



# Text Complexity Modulates Cross-Linguistic Sentence Integration in L2 Reading

Sibylla Leon Guerrero<sup>1\*</sup>, Veronica Whitford<sup>2</sup>, Laura Mesite<sup>1</sup> and Gigi Luk<sup>3</sup>

<sup>1</sup>Graduate School of Education, Harvard University, Cambridge, MA, United States, <sup>2</sup>Dept. of Psychology, University of New Brunswick, Fredericton, NB, Canada, <sup>3</sup>Dept. of Educational and Counselling Psychology, McGill University, Montreal, QC, Canada

Cross-linguistic influences (CLI) in first-language (L1) and second-language (L2) reading have been widely demonstrated in experimental paradigms with adults at the word and sentence levels. However, less is known about CLI in adolescents during naturalistic text reading. Through eye-tracking and behavioral measures, this study investigated expository reading in functionally English monolingual and Spanish (L1) - English (L2) bilingual adolescents. In particular, we examined the role of L1 (Spanish) sentence integration skills among the bilingual adolescents when L2 school texts contained challenging syntactic structures, such as complex clauses, elaborated noun phrases, and anaphoric references. Results of generalized multilevel linear regression modeling demonstrated CLI in both offline comprehension and online eye-tracking measures that were modulated by school text characteristics. We found a positive relationship (i.e., facilitation) between L1 sentence integration skills and L2 English text comprehension, especially for passages with greater clause complexity. Similar main, but not modulatory, effects of sentence integration skill were found in online eye-tracking measures. Overall, both language groups appeared to draw upon similar reading component skills to support reading fluency and comprehension when component skills were measured only in English. However, differential patterns of association across languages became evident when those skills were measured in both L1 and L2. Taken together, our findings suggest that bilingual adolescents' engagement of cross-linguistic resources in expository reading varies dynamically according to both language-specific semantic knowledge and language-general sentence integration skills, and is modulated by text features, such as syntactic complexity.

**Keywords:** bilingual (Spanish/English), cross-linguistic influence, reading comprehension, syntactic integration, second language reading, text complexity, eye-tracking

## INTRODUCTION

Reading and comprehending complex connected school texts can be challenging for both first-language (L1) (Cain and Oakhill, 2004; Lervåg et al., 2018) and second-language (L2) (Lesaux, et al., 2006; Melby-Lervåg and Lervåg, 2014) learners. Readers must draw upon linguistic knowledge and skills of different types: orthographic decoding, vocabulary, syntax, and discourse knowledge to rapidly create and update dynamic representations of meaning (Kintsch and Van Dijk, 1978; Perfetti

## OPEN ACCESS

### Edited by:

Anat Prior,  
University of Haifa, Israel

### Reviewed by:

Jorge R. Valdes Kroff,  
University of Florida, United States  
Brian MacWhinney,  
Carnegie Mellon University,  
United States

### \*Correspondence:

Sibylla Leon Guerrero  
sibylla\_leon\_guerrero@  
mail.harvard.edu

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Communication

**Received:** 11 January 2021

**Accepted:** 31 May 2021

**Published:** 18 June 2021

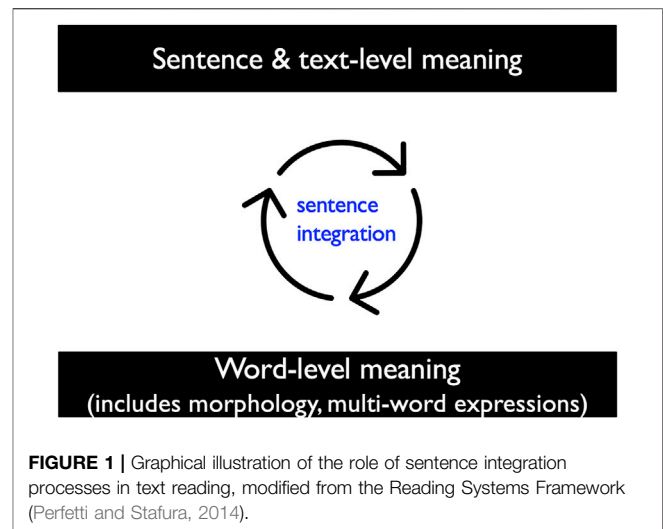
### Citation:

Leon Guerrero S, Whitford V, Mesite L  
and Luk G (2021) Text Complexity  
Modulates Cross-Linguistic Sentence  
Integration in L2 Reading.  
Front. Commun. 6:651769.  
doi: 10.3389/fcomm.2021.651769

and Stafura, 2014). For bilinguals who are also biliterate, the reading process is further complicated by the presence of two languages, drawing upon component skills that can include more than one orthography, lexicon, and syntactic system. Cross-linguistic influence (CLI) describes the effects that bilinguals' languages may have on each other, even when only one language is the target of communication (Alonso, 2019). Cummins' (1979) early Linguistic Interdependence Hypothesis proposed a common underlying linguistic proficiency that would allow bilinguals to transfer linguistic skills across languages given sufficient L1 proficiency. This hypothesis proposes that in addition to their emerging L2 abilities, young bilinguals draw upon L1 knowledge and skills in developing L2 literacy.

Educational studies of cross-linguistic influence using behavioral measures of oral language broadly defined (e.g., Lesaux et al., 2010; Nakamoto et al., 2012; Proctor et al., 2010; Relyea and Amendum, 2020) have generally found a weak or no contribution of L1 oral language to skilled text reading and comprehension in L2, or have found that L1 contributions are mediated by L2 language skills. In contrast, extensive behavioral research has found that L1 can support L2 text reading through shared component processes, such as phonological awareness (Durgugoglu et al., 1993; Bialystok et al., 2005; Prevoo et al., 2016) and orthographic decoding (Geva and Siegel, 2000; Lindsey et al., 2003; Lesaux and Siegel, 2003; Verhoeven and van Leeuwe, 2009; Melby-Lervåg and Lervåg, 2011; Kremin et al., 2019), particularly when L1 and L2 share writing systems. Similarly, cross-linguistic transfer of L1 vocabulary knowledge, especially as children move from the stages of "learning to read" to "reading to learn," can bolster L2 reading comprehension (e.g., Proctor et al., 2005; Pasquarella et al., 2012; van den Bosch et al., 2020). However, in social contexts such as the United States, where there is often attrition of the minority home language as children's schooling progresses in the societally dominant language (i.e., English), L1 proficiency can also exert a negative influence on L2 comprehension (e.g., Swanson et al., 2008; Kieffer, 2012; Ordóñez et al., 2002). In these contexts of subtractive bilingualism (Cummins, 2000), conditions which promote L2 proficiency may also exacerbate L1 attrition, giving the impression of negative interference as higher levels of L1 are correlated with lower levels of L2 language knowledge and vice versa.

Although CLI in reading has been widely examined at the level of sound and word representations (see Genesee and Geva, 2006; Chung et al., 2019), there has been comparatively little education research on CLI involving the integration of these word representations into sentence- and text-level meanings, a higher-order process needed for skilled reading and comprehension of school texts. While rich lexical representations are important building blocks in reading, comprehension requires more than understanding words in isolation. Readers must create a context or situation model of connected concepts and rapidly integrate new words as they are read (Van Dijk and Kintsch, 1983). The Reading Systems Framework (RSF, Perfetti and Stafura, 2014) is one account of this memory-dependent integration process, which postulates that successful readers rapidly access orthographic and rich



lexical representations from written text. These rich lexical representations include morphological and other word features, such as aspect and category, that contribute to "high-quality" lexical representations. In order to comprehend a text, readers hold these representations in memory while integrating them into a holistic representation of the sentence (or sentences), called the textbase. Sentence-level integration requires the reader to both compose and decompose meaning beyond the lexical level (e.g., using word order, referential rules, and other sentence-level cues). The present study investigates this integrative processing (illustrated in **Figure 1**) as a potential locus of L1 linguistic ability that may support L2 reading efficiency and comprehension.

Here, we investigated the nature of word-to-sentence integration using naturalistic reading tasks drawn from education curricula in order to investigate the complex reading behaviors of adolescents who must learn in a developing L2. According to the RSF, linguistic knowledge contributes to the construction of meaning beyond the word level; however, the mechanism by which word-level representations are integrated at the sentence and text level is not precisely specified in the RSF. There are currently two prevalent types of models of sentence integration: memory-based integration models and predictive processing models. On the one hand, memory-based accounts of complex meaning representation have traditionally been formulated around the "bottom-up" processing of separate representations of the lexicon, morphology, and syntax that are held together in working memory to create sentence-level meaning (e.g., Cunnings, 2017). On the other hand, computational and neurobiological accounts of complex meaning representation focus on the role of "top-down" predictive or expectation-based meaning construction (e.g., Levy, 2008). Tightly controlled experimental work can provide rigorous evidence validating or disproving these types of models. However, such work also trades off experimental control and ecological validity and insight into complex behavior in rich discourse contexts (e.g., see Vanderwal et al., 2019 for a discussion of naturalistic stimuli). In order to conduct transdisciplinary educational research, the processing account in the present

naturalistic study is grounded in an ecologically relevant approach in which we adopt a top-down perspective whereby different reading component skills are revealed in behavioral and eye movement measures of reading efficiency and comprehension.

