunity to acquire L2. Thus, reliance on L2 oral language assessment may result in overidentification of risk for reading difficulty. A true difficulty with acquiring oral language and/or literacy skills would affect skills in both languages.

Second, we currently lack appropriate standardized assessments to measure oral language and reading among bilingual children reliably (Boerma & Blom, 2017). Although it is common in educational practice to simply adapt measures developed for monolingual English speakers into Spanish, this approach does not ensure that the measure has adequate reliability and validity for measuring oral language among Spanish–English *bilinguals* (Karem et al., 2019). There is a critical need for measures that are normed on bilingual children in the USA. Additionally, there may be fundamental differences in how oral language proficiency manifests across monolingual and bilingual children (Kapantzoglou et al., 2012). For example, when given a measure of English expressive vocabulary, an English monolingual child may know more words than a Spanish–English bilingual child because of differences in the amount of exposure to English. Therefore, the bilingual child may score lower than the monolingual child because of fewer opportunities to acquire proficiency in English, rather than a lower underlying language ability. In fact, dual language assessment may reveal that the Spanish–English bilingual child knows the same number of words as the monolingual child, with word knowledge being distributed across the child's two languages. This nuance would be missed—and the bilingual child artificially labeled with low oral language proficiency—if a measure designed for assessing oral language among monolingual speakers was the only tool used to measure oral language proficiency for a bilingual child. There are also cultural considerations when adapting measures across languages. A word may be common in an English-speaking household but uncommon in a Spanish-speaking household and vice versa. Assessing both languages with a measure that takes these issues into consideration is important to accurately identify bilingual students with low oral language

proficiency who may be at risk for reading difficulties (Castilla-Earls et al., 2020; Francis, 2019).

Assessing oral language among bilingual children

Oral language is a complex developmental construct, which is further complicated by exposure to and use of multiple languages during early development. A series of recent studies conducted by the Language and Reading Research Consortium (LARRC) have documented the nature of oral language ability and how it relates to reading in young children, including Spanish–English bilingual children. In an initial study examining dimensionality of English in monolingual children, LARRC (2015a) reported that dimensionality of oral language skill changes across the early childhood years, such that in preschool and kindergarten, language is best represented as a unidimensional construct, but that by first grade, children’s discourse-level language skills (e.g., narrative language use) emerge as distinct constructs from lower-level language abilities (e.g., vocabulary, grammatical knowledge). This suggests that as young children acquire oral language skills, it becomes increasingly important to consider diverse aspects of oral language, rather than solely focusing on one aspect of language development (e.g., focusing on vocabulary knowledge in the preschool years but shifting to include other aspects of language such as grammatical knowledge or narrative ability in school-age children).

It is yet unclear if bilingual children follow the same developmental trajectories as monolingual children. Studies specifically investigating Spanish–English bilingual children’s oral language skills indicate that vocabulary knowledge, grammar, and higher-level language skills (particularly in Spanish) are separable as early as preschool or kindergarten (LARRC, 2015b; LARRC & Chiu, 2018). Additional evidence from LARRC indicates that preschool Spanish oral language contributes directly to English reading achievement (after controlling for English oral language and alphabet knowledge; LARRC et al., 2019; LARRC et al., 2021). Consequently, it is important to assess skills beyond vocabulary knowledge and consider bilingual children’s oral language skills in both of their languages when assessing oral language ability to predict reading achievement. Recent research suggests that for younger bilingual children, measures of morphosyntax knowledge are the strongest indicators of developmental language disorder (Peña et al., 2020). Therefore, pairing assessment of vocabulary with morphosyntax in Spanish and English should represent a relatively comprehensive approach to evaluating young bilingual children’s oral language skills.

Although it is important to assess language broadly, it is likely that certain measures of oral language represent stronger cross-language predictors of reading achievement than others. One factor that may help explain the interaction between oral language and reading is the hypothesis of cross-linguistic transfer (Cummins, 1979). This hypothesis states that concepts and topics learned in one language (e.g., L1) can be transferred and utilized in their other language (e.g., L2). When transfer occurs, a bilingual child—with the appropriate vocabulary—should be able to use what they learned in their L1 and express it in their L2 without ever having explicit instruction of that concept in their L2. For example, if a bilingual child learns how to count in their L1, they would not have to learn the concept of counting again in their L2 but just need to learn the names of the numbers. According to Cummins (1981), the mechanism by which transfer occurs is a *common underlying proficiency* about language or academic

skills that develops with acquisition of the skills itself. In the above example, learning to count in one language facilitates acquisition of general knowledge of principles of counting that can be applied in a new language. Transfer of skills that are relatively language-independent (e.g., counting, phonological awareness) may be more likely to occur than transfer of discrete knowledge that is highly specific to a given language (e.g., vocabulary). Consequently, young bilingual children's oral language abilities that represent underlying knowledge about language or an underlying capacity for acquiring language may have utility in predicting risk for reading difficulty.

Two commonly assessed oral language skills that might be important for predicting risk for reading difficulty are vocabulary knowledge and morphosyntactic ability. Bilingual children's vocabulary knowledge is highly specific to a given language (with the exception of cognates) and largely dependent on relative exposure to each language. Prior research has consistently reported that cross-language correlations of vocabulary among Spanish–English bilingual children are very small or even negative (e.g., Goodrich & Lonigan, 2018; Melby-Lervåg & Lervåg, 2011), supporting the idea that vocabulary is highly dependent on exposure. Consequently, individual differences in L1 vocabulary may not be particularly informative in the prediction of L2 reading skill. In contrast, measures that tap knowledge of grammatical domains such as syntax and morphology may be more indicative of children's underlying ability to acquire language. Evidence indicates that measures of morphosyntactic knowledge effectively discriminate between bilingual children with and without language disorders (e.g., Lazechnik et al., 2019; Paradis et al., 2013; Peña et al., 2020), suggesting that these tools are reflective of language learning ability. Therefore, measures of grammatical knowledge may be more predictive of reading across languages than are measures of language-specific skills such as vocabulary knowledge. Prior meta-analytic evidence indicates that cross-linguistic relations between oral language and reading among bilingual children vary depending on the type of oral language measure used (Prevo et al., 2016); however, the relations were not consistent with theoretical predictions derived from theory of cross-language transfer, as Prevo et al. reported that L1 vocabulary was more highly associated with L2 reading than was general language proficiency in L1. Consequently, additional research is needed to determine which indicators of oral language proficiency are the best early indicators of reading among bilingual children.

It is also possible that cross-language relations between oral language and reading vary for children with different levels of reading achievement. Among studies that evaluate the contributions of L1 and L2 oral language to reading achievement, evidence consistently indicates that within-language relations between oral language and reading are stronger than cross-language relations (LARRC et al., 2021; Leider et al., 2013; Prevo et al., 2016; Swanson et al., 2008). However, early in school, many Spanish–English bilingual children have only had limited experiences that would facilitate development of English oral language and reading skill. Consequently, low English reading achievement may be reflective of lack of opportunity to acquire English reading skill, rather than an underlying reading difficulty or disorder. In this case, it is possible that the relation between Spanish oral language and English reading is stronger for children with poor English reading skill than it is for children with strong English reading skills. Quantile regression (Koenker & Bassett, 1978) represents one approach through which to analyze whether the relation between bilingual oral language and reading changes across the continuum of reading achievement (i.e., at different quantiles of reading skill).

Current study

The purpose of this study was to evaluate differential relations between oral language ability and English reading achievement among Spanish–English bilingual children. Because oral language ability is more complex for bilingual children than it is for monolingual children, we were particularly interested in understanding how different aspects of oral language are related to reading achievement at different levels of English reading achievement. To do so, we measured Spanish and English oral language through assessments of vocabulary knowledge and morphosyntactic ability. We used quantile regression to examine which oral language skills were uniquely predictive of English reading at low (0.25 quantile), medium (0.50 quantile), and high (0.75 quantile) levels of English reading achievement. We addressed the following research questions:

RQ1: Does the relation between English oral language and English reading achievement differ for children with different levels of English reading achievement?