Findings are thus observational and seek to extend prior monolingual literature indicating that same-language sentence integration skills are important contributors to reading comprehension, particularly starting in late elementary and middle school (Low and Siegel, 2005; Nation and Snowling, 2000; Lesaux et al., 2006; Geva and Farnia, 2012; Proctor et al., 2012; Jeon and Yamashita, 2014; Gottardo et al., 2018; Babayigit and Shapiro, 2019; Brimo et al., 2017). In particular, little is understood about the cross-linguistic contributions of L1 sentence integration abilities when bilingual individuals are reading in their L2, a common requirement for bilingual learners with L2 English in schools with English as the medium of instruction. The current study focuses on CLI in sentence-level integration, examining how L1 (Spanish) sentence integration abilities during L2 (English) text reading in Spanish-English bilingual adolescents are associated with both online text processing efficiency (assessed by eye-tracking measures) and offline text comprehension (assessed by post-reading comprehension measures) with their functional monolingual peers who spoke and read primarily only in English.

## Educational Studies of Cross-Linguistic Influence in Text Comprehension

Although cross-linguistic influence involving the lexicon has been widely investigated, there is inconclusive evidence regarding potential CLI of syntax on L2 reading comprehension. Several behavioral studies involving children educated in L2 English-dominant environments found no relationship between L1 syntactic skills and L2 reading comprehension. In a study of adolescent Spanish-English bilinguals who were newcomers to the United States with prior schooling in Spanish, Garrison-Fletcher et al. (2019) found that general and academic reading comprehension, but not Spanish syntactic skills, were positive predictors of English text comprehension. Similarly, Swanson et al. (2008) assessed 68 third grade Spanish-English bilinguals with both Spanish and English reading component measures. Although L1 Spanish morphosyntactic knowledge positively predicted English reading comprehension, it did not do so uniquely—this relationship was explained by other English measures. With 123 children in grades 3–5, Leider et al. (2008) found a negative relationship between Spanish syntax and English sentence judgment performance but no relationship with English comprehension measured by cloze or multiple-choice tasks. Kieffer (2012) found a similar negative relationship between 295 kindergarten children's Spanish oral language skills more broadly measured and later English reading comprehension in nationally representative longitudinal data. These three studies were conducted in the United States, where the L2 (English) is the dominant language used almost exclusively in public schools. Only one study with participants in this setting produced a contrasting result. In a longitudinal

investigation of syntactic skills among 156 Spanish-English bilingual upper elementary students, Proctor et al. (2017) found that L1 syntax, measured with a sentence formulation task in the second grade, predicted English oral language and reading comprehension in the fifth grade. There is thus mixed evidence of cross-linguistic syntactic transfer in L2 English-dominant settings at different ages, with study findings variously suggesting interference, positive transfer, or no influence of L1 syntactic skills on L2 text comprehension.

In contrast, behavioral studies conducted in immersive bilingual education settings where bilingual literacy is explicitly instructed have found a consistent positive relationship between L1 syntax and L2 reading comprehension during the preschool to upper elementary years. Gabriele et al. (2009) examined a small sample of Spanish-English bilinguals in a bilingual education preschool program. Preschoolers' performance on native Spanish structures (i.e., more complex structures) predicted better performance on an English reading readiness measure. In unpublished studies, Sohail (2015) tested a cross-sectional sample of emergent Canadian English-French bilingual first graders and their third-grade peers who had completed 2 years of French immersion using a word order correction task. In the first grade, but not in the third grade, L1 English word order correction was positively associated with L2 French reading comprehension. Among bilingual Cantonese-English first through third graders studying a dual-language curriculum in Hong Kong, both cross-sectional (Siu and Ho, 2015) and longitudinal (Siu and Ho, 2020) studies found that L1 Chinese word order and morphosyntactic skills positively predicted L2 English text comprehension measured in the third grade. L2 English word order and morphosyntactic skills mediated both cross-sectional and longitudinal relationships when accounting for oral language and cognitive skills, suggesting that syntactic skills represent a common underlying linguistic proficiency contributing to reading comprehension. Similarly, for Spanish-English bilingual upper elementary students in dual immersion programs, Phillips Galloway et al. (2020) found that L2 (English) reading comprehension was positively associated with a composite measure of academic language in the L1 Spanish that included syntactic skills in addition to vocabulary and genre-related knowledge. Although each of these studies investigated different age groups (preschool to upper elementary) and different measures of syntactic skill (word order, morphosyntax, and sentence complexity), they share common findings of cross-linguistic syntactic transfer to L2 reading among children in bilingual immersion education settings. Consistent with Cummins' (1979) hypothesis that readers could draw upon a common underlying linguistic proficiency to support L2 comprehension only when they possessed adequate L1 proficiency, these findings in bilingual education settings suggest that school instruction in both L1 and L2 may support the positive influence of L1 skills on L2 reading, or at least may mitigate variability in L1 proficiency that impedes this cross-linguistic influence.

Particularly in the United States' educational context where English-only schooling has predominated in recent decades, these mixed results from studies investigating CLI may arise from

heterogeneity in either or both participant language characteristics and educational environments. For example, United States-centric studies of Spanish-English bilinguals often examine students designated in the public school system as English learners, who are predominantly United States born, but may in fact range from recent immigrants from Spanish-speaking countries with extensive Spanish-language educational experience to second generation heritage speakers with minimal Spanish proficiency (e.g. see Luk and Christodoulou, 2016 for a discussion). The present study, conducted in the United States context, seeks to understand how students for whom Spanish is the first spoken language (L1) may draw upon this L1 to support text processing and comprehension in English, their second spoken language (L2). We considered Spanish as the participant's L1 if it was the earliest language *spoken* by the participant on a daily basis at home, even for participants who may have been *exposed* to English at an early age by listening to media, community interactions, or family conversation in the context of the United States where English is the societally dominant language.

Variability found in CLI effects both within L2 English-dominant school environments and also across L2-dominant compared to immersive bilingual education settings additionally suggest that CLI is more likely to be observed in contexts where both L1 and L2 literacy are systematically taught and that skills with L1 oral vs. written syntax may have different influences on L2 written comprehension. L1 syntactic skills may thus support L2 reading comprehension; however, it is still unclear for whom and under what behavioral measures and environmental conditions this cross-linguistic support may take place.

## Eye-Tracking Studies of Cross-Linguistic Influence in Text Processing

Unlike traditional behavioral tasks, eye-tracking allows for a direct, naturalistic, and temporally sensitive measure of the cognitive processes underlying reading behaviors as they unfold (Rayner, 2009). A large body of eye-tracking literature on bilingual sentence processing has found that L1 syntactic knowledge impacts online L2 sentence reading most often by slowing text processing, for example in complex syntactic structures, such as causal connectives (van den Bosch et al., 2018), anaphora (see Godfroid, 2019 for a review), and referential clauses (see Dussias et al., 2010; Rossi et al., 2019 for a review). Eye-tracking studies of connected text reading (though not focused on CLI) have also found differences in L1 and L2 reading behavior, including longer fixations and more saccades in the L2, contributing to longer reading times. These differences can vary according to levels of self-reported L1 and L2 exposure (Whitford and Titone, 2012; Whitford and Titone, 2017) and objective L1 and L2 proficiency (Cop et al., 2015; Whitford and Joannis, 2018).

In addition to reader characteristics, effects of syntactic complexity of the text have been found in bilingual adults' adaptation to word category combinations (e.g., article-adjective-noun). Snell and Theeuwes (2020) presented a naturalistic narrative text in both Dutch and English and found that higher frequency structures facilitated Dutch-

English bilinguals' eye movement reading behavior in both languages. Differences for bilingual adults reading narrative passages in L1 vs. L2 have been found for syntactic structures such as gerunds, participial phrases, and relative clauses (De Groot, 2018). Among monolingual children, word position during expository text reading has been found to impact offline comprehension (de Leeuw et al., 2016). Furthermore, syntactic processing difficulties have been reported among monolingual children for anaphoric structures in short narrative passages (Joseph et al., 2015).

Eye-tracking studies of connected text reading have primarily focused on adults, leaving reading behavior in children relatively less understood. However, prior studies involving both monolinguals (Joseph et al., 2015; Reichle et al., 2013; Whitford and Joannis, 2018) and bilinguals (Whitford and Joannis, 2018) have reported age differences in reading performance, with children exhibiting more fixations and regressions, longer fixations, and shorter saccade amplitudes than adults (Whitford and Joannis, 2018). Although the maturation of oculomotor control appears to be largely complete by puberty, or around 12 years of age, children's eye movement reading behavior only approximates that of adults as their language proficiency and word-reading automaticity develops (Blythe and Joseph, 2011; Reichle et al., 2013). Adolescent middle schoolers in K-12 education contexts thus are likely to have adult-like oculomotor capacities as well as word-level decoding skills but are still developing in the language and higher-level complex reading skills needed for efficient text processing.

To our knowledge, studies have yet to investigate the interaction of reader characteristics, syntactic characteristics of text, and reading comprehension using both offline comprehension tasks and eye movement measures of reading in L2 adolescent readers. Kuperman et al. (2018) examined a broad range of reader and text characteristics in native English monolingual university students, and found at the passage level that syntactic complexity of passages predicted several measures of online eye movement behavior, as well as offline reading comprehension. Results for reader characteristics varied depending upon the specific eye movement measure examined, with word reading fluency and nonverbal reasoning emerging as the common predictors of online text processing. Overall in this study, there was little modulation of reader characteristics by syntactic complexity of the text when comparing less complex to more complex passages. Although prior eye-tracking studies have found that syntactic complexity of texts influences eye movement reading behavior in monolingual children (e.g., measured as sentence length in German-speaking children: Tiffin-Richards and Schroeder, 2018) and adults (e.g., English-speaking adults: Kuperman et al., 2018), none have focused on word-to-sentence integration in bilingual adolescents who are experiencing the transitional demand of learning complex academic knowledge through a developing second language.

## The Current Study

Prior CLI research, which has primarily focused on bilingual adults (and thus, lack generalizability to other populations) and



employed highly controlled sentences as stimuli (and thus, may lack ecological validity), has found that both complex syntactic features of the text and L1 proficiency influence L2 text processing and comprehension. Syntactic features prevalent in school texts, such as elaborated noun phrases, lengthy relative clauses, conceptual anaphora, and the use of distinctive connectives, have been found to pose particular challenges for written text comprehension (Scott and Balthazar, 2010; Uccelli et al., 2015a) in students generally and in L2 learners more specifically (Phillips Galloway and Uccelli, 2019). Cross-linguistically, L2 English academic skills and reading comprehension have been found to be positively associated with L1 Spanish academic vocabulary (Lubliner and Hiebert, 2011) and with academic language skills broadly measured as combined lexical, syntactic, and discourse skills (Aguilar et al., 2020; Phillips Galloway and Uccelli, 2019). However, how L2 learners process distinctive syntactic features of written school texts is not well understood, and little prior research has examined L2 syntactic processing beyond morphosyntax in written academic discourse comprehension. Concurrently, prior education research involving bilingual adolescents has found that complex syntactic text features impede L2 reading comprehension, as a function of L2 proficiency. However, there is still little work connecting these two bodies of literature and examining how syntactic characteristics of L2 texts modulate the cross-linguistic relationship of individual reader characteristics and the comprehension of naturalistic school-based texts. To address this critical gap, the current study asks: How do complex syntactic text features and individual differences in cross-linguistic sentence integration skills, as well as their interactions, affect 1) online processing and 2) offline comprehension of naturalistic school-based texts in bilingual adolescent readers?

We examined individual differences in sentence integration skills in both the L1 (Spanish) and L2 (English), as well as in word decoding and vocabulary knowledge in both languages in order to control for lexical effects. Based on prior research, we expected that better L1 and L2 sentence integration skills would be associated with higher levels of L2 text comprehension (Lesaux et al., 2006) and with more efficient online L2 reading (Cop et al., 2015; Gollan et al., 2008; Whitford and Titone, 2012; Whitford and Titone, 2015; Whitford and Titone, 2017). We included two adolescent participant groups: functionally monolingual native English speakers, and Spanish (L1) - English (L2) speakers, acquiring English in school.

Three categories of text features were included in the study: phrase complexity, clause complexity, and the degree of anaphoric reference. Based on the RSF model (Perfetti and Stafura, 2014), we expected that these features, which present challenges to sentence integration processes, such as those involved in parsing, sequencing, and combinatorial analysis/synthesis, would be a potential locus of cross-linguistic influence in which individual differences in L1 sentence integration skills might contribute to L2 text comprehension. Because syntactic complexity presents challenges to sentence integration, readers may draw upon integrative resources such as syntactic representations or processing biases developed in the

L1 to cope with these challenges. As previous research has found that complex syntactic text features are associated with reading comprehension difficulties in L2 child and adolescent readers (Uccelli et al., 2015a; Uccelli et al., 2015b; Phillips Galloway and Uccelli, 2019), we expected that phrase complexity, clause complexity, and anaphoric references would all negatively modulate the expected positive effect of cross-linguistic sentence integration skills on offline text comprehension. Based on similar findings in a self-paced L2 reading study in Spanish-English bilingual adolescents which found that words in more syntactically complex passages were read more slowly than words in simpler ones (Kim et al., 2018; Mulder et al., 2020), we expected a negative impact of these complex text features on online text processing. The study's hypotheses were that:

- 1) greater L1 proficiency in sentence integration would be associated with higher levels of L2 text reading efficiency and comprehension, and
- 2) syntactic complexity of the text would negatively modulate the expected positive association of L1 syntactic integration skills with L2 text reading efficiency and comprehension.

## MATERIALS AND METHODS

### Participants

Sixty-five typically developing adolescents, aged 11–15 years, participated in the current study. Participants resided in the Boston Metropolitan Area and attended English-instruction schools. In accordance with Institutional Review Board guidelines (Harvard University IRB16-0866), both child assent and parental consent were obtained. Participants were compensated \$50 for their time. The final sample included 59 participants; three were excluded due to eye-tracking issues and another three were excluded due to low nonverbal reasoning scores (standard scores <86.6 or > two standard deviations below the sample mean on the Kaufman Brief Intelligence Test-2; KBIT-2 Matrices; Kaufman and Kaufman, 2004), which can influence reading comprehension outcomes (Quinn and Wagner, 2018).

A parental version of the Language and Social Background Questionnaire (LSBQ; adapted from Luk and Bialystok, 2013; Anderson et al., 2018) assessed participants' demographic background, language history, and familial language use. All participants demonstrated heterogeneous language backgrounds reflective of minority language speakers found in United States classrooms. Twenty-nine participants were native L1 (English) speakers, with no L2 immersion experience and minimal proficiency in an additional language. Most were enrolled in an introductory foreign language class as part of the standard middle school curriculum, and several of these L1 English speakers had been enrolled in some form of beginning language class sporadically since an early age. Thus, they were all functionally monolingual.

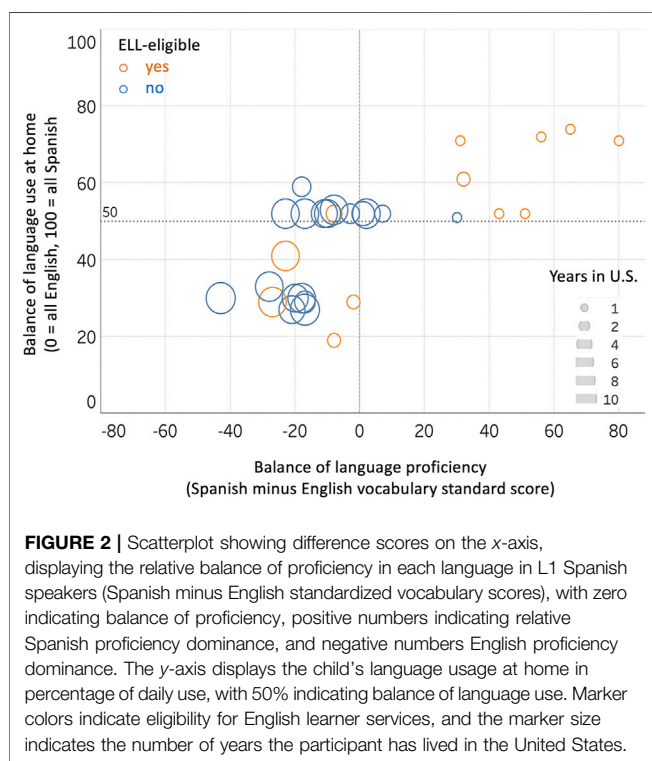
We considered Spanish as the participant's L1 if Spanish was the earliest language *spoken* on a daily basis at home, even for participants who may have been *exposed* to English at an early age by listening to media, community interactions, or family

**TABLE 1 |** Participant demographic characteristics.

	L1 English (n = 29) m (sd)	L1 Spanish (n = 30) m (sd)	Group difference (p of Kruskal-Wallis $\chi^2$ )
<i>Demographics</i>			
Gender (female/male)	20/9	17/13	n/a ( $p = 0.42$ )
Maternal education (median) (Categories: Elementary school; high school; 2-years (AA) degree; BA/BS; graduate degree)	Graduate degree	Bachelor's degree	n/a ( $p = 0.01$ )
Age (years)	12.6 (1.1)	13.1 (1.0)	0.54 (0.06)
<i>Language background</i>			
Age of english exposure in years (0 = "at birth")	0 (0)	7.8 (3.7)	7.8 (<0.001)
Non-english language <sup>b</sup> spoken daily (% time at home)	1.2 (2.7)	46.9 (15.3)	45.7 (<0.001)
Non-english language <sup>b</sup> heard daily (% time at home)	1.3 (2.7)	47.3 (14.5)	46.0 (<0.001)
Non-english language <sup>b</sup> read daily (% time at home)	1.2 (2.6)	33.8 (17.5)	32.6 (<0.001)

<sup>a</sup>Fisher's exact.