Hypothesis 1: We hypothesized that all measures of English oral language would more strongly predict English reading at high quantiles of English reading. Low English reading achievement among bilingual children may be indicative of limited exposure to English reading instruction. Limited exposure to English reading instruction should result in poor English decoding (i.e., word reading) skills. In the absence of sufficient decoding skills, children’s ability to use oral language ability during reading is limited.

RQ2: Does the relation between Spanish oral language and English reading achievement differ for children with different levels of English reading achievement?

Hypothesis 2a: We expected that Spanish oral language would more strongly predict English reading achievement at low quantiles of English reading. Children with lower levels of English reading achievement may also have relatively low English proficiency relative to Spanish proficiency (i.e., Spanish-dominant children). Therefore, we expected that variation in Spanish oral language at low quantiles of English reading will better reflect underlying ability to acquire language and literacy skills (whereas variation in English oral language at low quantiles of English reading may be an issue of opportunity/exposure).

Hypothesis 2b: We expected that oral language skills indicative of underlying language learning ability (e.g., morphosyntactic skills) would demonstrate greater cross-language prediction of English reading achievement than would skills that are highly dependent on exposure to a specific language (e.g., vocabulary knowledge).

RQ3: Is Spanish oral language associated with English reading achievement above and beyond the effect of English oral language at different quantiles of English reading achievement?

Hypothesis 3: Consistent with hypotheses for RQ1 and RQ2, we expected that Spanish oral language would only uniquely contribute to English reading achievement (beyond the effect of English oral language) at low quantiles of English reading achievement. We expected this trend to be consistent across both Spanish morphosyntax and Spanish vocabulary.

Method

Participants

The participant sample included a subsample of 117 Spanish–English speaking children enrolled in a larger study of bilingual language and reading development conducted in South Carolina and Nebraska during the 2019–2020 academic school year. The children were enrolled in kindergarten ($n=86$) or first grade ($n=31$) at the time of participation and were between the ages of 5–7 years ($M=6.28$, $SD=0.68$). Of the 117 participants, 76 were recruited from one South Carolina school, and 41 were recruited across ten Nebraska schools. The differences in recruitment rates were attributable to two factors. First, there is a greater density of Spanish-speaking children in the Midlands of South Carolina compared to southeastern Nebraska. Second, different consent procedures governed each site. Consistent with each site’s Institutional Review Board approval, passive consent was used in South Carolina, whereas active consent was used in Nebraska. Consequently, in addition to fewer eligible students per school in Nebraska, only some eligible students participated at each school in Nebraska, whereas nearly all eligible students at the South Carolina school participated.

The children were identified as Spanish–English speakers by their classroom teachers and were recruited to participate following procedures approved by the respective site Institutional Review Board. Children were recruited equally regardless of developmental language status or eligibility classification, to maximize the representativeness of the sample and for consistency with the heterogeneity that exists within the larger Spanish–English speaking population in the USA.

School demographics

In South Carolina, the participating school was a federal Title I public school and served students in preschool (4 K) through 5th grade during the 2019–2020 academic year. The school was located in a large suburb and served a total of 1,143 students with a student–teacher ratio of 15.45. The racial/ethnic makeup of the school included students whose parents identified their background(s) as the following: American Indian/Alaska Native (0.2%), Asian (13.1%), Black (23.8%), Hispanic/Latino (37.3%), White (33.3%), and two or more races (4.1%). School administrators reported that all enrolled students were from economically disadvantaged homes based on free/reduced price lunch qualification and reported household income.

In Nebraska, participants came from 10 schools. Twenty-five students were enrolled in a single school. All other schools had four or fewer children who participated. Among the seven schools for which data were available in the Common Core of Data (three schools were private schools), all were designated as Title I schools and served students in either preschool through 4th grade or preschool through 5th grade. Three schools were in an urban area, and four were in rural areas. Student enrollment ranged from 224 to 707, and student teacher ratio ranged from 9.87 to 15.92. Racial/ethnic makeup was as follows: American Indian/Alaska Native (0.0 to 2.0%), Asian (0.0 to 7.8%), Black/African American (0.4 to 11.9%), Hispanic/Latino (14.3 to 88.8%), White (7.4 to 68.4%), and two or more races (0.0 to 14.2%).

Procedure

Participants completed a battery of Spanish–English bilingual oral language measures and English reading assessments during the middle of the school year. The tests were administered by trained research assistants with native or near-native fluency in the language of assessment. Children were randomly assigned to Spanish-first or English-first test administration and completed the full assessment battery within a 2-week time window. Assessments were administered at children’s elementary schools, either during or after school hours.

Bilingual language measures

To assess students’ oral language abilities in Spanish and English, measures of vocabulary knowledge and morphosyntactic ability were administered. To measure vocabulary, the *Expressive One-Word Picture Vocabulary Test-4 Spanish-Bilingual Edition* (EOWPVT-4 SBE; Martin, 2013) was administered separately in Spanish and English following recommendations provided by Anaya et al. (2018) and Gross et al. (2014). For this assessment, children are asked to name pictures they are shown. Based on their responses, three sets of scores were obtained: Spanish-only vocabulary, English-only vocabulary, and conceptual vocabulary. Conceptual vocabulary scores are based on correct responses in Spanish or English, so that the child receives credit for each concept known in either language. For example, when assessed in Spanish and English, a child could know a word in Spanish only, English only, both languages, or neither language. In a conceptual vocabulary scoring framework, all children who know the word in at least one language receive equal credit (i.e., a child who knows both *apple* and *manzana* receives equal credit as a child who knows *manzana* but not *apple*). Consequently, conceptual vocabulary is thought to represent children’s complete vocabulary knowledge more accurately than measuring vocabulary in Spanish or English alone (e.g., Ehl et al., 2020). The EOWPVT-4 SBE was normed for administration with Spanish–English-speaking individuals in the USA and has high internal consistency reliability ($\alpha=0.95$).

To measure children’s bilingual morphosyntactic ability, the Sentence Repetition subtest of the *Bilingual English–Spanish Assessment* (BESA; Peña et al., 2014) was administered separately in Spanish and English. For this task, children are asked to repeat sentences verbatim and receive scores based on the accuracy of their repetitions. There is strong evidence for the validity of sentence repetition tasks as measures of morphosyntactic ability for children who speak multiple languages (Fitton et al., 2019; Pratt et al., 2021). The BESA Sentence Repetition subtest has high internal consistency ($a=0.96$ in Spanish and $a=0.95$ in English).

English reading assessment

Subtests from the reading cluster of the *Woodcock-Muñoz Language Survey III* (WMLS III; Woodcock et al., 2017) were administered to measure participants’ English reading achievement. The WMLS III is specifically designed to assess the academic language proficiency, including reading, of Spanish–English speaking children and has parallel forms in English and Spanish. For the present work, only English reading was assessed given that children were enrolled in English-only schools that prioritize English literacy development.

Children completed the Letter-Word Identification and Passage Comprehension subtests, which were administered individually by research assistants. The Letter-Word Identification subtest focuses on basic decoding skills including pointing to named letters or words, expressively labeling letters or words, and reading English words aloud. The Passage Comprehension subtest focuses on text comprehension skills including print awareness via pointing to named symbols, matching printed words to pictures, and expressively filling in the blank of printed sentences or paragraphs. Subtest-specific standardized scores and composite standardized reading scores were obtained for each child based on age norms.

Children were selected for inclusion in the present study based on completion of both subtests of the WMLS III prior to school shutdowns due to the COVID-19 pandemic in March 2020. Missing data on predictor variables is attributable to school closures that ended data collection prematurely, equally impacting all participating students. No students dropped out of the study. For transparency, missing data rates are provided in the results.

Analytic approach

Analyses and data visualizations were conducted in R Version 3.6.3 (R Core Team, 2020) using the packages *summarytools* (Comtois, 2021), *psych* (Revelle, 2021), *quantreg* (Koenker, 2021), *ggplot2* (Wickham, 2016), and *ggpubr* (Kassambara, 2020). The *summarytools* and *psych* packages were used to examine univariate and bivariate descriptive statistics for all raw and standardized, norm-referenced scores obtained from the measures of interest. The packages *ggplot2* and *ggpubr* were used to visually inspect data throughout the analytic process (e.g., residual plotting). All quantile regression analyses (described in further detail below) were conducted using the *quantreg* package. Ordinary least squares (OLS) regressions were completed using base R functions (e.g., `lm`).