<sup>b</sup>For L1 English speakers, this variable measured any non-English language use at home. For L1 Spanish speakers, this variable measured L1 Spanish language use at home.



conversation. As illustrated in **Table 1**, thirty participants were identified as L1 Spanish speakers, who spoke Spanish from birth at home and used it 47% of the time, on average. Twenty-two of these participants were from Spanish-speaking countries (Mexico = 8, Central America = 5, South America = 7, Spain = 2) and eight had only lived in the United States. The age at which participants were first exposed to English on a regular basis was correspondingly heterogeneous, ranging from birth to 14 years of age ( $m = 7.8$ ,  $sd = 3.7$ ); however, all participants spoke Spanish at home and did not start to use English regularly until at least school age, or approximately 5–6 years of age.

As illustrated in **Figure 2** showing on the x-axis the difference scores between Spanish and English vocabulary measures as a

proxy for language proficiency dominance, the sample displayed a continuous range of variation on multiple and intersecting, but not fully overlapping, dimensions of language experience. For example, one participant who had lived in the United States for 10 years had markedly strong English dominance with a difference between English and Spanish vocabulary scores of over 40 points. This student could thus be considered a heritage speaker with attrition in home language but was nonetheless reported by parents to speak and read Spanish at home approximately 30% of the day and was similar in this regard to other students with English learner designations and with fewer years of residence in the United States. The multiple and at times disparate dimensions of language experience displayed in **Figure 2**—relative language proficiency, relative language usage, and relative community immersion—are reflective of the ecological variation in minority language speaker backgrounds found in United States classrooms. Also typical of United States classrooms, most of our minority language speaking sample had been initially identified as English learners upon school entry and had exited from this designation at various points in time prior to their study participation. Because initial English learner classification in United States schools is often based upon speaking a home language other than English, this designation encompasses a heterogeneous range of language proficiencies and language dominance as represented in our sample. **Figure 2** illustrates this wide distribution of students eligible for English learner services along all the other dimensions of language use, including language dominance, daily Spanish spoken, and years in the United States.

The two language groups were matched on gender (Fisher's exact  $p = 0.42$ ), age (Kruskal-Wallis  $\chi^2 p = 0.06$ ), nonverbal IQ (KBIT Matrices, Kaufman and Kaufman, 2004, Kruskal-Wallis  $\chi^2 p = 0.07$ ), verbal working memory (Digits Backward, Reynolds and Voress, 2007, Kruskal-Wallis  $\chi^2 p = 0.05$ ), and rapid naming (Letter, Number, 2-set, and 3-set subtests; Wolf and Denckla, 2005; Kruskal-Wallis  $\chi^2 p = 0.16$ ). These tasks were administered in the participant's preferred language. The two language groups were also matched on timed English word reading (Sight Word Efficiency, Torgesen et al., 2011, Kruskal-Wallis  $\chi^2 p = 0.96$ ) and untimed English word reading (Word Identification, Woodcock-Muñoz Language Survey with Normative Update, WMLS-R-NU,

**TABLE 2 |** Reading and cognitive characteristics.

	L1 English (n = 29) m (sd)	L1 Spanish (n = 30) m (sd)	Group difference (p of Kruskal-Wallis $\chi^2$ )
<b>Administered in preferred language (Spanish or English):</b>			
KBIT-2 nonverbal reasoning <sup>a</sup>	114.6 (10.7)	109.7 (10.1)	5.0 (0.07)
TOMAL digits backward <sup>b</sup>	10.9 (3.4)	9.5 (2.4)	1.5 (0.05)
RAN/RAS (subtest average) <sup>a</sup>	107.8 (10.1)	111.1 (10.2)	-3.3 (0.16)
TOWRE-2 sight word fluency <sup>a</sup>	108.6 (14.9)	107.6 (15.6)	1.0 (0.96)
<b>Administered in English:</b>			
WMLS-R English picture vocabulary <sup>a</sup>	113.4 (10.2)	84.6 (23.4)	28.8 (<0.001)
WMLS-R English letter-word ID <sup>a</sup>	114.5 (12.1)	111.9 (11.1)	2.7 (0.46)
WMLS-R English cloze <sup>a</sup>	105.5 (14.6)	87.1 (22.3)	18.4 (0.004)
<b>Administered in Spanish:</b>			
WMLS-R Spanish vocabulary <sup>a</sup>	—	87.1 (12.1)	—
WMLS-R Spanish letter-word ID <sup>a</sup>	—	127.9 (17.6)	—
WMLS-R Spanish cloze <sup>a</sup>	—	88.6 (14.1)	—

<sup>a</sup>Standardized test with  $M = 100$  and  $SD = 15$ .

<sup>b</sup>Standardized test with  $M = 10$  and  $SD = 3$ .

Woodcock et al., 2005; Kruskal-Wallis  $\chi^2 p = 0.46$ ). Median maternal education was higher in the L1 English group (graduate degree) than the L1 Spanish group (bachelor's degree, Fisher's exact  $p = 0.01$ ). L1 English speakers also performed higher on average than L1 Spanish speakers on measures of English vocabulary knowledge (WMLS-R-NU Picture Vocabulary, Woodcock et al., 2005; Kruskal-Wallis  $\chi^2 p = 0.46$ ). Measures of Spanish proficiency (*Vocabulario sobre dibujos*, *Identificación de letras y palabras*, and *Comprensión de textos*, Woodcock et al., 2005) were only administered to L1 Spanish speakers due to L1 English speakers' reported lack of Spanish proficiency and language experience. L1 Spanish speakers' English and Spanish skills were balanced on average, with no significant difference between English and Spanish vocabulary (Wilcoxon signed-rank  $p = 0.76$ ) or cloze ( $p = 0.77$ ) scores. Word identification skills for this group were higher on average in Spanish than English ( $p < 0.001$ ) although English scores on this subtest remained significantly above the population mean of 100 ( $p < 0.001$ ). See **Table 2** for participants' cognitive and language proficiency characteristics and **Supplementary Table SA**, for psychometric information on the standardized assessments.

## Measures

### Cloze Integration Task

Cloze tasks are commonly used as measures of word predictability, as well as individual differences in vocabulary, prediction, and syntactic skills. These tasks were originally developed as an indicator of reading comprehension (Taylor, 1953) and are frequently used as such in both L1 (see Collins et al., 2018 for a meta-analysis) and L2 (Tremblay, 2011; Trace, 2020) education research and practice. However, there is ongoing debate regarding the precise skills that cloze tasks assess, with studies of concurrent validity suggesting that these measures do not align well with other forms of reading comprehension assessment, such as post-reading questioning, whether in multiple-choice or open-ended format (Cutting and Scarborough, 2006; Francis et al., 2006; Keenan et al., 2008; Keenan and Meenan, 2014). There is broad consensus in the

research literature that the manner in which the task is constructed strongly influences the skills it taps for both L1 (Gellert and Elbro, 2013) and L2 speakers (Alderson, 1979; Kleijn et al., 2019). At least for L1 speakers, multi-sentence thematic contexts tap into global, discourse knowledge (Clark and Kamhi, 2014), while cloze assessments such as the WMLS-R and other Woodcock passage comprehension formats with single-sentence stimuli and more weakly constraining thematic contexts draw more strongly into lexical knowledge (Leider et al., 2013), word familiarity (Cutting and Scarborough, 2006; Francis et al., 2006), and syntactic knowledge (Cutting and Scarborough, 2006; Keenan et al., 2008; Deacon and Kieffer, 2018).

We employed a cloze task to assess participants' sentence integration abilities in each language. We administered the English WMLS-R-NU Passage Comprehension to all participants and the Spanish *Comprensión de textos* to L1 Spanish speakers only (Woodcock et al., 2005). The Spanish subtest is a parallel and equated form of the WMLS-R English Passage Comprehension subtest (Woodcock et al., 2005). This cloze task asks the participant to read one or two sentences and orally supply a single missing word. The task starts with stimuli consisting of a single, very brief sentence, expanding to two sentences and/or a longer sentence stimulus as the difficulty of the task increases. Thus, the WMLS-R-NU cloze tasks primarily assess word-to-text integration at the sentence level, particularly for developmental levels of language proficiency. All regression analyses of the cloze tasks also included English and Spanish vocabulary and word identification scores to control for the role of lexical knowledge and word familiarity in sentence comprehension (Francis et al., 2006). Age-scaled standard scores were used in analysis.

### Experimental Stimuli

Stimuli design paralleled that of Whitford and Joannis (2018) and consisted of four expository paragraphs (~100 words each), taken from the Basic Reading Inventory (BRI), 10th ed, a grade-leveled reading inventory commonly used in schools to evaluate reading fluency and comprehension (Johns, 2008). One

**TABLE 3** | Word characteristics for passages at BRI text grade levels 5 through 8.

	Text level 5	Text level 6	Text level 7	Text level 8
Word count	99	100	100	100
Spanish cognates	28	25	26	26
Lexile level	680L	800L	1010L	1020L
SUBTLEX-US word frequency (log), mean, content words only	3.02	3.33	3.66	3.51
SUBTLEX-US word frequency (log), mean, all words	4.20	4.35	4.41	4.42
Mean word length (letters)	4.15	4.25	4.82	4.78
Mean word length (syllables)	1.31	1.32	1.47	1.47
Mean word AoA rating	5.01	4.71	5.29	5.19
Mean word predictability	0.31	0.32	0.24	0.24

paragraph was selected at each BRI level of difficulty corresponding to fifth to eighth grades. Paragraph topics focused on the natural world (ocean, plants, sunflowers, and the environment) and social sciences (indigenous people of the Americas). The BRI itself does not provide validity and reliability statistics for the 10th edition; however, independent analysis of BRI 10th and 11th ed. fluency measures with a younger sample (third through fifth graders) provide some evidence of high test-retest reliability (Pearson  $r = 0.89$  to  $0.93$ ), alternate forms reliability (Pearson  $r = 0.84$  to  $0.96$ ), and concurrent validity with the DIBELS (Pearson  $r = 0.85$  to  $0.97$ , Bieber et al., 2015), a screening assessment also widely used in schools (Good and Kaminsky, 2002). The six factual open-ended comprehension questions were drawn from a set of ten provided in the BRI for each passage and orally administered. Participants responded to questions orally and received dichotomous scores (correct = 1; incorrect = 0) for each comprehension question following the BRI scoring procedure (minimum score = 0 and maximum score = 6 per paragraph).

For each paragraph, we obtained lexical features known to influence reading behavior (Clifton et al., 2007): word age of acquisition (AoA), word frequency, word predictability, orthographic and phonological length, and the number of Spanish cognates (Table 3). Word AoA values were derived from Brysbaert and Biemiller (2017) test-based word AoA ratings. Word frequencies were obtained from the SUBTLEXUS corpus (Brysbaert and New, 2009). Orthographic and phonological length, as well as Levenshtein distances for cognates, were calculated using the R (R Core Team, 2013) package *stringdist* (van der Loo, 2014). Word predictabilities were obtained through a computerized cumulative cloze task (following Whitford and Titone, 2012; Whitford and Titone, 2014; Whitford and Titone, 2017), where a separate sample of adult L1 English speakers ( $n = 30$ ), guessed the words of each paragraph on a word-by-word basis. Accuracy scores were averaged across participants to create a word-level probability of cloze prediction. These word-level features were next averaged over paragraphs to produce mean characteristics for these linguistic units. As seen in Table 3, paragraphs had approximately the same number of words and Spanish cognates. Texts differed only in the average word frequency of their content words ( $p = 0.02$ ) with the lowest log frequency mean in text level 5 and the highest in text level 7. Texts did not differ in

average word frequency, predictability, AoA rating, length, proportion of content to function words, or syllable length (all Kruskal-Wallis  $p > 0.05$ ).

### Paragraph Syntactic Complexity

Each paragraph was evaluated using three classes of syntax measures: 1) classic measures that employ length, or number of words, as an indicator of syntactic complexity; 2) word-level syntax measures; and 3) syntax measures beyond the word level (e.g., examining phrases, clauses, sentences, etc.). Most sentences were composed of a single utterance called a T-unit, itself containing only a single clause. Classic syntactic complexity measures were first generated using the L2 Syntactic Complexity Analyzer (L2SCA; Lu, 2010). L2SCA output was reviewed manually. The greatest numeric variability across paragraphs is captured by sentence length (in number of words), verb phrases, coordinate phrases, and complex nominals. See **Supplementary Table SB**, for more on L2SCA measures. Next, word-level syntax measures were generated using the Tool for the Automatic Analysis of Cohesion 2.0 (TAACO; Crossley et al., 2019), which employs the Stanford Natural Language Parser (NLP). The word-level parses were inspected for errors manually and corrected where needed. TAACO analysis produces part-of-speech (POS) information as well as type-token measures for key function words at the lemma level. Semantic similarity is represented by Word2Vec (Mikolov et al., 2013), a neural network-derived representation of semantic distance among words and phrases. **Supplementary Table SC**, provides further information on the TAACO lexico-syntactic complexity measures. Finally, syntactic complexity measures beyond the word level were generated using the Tool for the Automatic Analysis of Syntactic Complexity (TAASC; Kyle and Crossley, 2018), which computes complexity indicators using Stanford NLP parses. Unlike the TAACO measures, TAASC provides syntactic complexity indicators that reflect relationships among words in a sentence and across sentences in a paragraph. TAASC indicators of syntactic sophistication (average lemma log frequency and average verb-argument construction log frequency) measure the overlap with lemmas and constructions found in the academic sub-corpus of the Corpus of Contemporary American English (COCA; Davies, 2009, 2010 as cited in Kyle and Crossley, 2018). See **Supplementary Table SD**, for more on TAASC phrase-level syntactic complexity measures.



Phrase elaboration, clause complexity, and anaphoric reference were quantified by conducting a nonlinear principal components analysis (PCA) for each characteristic. The nonlinear PCA for phrase complexity incorporated phrase measures from the parsing tools described above: adjective modifiers, nominal dependents, direct and prepositional object dependents, and prepositional phrases as nominal modifiers. The nonlinear PCA for clause complexity included clause measures: relative clause modifiers, dependent clauses, clausal direct objects, and clausal conjunctions. For anaphoric reference, the variables entered into PCA analysis were pronoun density, pronoun-noun ratio, and demonstratives. The first principal component from each of these PCA analyses was used as a corresponding text characteristic predictor in subsequent linear regression models. For phrase complexity, the first principal component accounted for 91% of total variance, for clause complexity 93%, and for anaphoric relations, 67%. See **Supplementary Table SE**, for details.

## Design and Procedure

After consenting, participants completed the eye-tracking task, where they silently read one practice paragraph and four experimental paragraphs at their own pace. After reading each paragraph, they answered six comprehension questions without being able to refer back to the passage. The order of the experimental paragraphs was counterbalanced across the two language groups in a Latin Square design. Lastly, participants completed the behavioral tasks, which were presented in a fixed order: KBIT-2, RAN/RAS, TOWRE-2, English WMLS-R, and then TOMAL-2 for L1 English speakers and KBIT-2, RAN/RAS, Spanish WMLS-R, TOMAL-2, English WMLS-R, and TOWRE-2 for L1 Spanish speakers. Total participation duration was about 3 hours.

## Language of Testing

Instructions for the nonverbal reasoning task (KBIT-2) as well as the full measures of lexical access/naming speed (RAN/RAS) and working memory (TOMAL-2) were administered in the participant's preferred language. Thirteen out of thirty L1 Spanish speakers chose to complete the RAN/RAS in English, and ten out of thirty did so for the TOMAL-2. English language measures (WMLS-R English) were administered in English only, and Spanish language measures (WMLS-R Spanish) were administered in Spanish only.

## Eye-Tracking Procedure

Participants binocularly viewed single paragraphs displayed in yellow text (14 pt, Courier New font, double-spaced) on a black background using Experiment Builder (SR Research Ltd., Ontario, Canada). Each paragraph was presented in its entirety on a 1,024 px × 786 px 21-in. screen positioned 70 cm from the participants, who maintained a fixed head position with the aid of a chin-rest. An EyeLink 1,000 desk-mounted eye-tracker (SR Research Ltd., Ottawa, Canada) collected right-eye monocular data at a 1000 Hz sampling rate. Calibration was performed before the start of each paragraph using a 9-point grid and repeated as necessary until the average fixation error was less than .5° of visual angle.

## Eye-Tracking Data Preprocessing

Trial data were inspected and corrected for vertical drift. Fixations shorter than 80 ms and those outside of word-level interest areas were deleted from the base data. No upper bound was applied to fixation durations; the maximum fixation duration was 3,330 ms (the only duration above 3,000 ms). The next seven largest fixations fell between 2000 and 2,550 ms with all remaining durations shorter than 2000 ms.

Fifteen eye-movement measures were examined based on prior literature (Cop et al., 2015; Rayner, 2012; Whitford and Joannis, 2018; Whitford and Titone, 2012; Whitford and Titone, 2017). Six of these were early stage, local (word-level) measures, which captured unconscious processing of the text during the first reading of the paragraph, also called the first pass or first run: first fixation duration and gaze duration, first pass mean saccade amplitude, first pass regressions out, and first pass word skipping. Nine late-stage eye-movement measures captured conscious integration of information and included all passes through the text: five at a local (word) level (mean fixation duration, total reading time, fixation count, regressions out, and mean saccade amplitude), and the remaining four at a global (trial) passage level (trial fixation count, saccade count, run count, and total trial time). Of these nine late measures, five (total fixation duration, total reading time, regressions out and run count and mean saccade amplitude) provided insight into late processing in specific, local areas of interest while four (total trial fixation count, saccade count, words skipped and trial duration) were indicators of global, or paragraph-level, text processing (see **Table 4** for calculation of these measures).