To prepare data for analyses, child age was regressed out of the raw scores. This approach was used, rather than analyzing data separately by age, for two reasons. First, children in both kindergarten and first grade receive instruction in foundational reading skills. Specifically, in both Nebraska and South Carolina, state English/Language Arts standards emphasize teaching and achievement of basic decoding skills using phonics in kindergarten. Second, our sample size is already relatively small, and reducing the analytic sample further by separating by grade level (e.g., only 31 first grade students) would likely produce unreliable results. The values were then z-scored to provide a standardized scale for interpretation ($M=0$, $SD=1.0$). For English reading, the z-scored values from the Letter-Word Identification subtest and Passage Comprehension subtest were averaged to create a single composite score for English reading. These z-scored values were used in the analyses.

To address the first research question, we first conducted single OLS regressions to examine the predictive relation between the English oral language measures and the English reading composite score. We examined English sentence repetition and English vocabulary separately as predictors of English reading. Then, we re-evaluated the models within a quantile regression framework using the Barrondale-Roberts algorithm to obtain estimates and associated confidence intervals (Koenker, 2021; Koenker & d'Orey, 1994). *P*-values were computed using Markov chain marginal bootstrapping (He & Hu, 2002), given that quantile regression does have any distributional assumptions for the dependent variable residuals, aside from continuity (Koenker & Hallock, 2001). We obtained regression estimates at the 0.25, 0.50, and 0.75 quantiles of English reading and conducted comparisons between the estimates at each quantile, using the Benjamini and Hochberg (1995)

procedure to correct for multiple comparisons with a false discovery rate of $Q=0.10$. The same procedures were used to address the second research question but with the Spanish oral language measures (i.e., Spanish sentence repetition and Spanish vocabulary) as predictors of English reading.

To address the final research question, we conducted both OLS and quantile regressions to examine the relations between the Spanish language measures beyond the English language measures for predicting English reading. We examined Spanish sentence repetition and Spanish vocabulary separately as predictors of English reading, while accounting for both English sentence repetition and English vocabulary as covariates. To assess the influence of possible multicollinearity between the English language predictors, we computed a variance inflation factor (VIF) for each estimate in the models.

For all modeling, missing data was addressed using casewise deletion, given that the missing data was entirely attributable to the beginning of the COVID-19 pandemic. No patterns were observed in the missingness of the data, and no participants dropped out of the study. To assess the robustness of this approach, sensitivity analyses were conducted based on a complete dataset including imputed values. Results were compared across each dataset for consistency.

Results

Descriptive statistics for the participants' standardized scores are provided by grade and by state in Table 1. Correlations and missing data rates for z-scored values used in the analyses are in Table 2. Students' English reading scores followed an overall normal distribution, though average performance within the sample ($M=87.18$, $SD=13.20$) was below the normative mean for the WMLS III. Within-language correlations among oral language skills were large and statistically significant, indicating that children with stronger vocabulary in a given language also had stronger morphosyntactic skills in that language. In contrast, there were no significant cross-language correlations among the oral language measures. All oral language measures were significantly correlated with English reading, with the exception of Spanish vocabulary knowledge. Within-language correlations between oral language and reading were stronger than were cross-language correlations. However, the positive correlation between Spanish morphosyntax and English reading indicates that a measure of general language ability such as morphosyntax may be a better indicator of English reading, consistent with our hypothesis (see hypothesis 2b).

English language predicting English reading

Results for the first research question are provided in the top half of Table 3. Single OLS regression revealed that English vocabulary alone predicted 28.6% of the variance in students' English reading scores. A 1.0 standard deviation increase in English vocabulary corresponded with a 0.51 (95% CI [0.37, 0.66], $p<0.001$) standard deviation increase in English reading scores. Quantile regression revealed, however, that this predictive relation was stronger at the 0.75 quantile of English reading compared to the lower quantiles of English reading, as indicated by a significant difference between the 0.75 and 0.50 quantile estimates: $F(1, 231)=4.63$, $p=0.032$. For children performing at the 0.75 quantile of English reading, English vocabulary predicted reading at 0.64 (95% CI [0.42, 0.82], $p<0.001$), whereas at the 0.50

Table 1 Descriptive statistics for standardized scores

Standardized assessments	Mean (SD) by grade			Mean (SD) by state			Full dataset		
	Kindergarten	First grade		South Carolina	Nebraska		Min	Max	Mean (SD)
English language measures									
Expressive vocabulary ¹	92.53 (21.58)	98.57 (19.38)		91.52 (21.60)	98.80 (18.89)		55	145	94.09 (20.89)
Morphosyntax ²	90.55 (18.39)	95.77 (14.56)		87.04 (16.75)	96.83 (18.12)		55	115	91.26 (17.94)
Spanish language measures									
Expressive vocabulary ¹	79.14 (19.42)	78.84 (16.22)		80.37 (17.10)	76.51 (21.08)		55	129	79.06 (18.54)
Morphosyntax ²	87.47 (17.23)	90.00 (14.43)		88.49 (15.30)	86.92 (18.83)		55	120	87.83 (16.81)
English reading measures ³									
Letter-word identification	83.57 (15.75)	88.35 (16.27)		84.64 (15.65)	85.20 (16.71)		50	126	84.84 (15.96)
Passage comprehension	90.60 (10.77)	90.06 (13.16)		91.66 (10.85)	88.24 (12.16)		55	115	90.46 (11.40)
Reading composite	86.45 (12.59)	89.19 (14.80)		87.63 (12.83)	86.34 (13.99)		52	120	87.18 (13.20)
Descriptive measures									
Conceptual Vocabulary ¹	101.75 (17.44)	104.23 (13.13)		100.77 (16.44)	105.54 (16.05)		55	145	102.40 (16.39)

¹Scores obtained from the Expressive One-Word Picture Vocabulary Test 4 Spanish-Bilingual Edition (Martin, 2013)

²Scores obtained from the sentence repetition task of the Bilingual English-Spanish Assessment (Peña et al., 2014)

³Scores obtained from the Woodcock Muñoz Language Survey III (Woodcock et al., 2017)

Table 2 Correlation table and missing data rates for z-scored variables

Assessment	% Missing	1	2	3	4
1. English vocabulary	0.9%				
2. English morphosyntax	6.0%	.74** [.64, .81]			
3. Spanish vocabulary	1.7%	-.07 [-.25, .11]	-.04 [-.23, .15]		
4. Spanish morphosyntax	11.1%	-.03 [-.22, .17]	.20 [-.01, .38]	.70** [.59, .79]	
5. English reading composite	0%	.54** [.40, .66]	.50** [.35, .63]	.18 [-.01, .35]	.23* [.04, .40]

Values in square brackets indicate the 95% confidence interval for the sample correlation. Bolded values meet $p < .05$ criteria for statistical significance. * indicates $p < .05$. ** indicates $p < .01$

Table 3 Results for single OLS and quantile regressions

OLS regressions					Quantile regressions			Model R ²
	Estimate	Std. error	<i>t</i> value	Pr(> t)	0.25	0.50	0.75	
(Intercept)	-0.01	0.07	-0.19	.851	-0.50	-0.04	0.42	28.6%
English vocabulary	0.51**	0.07	6.86	<.001	0.45**	0.48**	0.64**	
(Intercept)	0.01	0.08	0.10	.922	-0.56	-0.01	0.54	24.5%
English morphosyntax	0.47**	0.08	6.02	<.001	.40**	.45**	.46**	
(Intercept)	0.01	0.09	0.16	.877	-0.61	-0.02	0.65	2.4%
Spanish vocabulary	0.17	0.09	1.95	.053	0.05	0.09	0.30	
(Intercept)	0.02	0.09	0.26	.798	-0.64	0.03	0.67	4.2%
Spanish morphosyntax	0.22*	0.09	2.35	.021	0.25*	0.14	0.13	

* indicates $p < .05$. ** indicates $p < .01$

The estimate for Spanish morphosyntax predicting English reading at the 0.25 quantile of English reading was not robust to sensitivity analyses based on missing data procedures. Analyses based on the imputed dataset indicated that this estimate was not statistically significant: 0.15, $p = .236$. Quantile regression p -values were computed using a bootstrapping procedure (He & Hu, 2002)

and 0.25 quantiles, English vocabulary predicted reading at estimates of 0.48 (95% CI [0.35, 0.61], $p < 0.001$) and 0.45 (95% CI [0.38, 0.51], $p < 0.001$), respectively.