As illustrated by these related measures, eye-tracking output produces high-dimensional data with resulting analytic challenges. On the one hand, the multiple eye measures provide different insights into the timing of cognitive processing. Particularly for syntactic processing, the cognitive processing may be observed in some measures and not others (Clifton et al., 2007; Rayner, 1998). On the other hand, analysis with multiple, correlated outcomes or predictors augment the likelihood of type I error in eye-tracking analysis (von der Malsburg and Angele, 2017). Data-driven dimensional reduction techniques provide one way to navigate the problems of dimensionality and multiple comparisons (Kuperman et al., 2018). The current study thus employed ordinal principal components analysis (*princals* in the *Gifi* R package, Mair et al., 2017) to extract shared variance from eye-tracking measures; the first principal component from the analysis then served as the single outcome variable representing *reading efficiency* in subsequent regression analysis. Lower values indicate more efficient reading performance.

## Analytic Methods

### Descriptive Statistics

Inspection of the raw data revealed that in most cases, the distribution of behavioral and eye-tracking data violated assumptions of normality and variance homogeneity, and in regression models, of sphericity and homoscedasticity. This analysis thus utilized non-parametric tests implemented in R (R Core Team, 2018) to compute basic descriptive correlations and first-level group comparisons. Specifically, the analysis

**TABLE 4** | Definition of eye-tracking measures.

Measure	Description
<b>Early measures</b>	
First fixation duration	Duration of the 1st fixation on a word
Gaze duration	Summed duration of all 1st pass fixations on a word
First pass saccade amplitude	Amplitude of the 1st saccade entering a word in degrees of visual angle
First pass fixation count	Number of fixations on a word during the first pass through the text
First pass regression count	Number of regressions out of a word during the first pass through the text
Word skipping	Whether a word was skipped during the first pass (yes/no)
<b>Late measures</b>	
Mean fixation duration	Mean duration of all fixations on a word
Total reading time	Summed duration of all fixations on a word
Fixation count	Number of fixations on a word during the trial
Regression out	Number of regressions out of a word during the trial
Mean saccade amplitude	Mean amplitude of all saccades entering a word (in degrees of visual angle)
Trial fixation count	Total number of fixations during the trial
Trial saccade count	Total number of saccades in the trial (paragraph)
Run count	Total number of passes or runs through the text in the trial
Total trial time	Total trial time, from start to when participant presses a key to stop reading

employed Wilcoxon (Mann Whitney) signed rank tests (R package *coin*; Zeileis et al., 2008) and BCa (bias corrected and adjusted) bootstrapped 95% confidence intervals to test differences in sample means (R package *boot*, Canty, 2002); Kendall's test of association (tau) to examine pairwise correlations among numeric variables of interest (in *base R*); one-way Kruskal-Wallis ANOVAs to examine associations between numeric and categorical variables of interest (in *base R*); and repeated measures, robust ANOVA (R package *WRS2*; Mair and Wilcox, 2016).

### Linear Regression Modeling

In order to construct a regression taxonomy for online eye-tracking and offline comprehension outcomes separately, the best-fitting distribution of the outcome variable was first determined through visual inspection, substantive alignment and likelihood ratio tests in R package family *gamlss* (Rigby and Stasinopoulos, 2005). The model taxonomy for analysis comprised a set of multilevel linear regression models using this best-fit outcome distribution with crossed random intercepts for subjects at the paragraph level through R package *lme4* (Bates et al., 2015) and R package family *gamlss* (Stasinopoulos et al., 2017). To examine autocorrelation effects in the eye-tracking data (Baayen et al., 2017), the linear baseline model in each taxonomy was also fitted with cubic polynomial and p-spline smoothers for the behavioral variables of interest using *gamlss*. In each case, BIC model evaluation indicated that the linear model provided the best fit to the data, hence all taxonomies represent linear mixed effects models.

For the taxonomy examining paragraph comprehension as outcome, model predictors were selected in three steps: first, the a priori selection of age, maternal education, and nonverbal reasoning were entered as control variables based on established relationship between these variables and reading outcomes (Burchinal et al., 2002; Hoff, 2013; Peterson et al., 2017; Auerbach et al., 2019; Sorenson Duncan and Paradis, 2020). Second, the full predictor dataset was reduced through

bidirectional stepwise regression minimizing generalized AIC (GAIC). Finally, variables identified in step two were corroborated using ridge regression with 10-fold cross-validation. Optimal model lambda was identified as the penalization factor yielding the lowest mean-squared error out of a range from 0.1 to 50. Ridge regression was considered to corroborate stepwise regression results if variables with ridge coefficients (i.e., effect sizes) greater than 0.10 were the same as those retained in the stepwise regression.

The model taxonomy examining reading efficiency as outcome included: 1) a baseline model in the full L1 English and L1 Spanish sample with word length, frequency, and predictability as control variables and scaled English behavioral predictors; 2) scaled text characteristic predictors were then added to the full sample model and replicated in 3) the sample of L1 Spanish speakers only. The final eye-tracking model for the L1 Spanish sample included scaled Spanish language behavioral assessments and statistically significant interactions between the Spanish language measures and text syntactic characteristics. Residual plots (residuals vs. fitted values, quantile, and residuals vs. leverage) for all models were examined to ensure that model assumptions were not violated.

## RESULTS

### Reading Comprehension

Our first hypotheses concerned cross-linguistic influence of component reading processes on offline comprehension of naturalistic English school texts of varying syntactic complexity. We first identified patterns in children's paragraph comprehension and the association of these patterns with specific syntactic characteristics of the paragraphs using multidimensional scaling (Borg et al., 2012). Next, behavioral measures that characterized paragraph comprehension outcomes were jointly plotted with these syntactic characteristics. Finally,

**TABLE 5 |** Paragraph comprehension accuracy.

	Combined (n = 59) m (sd)	L1 English (n = 29) m (sd)	L1 Spanish (n = 30) m (sd)	Group mean difference (p of Kruskal-Wallis $\chi^2$ )
All paragraphs	0.68 (0.19)	0.72 (0.17)	0.63 (0.21)	0.09 (0.11)
Level 5	0.76 (0.21)	0.82 (0.18)	0.71 (0.23)	0.11 (0.03)*
Level 6	0.78 (0.29)	0.83 (0.21)	0.74 (0.35)	0.09 (0.69)
Level 7	0.62 (0.26)	0.67 (0.26)	0.57 (0.25)	0.10 (0.09)
Level 8	0.54 (0.26)	0.56 (0.30)	0.52 (0.22)	0.04 (0.60)

\*p < .05.

**TABLE 6 |** Results of fitting a taxonomy of multi-level gaussian regression models for English paragraph comprehension scores as a function of control variables age, nonverbal reasoning and maternal education, and predictor variables English word fluency, English vocabulary, English cloze, Spanish vocabulary, and Spanish cloze accuracy, n = 59 middle-schoolers, displaying standardized coefficients).

	Model 1.1 $\beta$ (se)	Model 1.2 $\beta$ (se)	Model 2.1 $\beta$ (se)	Model 2.2 $\beta$ (se)	Model 2.3 $\beta$ (se)
Intercept	0.67 (0.03)***	0.64 (0.02)***	0.73 (0.02)***	0.94 (0.07)***	0.92 (0.06)***
<b>Control variables</b>					
Age	0.02 (0.01)	—	—	—	—
Nonverbal reasoning	0.11 (0.01)***	0.03 (0.01)*	0.00 (0.02)	0.02 (0.02)	—
Maternal education (0 = pre-BA/BS; 1 = BA/BS 2 = MA/MS/PhD)	0.03 (0.04)	—	—	—	—
Group (1 = Spanish)	—	0.07 (0.03)*	—	—	—
<b>Text-level predictors</b>					
Phrase elaboration	—	0.00 (0.03)	0.02 (0.04)	—	—
Clause complexity	—	-0.10 (0.03)***	-0.07 (0.04)	-0.09 (0.02)***	-0.27 (0.05)***
Anaphoric reference	—	-0.00 (0.01)	-0.01 (0.02)	—	—
<b>Participant-level predictors</b>					
English word fluency	—	0.01 (0.01)	-0.04 (0.02)	-0.01 (0.03)	—
English vocabulary	—	0.08 (0.03)**	0.11 (0.03)***	0.10 (0.04)**	0.13 (0.02)***
English cloze	—	0.06 (0.02)*	0.07 (0.03)*	0.03 (0.04)	—
Spanish word reading	—	—	—	-0.02 (0.08)	—
Spanish vocabulary	—	—	—	-0.47 (0.12)***	-0.51 (0.10)***
Spanish cloze	—	—	—	0.26 (0.10)**	0.30 (0.08)***
Clause complexity X Spanish cloze	—	—	—	—	0.19 (0.05)***
$N_{ID}$	59	59	30	30	30
Variance <sub>ID</sub>	0.01	0.01	0.01	0.00	0.00
Observations	235	235	119	119	119
AIC	1.3	-68.4	-35.8	-38.2	-58.4
BIC	132.6	69.8	34.1	18.7	-0.03

\*p < .05, \*\*p < .01, \*\*\*p < .001.

regression taxonomies using behavioral predictors modeled paragraph comprehension outcomes.

### Descriptive Statistics

On average for each paragraph, participants answered four out of the six open-ended questions correctly (mean accuracy = 0.68, sd = 0.19). As **Table 5** illustrates, scores in the full sample were numerically lower for level 7 and 8 paragraphs than for levels 5 and 6, but scores on level 5 and 6 (post-hoc Hochberg family wise error correction,  $p = 0.58$ ) and on level 7 and 8 (Hochberg  $p = 0.06$ ) paragraphs did not differ significantly from each other. When examined pairwise by paragraph without multiple comparisons correction, accuracy means differed across language groups only for the level 5 paragraph (Kruskal-Wallis  $\chi^2 p = 0.03$ ). Variability in scores differed

across paragraphs (Mauchly sphericity test  $p = <0.001$ ), but not across groups (Mauchly's  $p = 0.09$ ), and there was no group by paragraph interaction (Mauchly's  $p = 0.77$ ).

### Linear Regression Modeling

Because multidimensional scaling is primarily an exploratory method that quantifies and visualizes dissimilarities among object scores, which in this study represented participant's accuracy in responding to paragraph questions, we next fit a taxonomy of regression models to the data in order to examine associations among behavioral predictors and paragraph comprehension in a linear regression framework. The first three models in the regression taxonomy focused solely on English language. The final model additionally incorporated behavioral measures of Spanish language skill in asking

whether cross-linguistic vocabulary knowledge and cloze abilities explained variance in paragraph comprehension over and above English measures.

As L1 Spanish and English speakers differed on average in levels of maternal education, the first model in taxonomy one included the control variables of age, nonverbal reasoning (KBIT-2), and maternal education. As displayed in **Table 6**, Model 1.1, no coefficients for age nor any level of maternal education reached significance (all  $p > 0.05$ ), while nonverbal reasoning was positively associated with paragraph comprehension ( $\beta = 0.11, p < 0.001$ ).

Model 1.2 next determined whether paragraph syntactic characteristics and behavioral measures predicted English paragraph comprehension when controlling for nonverbal reasoning, the only significant control variable from Model 1.1. Reading component predictors were word reading fluency, vocabulary and cloze performance, all in English. The three syntactic characteristics of paragraphs in **Table 5** taxonomy, phrase and clause complexity, and anaphoric reference were quantified by conducting a nonlinear principal components analysis (PCA) for each characteristic.

As displayed in Model 1.2, L1 Spanish speakers overall ( $\beta = 0.07, p = 0.02$ ) scored higher on average on paragraph responses when controlling for paragraph syntactic characteristics, nonverbal reasoning, English word fluency, English vocabulary, and English cloze. English vocabulary ( $\beta = 0.08, p = 0.003$ ) and cloze ( $\beta = 0.06, p = 0.01$ ) were positively associated with paragraph comprehension. Clause complexity was negatively related to paragraph response accuracy such that paragraphs with higher clause complexity were associated on average with lower accuracy ( $\beta = -0.10, p < 0.001$ ). Phrase complexity ( $\beta = 0.001, p = 0.96$ ) and anaphoric relations ( $\beta = -0.003, p = 0.82$ ) were not significantly associated with paragraph comprehension.

Because prior research provides conflicting findings on the role of word reading performance in L2 readers, Model 2.1 next examined the Model 1.2 predictors in a reduced sample of only L1 Spanish speakers. Model 2.1 demonstrates that across the subsample and when considering English language performance and cognitive and demographic measures, results mirror those of the full sample, with English vocabulary ( $\beta = 0.11, p = 0.001$ ) and cloze ( $\beta = 0.07, p = 0.02$ ) as the only significant behavioral predictors of accurate paragraph comprehension when controlling for nonverbal reasoning and English word fluency abilities.

Model 2.2 then determined whether Spanish vocabulary and cloze accuracy additionally contributed to explaining variance in English paragraph comprehension for L1 Spanish speakers. When these predictors were entered into a new model that excluded the non-significant syntactic characteristics in Model 2.1, English vocabulary positively predicted paragraph comprehension ( $\beta = 0.10, p = 0.007$ ) just as in the prior model, while Spanish vocabulary was negatively associated with comprehension ( $\beta = -0.47, p < 0.001$ ), such that higher levels of Spanish vocabulary knowledge predicted lower accuracy on the comprehension questions. In contrast to Spanish vocabulary, however, Spanish cloze accuracy in Model 2.2

indicated a positive relationship, such that stronger Spanish cloze abilities predicted greater paragraph comprehension in L2 English ( $\beta = 0.26, p = 0.009$ ) when controlling for English and Spanish vocabulary levels. Furthermore, as seen in the final model of the taxonomy, Model 2.3, which removes non-significant predictors in the interest of parsimony, there was a significant interaction, illustrated in **Figure 3**, between Spanish cloze scores and clause complexity ( $\beta = 0.19, p < 0.001$ ), such that higher Spanish cloze scores attenuated the negative association of paragraph clause complexity with paragraph comprehension.

## Eye Movement Reading Behavior

Using syntactic elements identified based on the text analysis, we next asked what kind of relationship between eye movement measures, syntax, and L1 and L2 reading skills is observed in adolescents when reading naturalistic school texts in English. This second set of analyses identified patterns in adolescents' eye-tracking measures and the association of these patterns with specific syntactic characteristics of the stimulus paragraphs using multidimensional scaling (Borg et al., 2012). Next, behavioral measures that characterized paragraph comprehension outcomes were jointly plotted with these syntactic characteristics. Finally, regression taxonomies using syntactic and behavioral predictors were used to model eye-tracking outcomes.

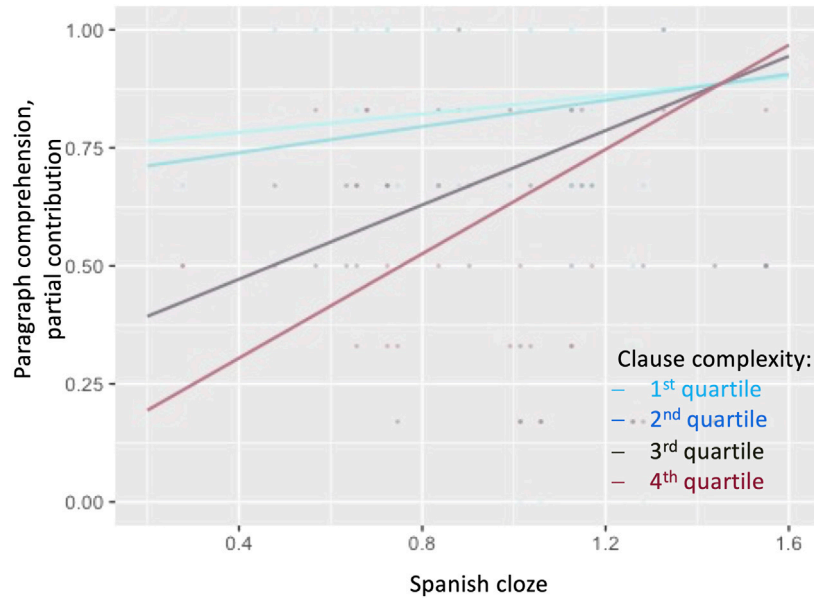
## Descriptive Statistics

Overall, first fixation and mean fixation durations at the word level were comparable for both language groups (Kruskal-Wallis  $p = 0.87$ ). In addition, they did not differ in either first pass ( $p = 0.84$ ) or total ( $p = 0.96$ ) regressions at the word level. However, as **Table 7** demonstrates, and consistent with prior findings on L2 reading (Whitford and Joannis, 2018; Cop et al., 2015), L1 Spanish speakers engaged in significantly more fixations on average at the word ( $p = 0.001$ ) and trial ( $p = 0.002$ ) levels, contributing to a concomitantly longer average trial time ( $p = 0.003$ ). L1 Spanish speakers also skipped fewer words, on average, in both first pass ( $p = 0.008$ ) and total reading ( $p = 0.001$ ), and their saccades were correspondingly shorter in both first pass ( $p = 0.002$ ) and total reading ( $p = 0.002$ ) than those of their L1 English counterparts. The group differences displayed on paragraphs overall also held true for paragraphs when examined separately (all uncorrected Kruskal-Wallis  $p < 0.05$ ).