Estimates for English morphosyntax predicting English reading were more stable across the quantiles of reading, as shown in the second line of models in Table 3. OLS regression yielded an estimate of 0.47 (95% CI [0.31, 0.62], $p < 0.001$) and R² of 24.5% for English morphosyntax predicting reading. Similar estimates were obtained at the 0.25, 0.50, and 0.75 quantiles of English reading: 0.40 (95% CI [0.25, 0.67], $p < 0.001$), 0.45 (95% CI [0.33, 0.50], $p < 0.001$), and 0.46 (95% CI [0.24, 0.69], $p = 0.003$) respectively. No significant differences in the estimates by quantile were observed.

Spanish language predicting English reading

Results for the second research question are provided in the lower half of Table 3. The OLS regression model revealed a generally positive association between Spanish vocabulary and English reading, though the estimate did not meet criteria for statistical significance: 0.17 (95% CI [-0.01, 0.35], $p=0.053$) with $R^2=2.4\%$. Quantile regression results revealed limited evidence of a relation between Spanish vocabulary and English reading at the 0.25 or 0.50 quantiles of reading, with estimates not meeting criteria for significance: 0.05 (95% CI [-0.11, 0.31], $p=0.731$) and 0.09 (95% CI [-0.02, 0.28], $p=0.410$) respectively. At the 0.75 quantile of English reading, the estimate was 0.30 (95% CI [-0.04, 0.49], $p=0.069$), suggesting that any association between Spanish vocabulary and English reading occurred only among children scoring in the upper quantiles of English reading.

OLS results for Spanish morphosyntax predicting English reading, however, did reveal an overall positive predictive relation, with a 1.0 standard deviation increase in Spanish morphosyntax corresponding with a 0.22 (95% CI [0.03, 0.40], $p=0.021$) standard deviation increase in English reading with $R^2=4.2\%$. Estimates obtained from the quantile regression models suggested a stronger relation between Spanish morphosyntax and English reading at the lowest quantile of reading (0.25, 95% CI [0.01, 0.36], $p=0.043$), though the significance of this estimate was not robust across sensitivity analyses for missing data. In all analyses, neither the estimate at the 0.50 nor the 0.75 quantile of English reading met criteria for statistical significance: 0.14 (95% CI [-0.06, 0.41], $p=0.385$) and 0.13 (95% CI [-0.08, 0.46], $p=0.383$), respectively.

Multiple predictors of English reading

Examination of estimates and variance inflation factors revealed evidence of multicollinearity between English vocabulary and morphosyntax for predicting English reading. Specifically, the estimates for vocabulary and morphosyntax varied substantially across models, the standard errors for the estimates were large, and VIF values for the English language measures were all above 2.25. To address this, a composite score for English language was constructed by averaging students' z-scores for English vocabulary and morphosyntax. This single composite English language score was then used in the statistical models conducted to address the third research question.

Spanish vocabulary predicting English reading beyond English language

Results from the multiple regression models are provided in Table 4. The top row of the table provides results from the first set of models, for which we examined Spanish vocabulary as a predictor of English reading while accounting for English language. OLS results revealed Spanish vocabulary as a unique contributor above and beyond the English language measures (0.21, 95% CI [0.07, 0.35], $p=0.004$). This estimate was similar to that obtained from the single OLS regression model of Spanish vocabulary predicting English reading. English language also significantly predicted English reading in the OLS model (0.61, 95% CI [0.46, 0.76], $p<0.001$). The overall model yielded an adjusted R^2 of 37.6%, with English language alone predicting 31.9% of the variance in English reading and Spanish vocabulary contributing an additional 5.7% to the model R^2 .

Table 4 Results for multiple OLS and quantile regressions

OLS regressions					Quantile regressions			Model R ²
	Estimate	Std. error	<i>t</i> value	Pr(> <i>t</i>)	0.25	0.50	0.75	
(Intercept)	0.02	0.07	0.28	.782	-0.50	-0.04	0.51	37.6%
English language	0.61**	0.08	8.05	< .001	0.52**	0.49**	0.69**	
Spanish vocabulary	0.21**	0.07	2.95	.004	0.17**	0.11	0.24*	
(Intercept)	0.02	0.07	0.22	.825	-0.51	-0.03	0.52	37.9%
English language	0.58**	0.08	7.51	< .001	0.51**	0.48**	0.61**	
Spanish morphosyntax	0.17*	0.07	2.22	.028	0.12	0.06	0.13	

* indicates $p < .05$. ** indicates $p < .01$

For Spanish morphosyntax predicting English reading in the multiple OLS regression, the estimate was not robust to sensitivity analyses based on missing data procedures: 0.14, $p = .054$. Quantile regression p -values were computed using a bootstrapping procedure (He & Hu, 2002)

Multiple quantile regression revealed that English language was a significant, consistent predictor of English reading, with estimates of 0.52 (95% CI [0.48, 0.68], $p < 0.001$), 0.49 (95% CI [0.43, 0.80], $p < 0.001$), and 0.69 (95% CI [0.51, 0.74], $p < 0.001$) at the 0.25, 0.50, and 0.75 quantiles, respectively. However, some differences in the predictive relations between Spanish vocabulary and English reading were observed at the different quantiles of reading. Spanish vocabulary significantly contributed to predicting English reading above and beyond English language at the 0.25 and 0.75 quantiles of English reading but was marginally predictive at the 0.50 quantile. At the 0.25 quantile, a 1.0 standard deviation increase in Spanish vocabulary corresponded with a 0.17 (95% CI [0.10, 0.26], $p = 0.012$) standard deviation increase in English reading, holding English language constant. At the 0.75 quantile, Spanish vocabulary also was a significant, positive predictor of reading: 0.24 (95% CI [0.01, 0.34], $p = 0.019$). At the 0.50 quantile, Spanish vocabulary did not meet criteria as a statistically significant predictor of English reading based on the p -value computed via Markov chain marginal bootstrapping (0.11, 95% CI [0.02, 0.30], $p = 0.286$). The 95% confidence intervals, computed based on an inversion of a rank test (Koenker and d'Orey, 1994), suggested at most a small association between Spanish vocabulary and English reading.

Spanish morphosyntax predicting English reading beyond English language

The lower row of Table 4 provides the results from models examining Spanish morphosyntax as a predictor of English reading while accounting for children's scores on the English language composite. Multiple OLS regression suggested that both English language and Spanish morphosyntax were unique, significant predictors of English reading. A 1.0 standard deviation increase in English language corresponded with a 0.58 (95% CI [0.43, 0.74], $p < 0.001$) standard deviation increase in English reading, whereas a 1.0 increase in Spanish morphosyntax corresponded with a 0.17 (95% CI [0.02, 0.32], $p = 0.028$) standard deviation increase in English reading. The overall model yielded an adjusted R² of 37.9%, with Spanish morphosyntax contributing an additional 6.0% to the 31.9% variance predicted by English language alone.

Multiple quantile regression provided further evidence of English language as a significant predictor of English reading across the quantiles of reading. Estimates for this relation

were 0.51 (95% CI [0.42, 0.70], $p < 0.001$), 0.48 (95% CI [0.45, 0.67], $p < 0.001$), and 0.61 (95% CI [0.43, 0.84], $p < 0.001$) at the 0.25, 0.50, and 0.75 quantiles of English reading, respectively. Spanish morphosyntax, however, did not meet criteria as a statistically significant predictor of reading above and beyond English language at any of the quantiles examined. Estimates obtained were 0.12 (95% CI [-0.02, 0.22], $p = 0.195$), 0.06 (95% CI [-0.04, 0.25], $p = 0.492$), and 0.13 (95% CI [-0.08, 0.39], $p = 0.238$) at the 0.25, 0.50, and 0.75 quantiles of English reading, respectively.

Discussion

The purpose of this study was to examine the relation between L1 and L2 oral language proficiency and English reading achievement in young Spanish–English bilingual children. Oral language development is more complex for bilingual children than monolingual children, and because of the important role oral language plays in reading development, additional research is needed to determine how bilingual children’s developing proficiencies in both languages contribute to their reading achievement. In this study, we assessed two components of oral language (i.e., vocabulary knowledge and morphosyntactic ability) and used OLS and quantile regression to examine links between oral language and English reading achievement. We explored the relations within language, across languages, and assessed whether Spanish oral language proficiency is related to English reading achievement beyond what is accounted for by English oral language proficiency.

Within-language relations between oral language and reading

Overall, English vocabulary knowledge accounted for a significant amount of the variance in participants’ reading scores. This finding is consistent with prior work with Spanish–English bilingual learners (Grimm et al., 2018; Mancilla-Martinez & Lesaux, 2010). Additionally, there was a significant difference in the relation between English vocabulary and reading across quantiles of English reading, such that the correlation between English vocabulary knowledge and English reading achievement was larger at the high end of the distribution of reading skills—indicating that English vocabulary knowledge was more strongly associated with reading for children who are good readers (relative to other children in this sample). The difference was substantial, as the relation observed among children with low reading scores was moderate and the relation observed among those with higher reading scores was moderate to large.

Prior research indicates that vocabulary development is largely dependent on exposure to the language (Goodrich & Lonigan, 2018; Melby-Lervåg & Lervåg, 2011). Children who have strong English reading abilities are likely to have had more exposure to English compared to children who are weaker English readers; hence, English vocabulary was more highly correlated with English reading among stronger readers. Children who have weaker English reading skills may have had limited exposure to English. Limited English exposure would reduce their opportunity to strengthen their English oral language proficiency, including their English vocabulary development. However, it is also possible that children with weaker English reading skills have lower overall oral language ability across their languages. Given that more than one explanation for weaker reading is viable, it is reasonable that bilingual children’s English vocabulary may not contribute as substantially to predicting English reading for less-skilled readers.

It is also important to consider that children with weaker reading skills are still building proficiency in decoding. If children are still learning to decode, a larger, more complex vocabulary would not necessarily be beneficial to reading skill, as cognitive resources such as working memory may be almost exclusively dedicated to the task of decoding words sound by sound. In other words, inability to read words on the page limits reading outcomes, regardless of oral language ability. However, once children are proficient in decoding, then vocabulary knowledge becomes important because children no longer need to allocate cognitive resources to decoding but rather can use them to draw on oral language proficiency and better comprehend what they are reading. These results can be understood using the Simple View of Reading model (Gough and Tunmer, 1986; Hoover & Gough, 1990). According to this model, it is necessary for both decoding and linguistic comprehension to be present for reading comprehension to be possible, and as the proficiency in each of these factors increases, so does reading comprehension ability. Prior research on the Simple View of Reading indicates that oral language skills play a larger role in predicting reading achievement at later ages or for more advanced readers, when decoding skills have solidified and become relatively automatized (Lonigan et al., 2018). Our findings indicate that the relatively greater importance of oral language for more advanced readers applies to bilingual children learning to read in English and begins to emerge early in the process of acquiring reading skills.

Overall, English morphosyntactic skill accounted for a moderate amount of variance in English reading achievement. As with English vocabulary, we used quantile regression to explore whether the relation between English morphosyntactic awareness and English reading achievement differed depending on the level of English reading achievement. Based on the results of the analyses, each quantile was significantly correlated with morphosyntactic awareness, but there was no significant difference in these correlations across the three quantiles. That is, regardless of English reading achievement, morphosyntactic skill is an important component of reading achievement. One possible explanation for the discrepant findings between English vocabulary and morphosyntactic skill is that morphosyntax is important at all stages of reading development because it encompasses knowledge of grammatical rules of language and tends to develop in complexity concurrent to reading achievement. As children become better readers, they begin to read more difficult texts with more complex syntactic and morphological forms. Additionally, it is possible that morphosyntax is more indicative of bilingual students' underlying ability to acquire language, rather than strictly of language exposure. Thus, at least in the early stages of learning to read, morphosyntactic ability consistently predicts children's reading achievement across the distribution of reading skill.

Our results are complimentary to those of van den Bosch et al. (2020) who also examined factors that contribute to bilingual students' reading comprehension. These researchers assessed Dutch bilingual children's reading comprehension in addition to factors related to reading comprehension (e.g., vocabulary, morphosyntactic awareness, decoding). van den Bosch et al. reported that vocabulary and morphosyntactic knowledge were significantly related to reading comprehension. Similar to the present work, they also found that morphosyntax was consistently related to reading across the quantiles of reading achievement. However, they reported that vocabulary was only predictive of reading achievement at low quantiles of reading. This contrasts with our findings wherein vocabulary was strongly correlated with the highest quantile but not the middle and low quantile.

Several key differences between the studies may explain this difference in findings. First, in the current study, the majority of the participants were in kindergarten with a minority in first grade; however, the sample in van den Bosch et al. (2020) solely consisted of second

grade students. Children in kindergarten and first grade are typically at the start of transitioning from pre-literacy to literacy, while children in second grade are firmly in the literacy stage, assuming the children are typically developing (McConnell & Wackerle-Holman, 2016). As children get older, decoding skills become relatively automatized, which provides children the opportunity to focus on more complex skills such as comprehension. The disparity between participants' stages of literacy development likely contributed to the diverging results. Further, there were differences in the studies' analytic approaches. In the current study, decoding was used as an outcome variable, but van den Bosch and colleagues (2020) used it as a controlling factor in their analysis of vocabulary and reading comprehension, which may affect the distribution and the relation between vocabulary and different levels of reading achievement. Finally, L1 and L2 oral language distribution could be a contributing factor to differences in results, as van den Bosch and colleagues (2020) included bilingual students from multiple language backgrounds (e.g., some students spoke Dutch as L1 and Turkish as L2, and others spoke Turkish as L1 and Dutch as L2). For our study, all participating children were identified as having a home language of Spanish, learning English in school. Such differences in the nature of the bilingual sample may have important implications for how language is related to reading among bilingual children. Regardless of the differences in results across our study and those of van den Bosch et al., the current study and similar research make it clear that oral language skills such as vocabulary and morphosyntactic awareness are important elements in the development of reading achievement among bilingual children.

Cross-language relations between oral language and reading

Our primary hypothesis regarding the potential for Spanish oral language to predict English reading was that Spanish oral language skills would be most strongly related to English reading achievement at low quantiles of English reading; however, the pattern of results obtained did not consistently support this hypothesis. For young bilingual children, development of language and literacy skills is highly dependent on opportunity to acquire those skills. For example, recent research on identification of language and literacy disorders among bilingual children has highlighted that language of instruction is a "non-ignorable factor" in the identification of reading disability (Francis et al., 2019). We believed that lower reading achievement among kindergarten and first grade students in our sample may have reflected lack of opportunity to acquire English language and literacy skills. Lack of opportunity to acquire English could result in stronger influences of Spanish proficiency (and correspondingly weaker influences of English proficiency) on English reading achievement for children with poor English reading skills. For example, it is possible that some children in our sample attended English language preschool settings, whereas others did not. Although we do not have specific data on preschool attendance or language of preschool instruction, we thought that individual differences in kindergarten English reading achievement may reflect, to some degree, differences in opportunity to acquire reading-related skills in English. Only two of four cross-language analyses demonstrated that Spanish language skills predicted English reading achievement at low quantiles of English reading.

Another unexpected pattern that emerged was the finding that, when English language was accounted for, Spanish vocabulary contributed to predicting English reading achievement at the low and high end, but not in the middle, of the distribution of English reading. The magnitude of these contributions was small-to-moderate. It is possible that inclusion of measures of L1 vocabulary (in addition to L2 vocabulary) improves our measurement

of children's underlying language learning capacity. If so, this could explain why Spanish vocabulary knowledge is important for reading at the opposite ends of the continuum of reading achievement for different reasons. As we originally hypothesized, it is possible that Spanish vocabulary supports reading for children with poor English reading skills because it represents underlying language learning ability, and deficits in English reading are reflective of limited opportunity. Similarly, based on the simple view of reading, we would expect that underlying language skills are increasingly important at high levels of reading achievement, when decoding is more automatized (e.g., Lonigan et al., 2018). If Spanish vocabulary represents unique information about bilingual children's underlying language learning capacity that is not captured by measuring English vocabulary, it should uniquely predict English reading at the high end of decoding ability.

Our secondary hypothesis was that language-independent skills such as morphosyntactic abilities would demonstrate stronger cross-language relations with English reading achievement than would language-specific skills such as vocabulary knowledge. We formed this hypothesis based on Cummins' developmental interdependence theory (e.g., Cummins, 1981) and prior evidence that language-independent skills are more strongly correlated across languages (e.g., Goodrich & Lonigan, 2018; Melby-Lervåg & Lervåg, 2011). We considered morphosyntax to be more language independent, relative to vocabulary, because there is overlap in syntactic structures and morphological derivations and rules across Spanish and English, whereas labels for concepts in each language are largely arbitrary, with the exception of cognates. Results of OLS regression analysis did not support our hypothesis, as both Spanish vocabulary and morphosyntactic skill significantly contributed to English reading achievement, after controlling for English oral language skills. In the quantile regression framework, Spanish vocabulary emerged as a more robust unique predictor of English reading after controlling for English oral language. The unique associations between Spanish vocabulary and English reading were small-to-moderate at the lower and upper quantiles of reading. Conversely, quantile regression suggested at most a small unique association between Spanish morphosyntax and English reading across the quantiles, and these associations did not meet criteria for statistical significance.

Although this pattern of results was contradictory to our theoretically driven hypothesis, it is consistent with prior empirical findings that L1 vocabulary knowledge is a better indicator of L2 reading than is general oral language proficiency in L1 (e.g., Prevoe et al., 2016; Proctor et al., 2006). There are several potential explanations for these findings. First, these findings are reflective of the specific sample. This is particularly important to consider given the modest sample size, which provided limited power to detect small effects with confidence. Second, although vocabulary knowledge is highly sensitive to exposure and is language-specific, it is possible that the mechanisms underlying vocabulary acquisition and those underlying early literacy acquisition are similar or the same. Individual differences in children's ability to learn new words may closely align with how children learn foundational literacy skills (e.g., letter names and sounds, early decoding).

As described above, it is possible that differences in the extent to which measures of morphosyntactic skill and vocabulary knowledge index individual differences in underlying language learning capacities can explain the unexpected pattern of results. Many prior investigations of cross-language correlations attribute correlational evidence to "cross-linguistic transfer" (e.g., Dickinson et al., 2004; Melby-Lervåg & Lervåg, 2011). However, in the absence of experimental data, cross-language correlations are open to alternative interpretations, such as common linguistic environments across L1 and L2 or underlying language learning ability (Castilla et al., 2009). Our correlational analysis (see Table 2) indicates that there is greater overlap across languages in morphosyntax than there is for

vocabulary knowledge. Therefore, it is possible that Spanish and English morphosyntax potentially overlap in the specific variance in English reading comprehension that they explain. Measuring vocabulary knowledge in both languages might provide more *unique* information about children's underlying language learning capacity because vocabulary knowledge is highly dependent on language input in a specific language (e.g., Hoff et al., 2014). In contrast, measuring morphosyntax in both languages may introduce some measurement redundancy, as a common underlying proficiency about morphosyntax may be facilitated by input in either language (Cummins, 1981). Consequently, the unique predictive value of Spanish oral language (above and beyond the influence of English language) may be greater for vocabulary than it is for morphosyntax.

Implications for practice

Results of this study contribute valuable information on language and literacy development that have important implications for the assessment and instruction of reading among young Spanish–English bilingual children. First, using quantile regression to analyze the data allowed a more detailed look at how oral language is related to reading for children with more or less advanced reading skills. There was a stronger link between reading and vocabulary for children who were high performers in English reading than children who were low or average performers. This highlights the importance of considering individual differences among multilingual learners in decision-making for assessment and potentially for intervention. Different oral language skills may contribute more or less substantially to reading development at different stages of literacy acquisition. Consequently, this may lead to different assessment and/or intervention needs depending on children's level of reading development.

Second, evidence from the current study suggests that bilingual children's level of proficiency in the home language is important in predicting English reading achievement. Therefore, it is not sufficient to only assess bilingual children's English oral language skills to predict English reading achievement. To ensure that difficulties in the acquisition of oral language or literacy skills are due to a true disorder or disability, practitioners should evaluate children's oral language and literacy skills in both Spanish and English. Our results re-emphasize the importance of bilingual assessment, as we observed that measures of Spanish oral language do contribute to predicting English reading, similar to recent studies that have highlighted the importance of Spanish reading skills for predicting English reading (Relyea & Amendum, 2020).

Limitations and future research

Although this study represented a novel investigation into the cross-linguistic relations between oral language and reading achievement for young bilingual children, it had several limitations. First, the lack of inclusion of Spanish reading measures in this study represents a substantial limitation. Given potential for cross-linguistic relations in reading (especially in decoding skills for languages with overlapping alphabets like Spanish and English), future research is needed to explore the extent to which influences of Spanish oral language on English reading are mediated by Spanish reading. In fact, some recent evidence suggests that Spanish reading may more strongly predict English reading than does English oral language proficiency (Relyea & Amendum, 2020). Similarly, research in communication sciences and disorders has consistently highlighted that the only way

to accurately assess underlying language ability for bilingual children is to assess in both of their languages (e.g., Peña et al., 2016). It is possible that dual language assessment is equally important for examining bilingual children's reading abilities as well. Although this limitation is somewhat mitigated by the fact that children were enrolled in English language instruction (and likely received little formal Spanish reading instruction), future research should explore assessment of bilingual children's reading skills in all their languages, especially when exploring potential evaluation of reading disabilities.

We were unable to explore effects of SES on children's Spanish and English oral language development. While ample evidence suggests that oral language skills and academic achievement differ for children from different socioeconomic backgrounds (e.g., Fernald et al., 2013; Hoff, 2003; Reardon, 2011), our sample had minimal variability in SES, as all children attended Title I schools. Nevertheless, SES may be one factor that is contributing to oral language and reading development in our sample, as children had reading skills in the low-average range despite English and conceptual vocabulary knowledge in the average range. Future research should explore the specific impacts of SES on bilingual children's oral language and reading development in Spanish and English.

Although our assessment and conceptualization of oral language was broader than most prior research that has focused exclusively on vocabulary knowledge (e.g., Grimm et al., 2018) when predicting reading among bilingual children, our assessment of oral language ability is still somewhat limited, as we did not include a measure of more complex oral language skills, such as narrative language production, in our analysis. Although narrative ability represents an important aspect of oral language skill, little research exists on the dimensionality of narrative language sampling (i.e., which microstructural and macrostructural indices derived from language samples contribute the most to the measurement of narrative language ability?). Consequently, additional research is needed to evaluate which indices of narrative language would be most relevant or how depth of vocabulary knowledge contributes to predicting reading achievement among bilingual children.

Finally, we only examined reading achievement in kindergarten and first grade. This represents the very beginning stages of learning to read, and early differences in reading skill are likely to just represent individual differences in decoding. Relations between oral language and reading among bilingual children may shift for older children who have more automatized decoding skills (e.g., Tilstra et al., 2009). Given the relatively small sample limited to two US states, we had limited statistical power to evaluate the relations between language and reading skill across the distribution of reading achievement. Future research should explore these questions using larger samples that are more representative of bilingual children throughout the USA to allow for more precise quantification of the relations between language and reading measures across quantiles of reading.

Conclusions

Overall, our results indicate that children's home language skills make a significant contribution to English reading, above and beyond the influence of English oral language. Understanding children's oral proficiency in Spanish is likely to help better identify bilingual children who may be at risk for difficulty acquiring English reading skills. The results of our quantile regression analyses have important theoretical implications. Although not entirely consistent with our hypotheses, our results suggest that children's L1 vocabulary knowledge may provide more unique information about individual differences in

underlying language ability, whereas assessment of relatively language-independent skills in L1 may be somewhat redundant with measurement of corresponding skills in L2. Although language-independent skills may have a greater propensity for “cross-language transfer,” skills such as vocabulary knowledge may provide more utility in predicting future language and literacy development among bilingual children.

Author contribution J. Marc Goodrich and Lisa Fitton were responsible for the design and conduct of the research, data collection, and writing the report for publication. Lisa Fitton conducted all data analysis, and Lauren Thayer assisted with data collection and writing the report for publication.

Funding The research reported in this paper was supported by a New Investigators Research Grant awarded to Lisa Fitton and J. Marc Goodrich by the American Speech-Language-Hearing Foundation.

Data availability De-identified data used in the analyses will be provided upon request to the corresponding author.

Code availability Upon acceptance for publication, analytic syntax will be made available on Open Science Framework.

Declarations

Ethics approval This study was reviewed and approved by the Institutional Review Boards at the University of Nebraska-Lincoln and the University of South Carolina.

Consent to participate All participants provided informed consent prior to participation.

Consent for publication Each university’s respective guidelines from the Institutional Review Board were followed in informing participants that de-identified data would be published.

Conflict of interest The authors declare no competing interests.

References

- Anaya, J. B., Peña, E. D., & Bedore, L. M. (2018). Conceptual scoring and classification accuracy of vocabulary testing in bilingual children. *Language, Speech, and Hearing Services in Schools*, 49(1), 85–97. https://doi.org/10.1044/2017_LSHSS-16-0081
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (methodological)*, 57(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Boerma, T., & Blom, E. (2017). Assessment of bilingual children: What if testing both languages is not possible? *Journal of Communication Disorders*, 66, 65–76. <https://doi.org/10.1016/j.jcomdis.2017.04.001>
- Castilla, A. P., Restrepo, M. A., & Perez-Leroux, A. T. (2009). Individual differences and language interdependence: A study of sequential bilingual development in Spanish-English preschool children. *International Journal of Bilingual Education and Bilingualism*, 12(5), 565–580. <https://doi.org/10.1080/13670050802357795>
- Castilla-Earls, A., Bedore, L., Rojas, R., Fabiano-Smith, L., Pruitt-Lord, S., Restrepo, M. A., & Peña, E. (2020). Beyond scores: Using converging evidence to determine speech and language Services eligibility for dual language learners. *American Journal of Speech-Language Pathology*, 29(3), 1116–1132. https://doi.org/10.1044/2020_AJSLP-19-00179
- Comtois D (2021) *summarytools: Tools to quickly and neatly summarize data*. R package version 1.0.0. <https://cran.r-project.org/web/packages/summarytools/index.html>
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49(2), 222–251. <https://doi.org/10.3102/00346543049002222>

- Cummins J (1981) The role of primary language development in promoting educational success for language minority students. In Office of Bilingual Bicultural Education, California State Department of Education (Ed.), *Schooling and language minority students: A theoretical framework* (pp. 3–49). California State Department of Education, Sacramento, CA. <https://doi.org/10.13140/2.1.1334.9449>
- Dickinson, D. K., McCabe, A., Clark-Chiarelli, N., & Wolf, A. (2004). Cross-language transfer of phonological awareness in low-income Spanish and English bilingual preschool children. *Applied Psycholinguistics*, 25(3), 323–347. <https://doi.org/10.1017/S0142716404001158>
- Duke, N. K., & Cartwright, K. B. (2021). The science of reading progresses: Communicating advances beyond the simple view of reading. *Reading Research Quarterly*, 56(S1), S25–S44. <https://doi.org/10.1002/rrq.411>
- Ehl, B., Bruns, G., & Grosche, M. (2020). Differentiated bilingual vocabulary assessment reveals similarities and differences compared to monolinguals: Conceptual versus single-language scoring and the relation with home language and literacy activities. *International Journal of Bilingualism*, 24, 715–728.
- Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, 16, 234–248.
- Fitton, L., Hoge, R., Petscher, Y., & Wood, C. (2019). Psychometric evaluation of the bilingual English-Spanish assessment sentence repetition task for clinical decision making. *Journal of Speech, Language, and Hearing Research*, 62(6), 1906–1922. https://doi.org/10.1044/2019_JSLHR-L-18-0354
- Florit, E., & Cain, K. (2011). The simple view of reading: Is it valid for different types of alphabetic orthographies? *Educational Psychology Review*, 23(4), 553–576. <https://doi.org/10.1007/s10648-011-9175-6>
- Francis, D. J. (2019). Identification, classification, and treatment of reading and language disabilities in Spanish-speaking EL students: Introduction to the special issue. *New Directions for Child and Adolescent Development*, 2019(166), 7–14. <https://doi.org/10.1002/cad.20301>
- Francis, D. J., Rojas, R., Gusewski, S., Santi, K. L., Khalaf, S., Hiebert, L., & Bunta, F. (2019). Speaking and reading in two languages: On the identification of reading and language disabilities in Spanish-speaking English learners. *New Directions for Child and Adolescent Development*, 166, 5–41. <https://doi.org/10.1002/cad.20306>.
- Goodrich, J. M., & Lonigan, C. J. (2018). Development of first- and second-language vocabulary knowledge among language-minority children: Evidence from single language and conceptual scores. *Journal of Child Language*, 45(4), 1006–1017. <https://doi.org/10.1017/S0305000917000538>
- 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>
- Grimm, R. P., Solari, E. J., & Gerber, M. M. (2018). A longitudinal investigation of reading development from kindergarten to grade eight in a Spanish-speaking bilingual population. *Reading and Writing*, 31(3), 559–581. <https://doi.org/10.1007/s11145-017-9798-1>
- Gross, M., Buac, M., & Kaushanskaya, M. (2014). Conceptual scoring of receptive and expressive vocabulary measures in simultaneous and sequential bilingual children. *American Journal of Speech-Language Pathology*, 23(4), 574–586. https://doi.org/10.1044/2014_AJSLP-13-0026
- He, X., & Hu, F. (2002). Markov chain marginal bootstrap. *Journal of the American Statistical Association*, 97(459), 783–795. <https://doi.org/10.1198/016214502388618591>
- Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, 74, 1368–1378.
- Hoff E, Welsh S, Place S, & Ribot KM (2014) Properties of dual language input that shape bilingual development and properties of environments that shape dual language input. In J. Paradis & T. Grüter (Eds.), *Input and Experience in Bilingual Development* (pp. 119-140). Amsterdam: Benjamins.
- Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing*, 2(2), 127–160. <https://doi.org/10.1007/BF00401799>
- Kapantzoglou, M., Restrepo, M. A., & Thompson, M. S. (2012). Dynamic assessment of word learning skills: Identifying language impairment in bilingual children. *Language, Speech, and Hearing Services in Schools*, 43(1), 81–96. [https://doi.org/10.1044/0161-1461\(2011/10-0095\)](https://doi.org/10.1044/0161-1461(2011/10-0095))
- Karem, R. W., Washington, K. N., Crowe, K., Jenkins, A., Leon, M., Kokotek, L., et al. (2019). Current methods of evaluating the language abilities of multilingual preschoolers: A scoping review using the International Classification of Functioning, Disability and Health – Children and Youth Version. *Language, Speech, and Hearing Services in Schools*, 50(3), 434–451. https://doi.org/10.1044/2019_LSHSS-18-0128
- Kassambara A (2020) *ggpubr: “ggplot2” Based publication ready plots*. R package version 0.4.0. <https://cran.r-project.org/package=ggpubr>

- Kendeou, P., van den Broek, P., White, M. J., & Lynch, J. S. (2009). Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of Educational Psychology, 101*(4), 765–778. <https://doi.org/10.1037/a0015956>
- Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica, 46*(1), 33–50. <https://doi.org/10.2307/1913643>
- Koenker, R., & D'orey, V. (1994). Remark AS R92: A remark on algorithm AS 229: Computing dual regression quantiles and regression rank scores. *Journal of the Royal Statistical Society. Series C (Applied Statistics), 43*(2), 410–414. <https://doi.org/10.2307/2986030>
- Koenker, R., & Hallock, K. F. (2001). Quantile regression. *Journal of Economic Perspectives, 15*(4), 143–156. <https://doi.org/10.1257/jep.15.4.143>
- Koenker R (2021) *quantreg: Quantile regression*. R package version 5.86. <https://CRAN.R-project.org/package=quantreg>
- Language and Reading Research Consortium. (2015). The dimensionality of language ability in young children. *Child Development, 86*(6), 1948–1965. <https://doi.org/10.1111/cdev.12450>
- Language and Reading Research Consortium. (2015). The dimensionality of Spanish in young Spanish-English dual-language learners. *Journal of Speech, Language, and Hearing Research, 58*(3), 754–766. https://doi.org/10.1044/2015_JSLHR-L-13-0266
- Language and Reading Research Consortium, & Chiu, Y. D. (2018). The simple view of reading across development: Prediction of grade 3 reading comprehension from prekindergarten skills. *Remedial and Special Education, 39*(5), 289–303. <https://doi.org/10.1177/0741932518762055>
- Language and Reading Research Consortium, Jiang, H., & Logan, J. (2019). Improving reading comprehension in the primary grades: Mediated effects of a language-focused classroom intervention. *Journal of Speech, Language, and Hearing Research, 62*(8), 2812–2828. https://doi.org/10.1044/2019_JSLHR-L-19-0015
- Language and Reading Research Consortium, Mesa, C., & Yeomans-Maldonado, G. (2021). English and Spanish predictors of grade 3 reading comprehension in bilingual children. *Journal of Speech, Language, and Hearing Research, 64*(3), 889–908. https://doi.org/10.1044/2020_JSLHR-20-00379
- Lazewnik, R., Craghead, N. A., Smith, A. B., Prendeville, J.-A., Raiser-Becker, L., & Silbert, N. (2019). Identifiers of language impairment for Spanish-English dual language learners. *Language, Speech, and Hearing Services in Schools, 50*(1), 126–137. https://doi.org/10.1044/2018_LSHSS-17-0046
- Leider, C. M., Proctor, C. P., Silverman, R. D., & Haring, J. R. (2013). Examining the role of vocabulary depth, cross-linguistic transfer, and types of reading measures on the reading comprehension of Latino bilinguals in elementary school. *Reading & Writing, 26*, 1459–1485.
- Lepola, J., Lynch, J., Kiuru, N., Laakkonen, E., & Niemi, P. (2016). Early oral language comprehension, task orientation, and foundational reading skills as predictors of grade 3 reading comprehension. *Reading Research Quarterly, 51*(4), 373–390. <https://doi.org/10.1002/rq.145>
- Lonigan, C. J., Burgess, S. R., & Schatschneider, C. (2018). Examining the simple view of reading with elementary school children: Still simple after all these years. *Remedial and Special Education, 39*(5), 260–273. <https://doi.org/10.1177/0741932518764833>
- 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>
- Martin N (2013) Expressive one-word picture vocabulary test-4: Spanish-bilingual edition (EOWPVT-4:SBE). Novato, CA: Academic Therapy Publications.
- McConnell, S., & Wackerle-Hollman, A. (2016). Can we measure the transition to reading? General outcome measures and early literacy development from preschool to early elementary grades. *AERA Open, 2*(3), 1–15. <https://doi.org/10.1177/2332858416653756>
- Melby-Lervåg, M., & Lervåg, A. (2011). Cross-linguistic transfer of oral language, decoding, phonological awareness and reading comprehension: A meta-analysis of the correlational evidence. *Journal of Research in Reading, 34*(1), 114–135. <https://doi.org/10.1111/j.1467-9817.2010.01477.x>
- Paradis, J., Schneider, P., & Duncan, T. S. (2013). Discriminating children with language impairment among English-language learners from diverse first-language backgrounds. *Journal of Speech, Language, and Hearing Research, 56*(3), 971–981. [https://doi.org/10.1044/1092-4388\(2012\)12-0050](https://doi.org/10.1044/1092-4388(2012)12-0050)
- Peña, E. D., Bedore, L. M., & Kester, E. S. (2016). Assessment of language impairment in bilingual children using semantic tasks: Two languages classify better than one. *International Journal of Language & Communication Disorders, 51*(2), 192–202. <https://doi.org/10.1111/1460-6984.12199>
- Peña, E. D., Bedore, L. M., Lugo-Neris, M. J., & Albudoor, N. (2020). Identifying developmental language disorder in school age bilinguals: Semantics, grammar, and narratives. *Language Assessment Quarterly, 17*, 541–558. <https://doi.org/10.1080/15434303.2020.1827258>

- Peña E, Gutiérrez-Clellen VF, Iglesias A, Goldstein BA, & Bedore LM (2014) *Bilingual English-Spanish Assessment (BESA)*. San Rafael, CA: AR Clinical Publications.
- Pratt, A. S., Peña, E. D., & Bedore, L. M. (2021). Sentence repetition with bilinguals with and without DLD: Differential effects of memory, vocabulary, and exposure. *Bilingualism: Language and Cognition*, 24(2), 305–318. <https://doi.org/10.1017/S1366728920000498>
- Prevo, M. J. L., Malda, M., Mesman, J., & van Ijzendoorn, M. H. (2016). Within- and cross-language relations between oral language proficiency and school outcomes in bilingual children with an immigrant background: A meta-analytical study. *Review of Educational Research*, 86, 237–276.
- Proctor, C. P., August, D., Carlo, M. S., & Snow, C. (2006). The intriguing role of Spanish language vocabulary knowledge in predicting English reading comprehension. *Journal of Educational Psychology*, 98, 159–169.
- R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.r-project.org/>
- Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In J. Duncan & R. J. Murnane (Eds.), *Whither opportunity? Rising inequality, schools, and children's life chances* (pp. 91–116). Russell Sage Foundation.
- Relyea, J. E., & Amendum, S. J. (2020). English reading growth in Spanish-speaking bilingual students: Moderating effect of English proficiency on cross-linguistic influence. *Child Development*, 91, 1150–1165.
- Revelle W (2021). psych: Procedures for psychological, psychometric, and personality research. R package version 2.1.6. <https://cran.r-project.org/web/packages/psych/index.html>
- Rojas, R., Hiebert, L., Gusewski, S., & Francis, D. J. (2019). Moving forward by looking back: Understanding why some Spanish-speaking English learners fall behind. *New Directions for Child and Adolescent Development*, 2019(166), 43–77. <https://doi.org/10.1002/cad.20305>
- Snowling, M. J., Hayiou-Thomas, M. E., Nash, H. M., & Hulme, C. (2020). Dyslexia and developmental language disorder: Comorbid disorders with distinct effects on reading comprehension. *Journal of Child Psychology and Psychiatry*, 61(6), 672–680. <https://doi.org/10.1111/jcpp.13140>
- Swanson, H. L., Rosston, K., Gerber, M., & Solari, E. (2008). Influence of oral language and phonological awareness on children's bilingual reading. *Journal of School Psychology*, 46, 413–429.
- Tilstra, J., McMaster, K., den Broek, P. V., Kendeou, P., & Rapp, D. (2009). Simple but complex: Components of the simple view of reading across grade levels. *Journal of Research in Reading*, 32(4), 383–401. <https://doi.org/10.1111/j.1467-9817.2009.01401.x>
- van den Bosch, L. J., van den Segers, E., & Verhoeven, L. (2020). First and second language vocabulary affect early second language reading comprehension development. *Journal of Research in Reading*, 43(3), 290–308. <https://doi.org/10.1111/1467-9817.12304>
- Wickham H (2016). ggplot2: Elegant graphics for data analysis (2nd ed.). Switzerland: Springer. https://doi.org/10.1007/978-3-319-24277-4_12
- Woodcock RW, Alvarado CG, Ruef ML, & Schrank FA (2017) *Woodcock-Muñoz Language Survey III*. Itasca, IL: Houghton Mifflin Harcourt

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.