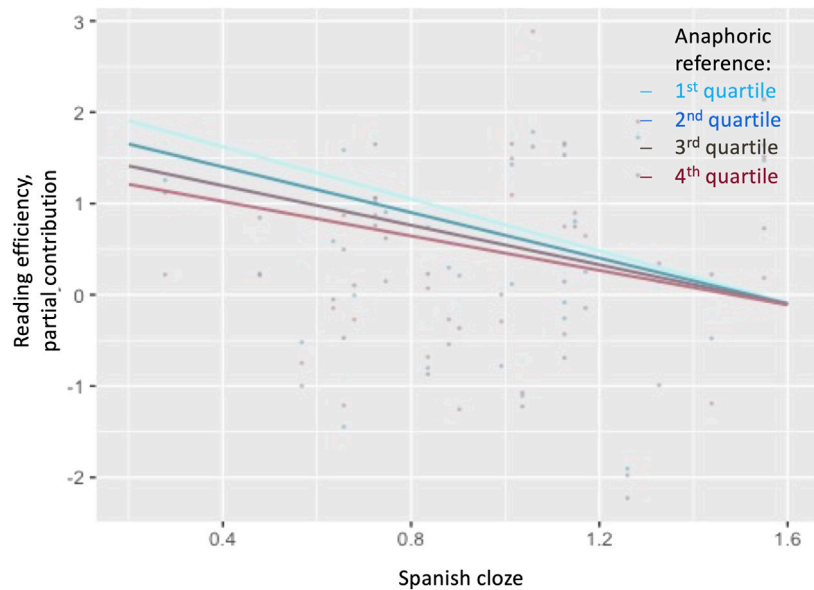
## Regression Modeling

Given the substantively interrelated nature of these measures and their high correlation, as well as loadings on a single PCA dimension, the eye-tracking measurement regression taxonomies used the first principal components dimension, accounting for 53.4% of variance in the 15 eye-tracking measures described above as the outcome variable. All eye measurement variables loaded on this first dimension such that variables typically positively associated with faster and more proficient reading (e.g., word-skipping, saccade amplitude) loaded with opposite sign to variables typically negatively associated with proficient reading (e.g., gaze durations, regressions, number of fixations), which loaded negatively. Secondary analysis using single eye





**FIGURE 3** | Model 2.3 displaying the relationship of L1 Spanish cloze skills with paragraph comprehension for different quartiles of clause complexity ( $\beta = 0.19, p < 0.001$ ).



**FIGURE 4** | Model 4.3 displaying the relationship of L1 Spanish cloze skills with reading efficiency for different quartiles of anaphoric reference (ns interaction).

movement measures (first run gaze duration and total gaze duration) aligned overall with the dimension one regression findings. Lower values of the dimension one measure were therefore indicative of faster and more efficient reading, while higher values indicated slower reading with more fixations and regressions.

Because word characteristics such as length (e.g., Rayner, 2009), frequency (e.g., Rayner and Raney, 1996), and predictability (e.g., Ehrlich and Rayner, 1981) have been widely shown to impact a variety of eye movement measures, Model 3.1 tested the importance of these measures when aggregated at the paragraph level in these experimental stimuli

**TABLE 7 |** Descriptive statistics for eye-tracking measures, by group.

	Combined (n = 59) m (sd)	L1 English (n = 29) m (sd)	L1 Spanish (n = 30) m (sd)	Absolute group mean difference (p of Kruskal-Wallis $\chi^2$ )
<i>Local/Word level—Early</i>				
First fixation duration (ms)	235.5 (29.2)	234.8 (312)	236.1 (27.7)	1.3 (0.63)
Gaze duration (ms)	290.6 (49.7)	278.8 (49.7)	302.0 (47.7)	23.2 (0.02)*
1st pass fixation count	1.24 (0.11)	1.19 (0.08)	1.29 (0.11)	0.10 (<0.001)***
1st pass mean saccade amplitude (°)	6.21 (0.95)	6.55 (1.09)	5.88 (0.65)	0.67 (0.003)**
1st pass regressions out	0.19 (0.07)	0.19 (0.06)	0.20 (0.07)	0.01 (0.84)
1st pass word-skipping rate	0.32 (0.10)	0.36 (0.11)	0.28 (0.07)	0.07 (0.009)**
<i>Local/Word level—Late</i>				
Mean fixation duration (ms)	232.0 (27.6)	232.0 (29.9)	232.0 (25.7)	.02 (0.87)
Total reading time (ms)	366.5 (119.0)	322.5 (89.8)	409.1 (129.3)	86.6 (0.006)**
Mean fixation count	1.57 (0.45)	1.38 (0.27)	1.76 (0.51)	0.38 (0.001)**
Total regressions out	0.33 (0.13)	0.33 (.13)	0.33 (0.13)	0.00 (0.96)
Mean saccade amplitude (°)	5.13 (0.10)	5.51 (1.03)	4.77 (0.82)	0.73 (0.002)**
<i>Global/Text level</i>				
Trial fixation count	156.51 (44.84)	137.14 (27.28)	175.24 (50.64)	38.10 (0.001)**
Trial saccade count	164.17 (46.24)	144.05 (28.90)	183.63 (51.70)	39.57 (0.001)**
Words skipped (%)	19.20 (7.27)	22.07 (6.67)	16.43 (6.83)	5.64 (0.001)**
Total trial duration (ms)	47655 (14404)	42043 (11383)	53079 (15097)	11036 (0.003)**

\*p < .05, \*\*p < .01, \*\*\*p < .001.

**TABLE 8 |** Results of fitting a taxonomy of multi-level gaussian regression models for paragraph reading efficiency (displaying standardized coefficients).

	Model 3.1 $\beta$ (se)	Model 3.2 $\beta$ (se)	Model 4.1 $\beta$ (se)	Model 4.2 $\beta$ (se)	Model 4.3 $\beta$ (se)
Intercept	-0.30 (0.04)***	-0.30 (0.04)***	0.35 (0.04)***	0.48 (0.14)**	0.45 (0.12)**
<b>Control variables</b>					
Group (1 = Spanish)	0.60 (0.07)***	0.60 (0.06)***	—	—	—
Nonverbal reasoning	-0.11 (0.03)***	-0.11 (0.03)***	-0.01 (0.04)	-0.04 (0.04)	—
<b>Word—level predictors</b>					
Word predictability	-0.80 (0.40)*	—	—	—	—
Word frequency (log)	0.22 (0.18)	—	—	—	—
Word length	-1.03 (0.53)	—	—	—	—
<b>Text—level predictors</b>					
Phrase elaboration	—	-0.04 (0.05)	-0.07 (0.08)	—	—
Clausal complexity	—	-0.05 (0.06)	-0.07 (0.08)	—	—
Anaphoric reference	—	-0.08 (0.03)**	-0.10 (0.04)*	-0.12 (0.03)***	-0.30 (0.10)**
<b>Participant—level predictors</b>					
English word fluency	-0.48 (0.03)***	-0.48 (0.03)***	-0.60 (0.05)***	-0.64 (0.05)***	-0.65 (0.05)***
English vocabulary	0.05 (0.05)	0.05 (0.05)	0.25 (0.06)***	0.13 (0.07)	0.15 (0.07)*
English cloze	-0.13 (0.05)**	-0.13 (0.05)**	-0.32 (0.07)***	-0.17 (0.07)*	-0.20 (0.06)**
Spanish word reading	—	—	—	-0.07 (0.16)	—
Spanish vocabulary	—	—	—	1.13 (0.24)***	1.06 (0.22)***
Spanish cloze	—	—	—	-1.18 (0.19)***	-1.17 (0.19)***
Anaphoric reference X sp. cloze	—	—	—	—	0.18 (0.10)†
$N_{ID}$	59	59	30	30	30
Variance <sub>ID</sub>	0.35	0.35	0.36	0.32	0.32
Observations	235	235	119	119	119
AIC	307.1	307.1	157.5	160.2	155.1
BIC	524.6	524.6	257.7	262.3	254.6

†p < .10, \*p < .05, \*\*p < .01, \*\*\*p < .001.

(Table 7). Behavioral and syntactic predictors for the taxonomy in Table 8 were constructed in the same way as for the preceding analysis in Table 5. In Model 3.1, only average word predictability was significantly associated with reading efficiency ( $\beta = -0.80, p = 0.04$ ); however, variation accounted

for by word predictability was colinear with syntactic complexity measures and was therefore not modeled in the remainder of the taxonomy. L1 Spanish speakers, on average, demonstrated reading efficiency outcomes 0.60 of a standard deviation higher than that of L1 English speakers ( $\beta = 0.60, p < 0.001$ )

in Model 3.1 when controlling for nonverbal reasoning, English word fluency and English vocabulary. Unlike models in the paragraph comprehension, English word fluency in Model 3.1 was a significant predictor of reading efficiency ( $\beta = -0.48$ ,  $p < 0.001$ ). However, English vocabulary was not ( $\beta = 0.05$ ,  $p = 0.33$ ), and only English word fluency was associated with lower values of reading efficiency (i.e., shorter gaze durations, faster reading times, fewer saccades, and longer saccade amplitudes) when controlling for L1 and for nonverbal reasoning.

Model 3.2 next determined whether the syntactic predictors of paragraph comprehension were also related to reading efficiency. In this model, neither phrase ( $\beta = -0.04$ ,  $p = 0.45$ ) nor clause ( $\beta = -0.04$ ,  $p = 0.39$ ) complexity predicted reading efficiency; however, on average in the full sample, anaphoric reference (i.e., the proportion of pronouns and demonstratives, and the ratio of pronouns to nouns) was associated with better reading efficiency ( $\beta = -0.08$ ,  $p = 0.006$ ). When re-examined in Model 4.1 with L1 Spanish speakers only, results mirrored those from the full sample, with anaphoric reference similarly associated with better reading efficiency in L2 ( $\beta = -0.10$ ,  $p = 0.01$ ).

As in the paragraph response taxonomy, Model 4.2 next determined whether Spanish language and reading skills additionally contributed to explaining variance in the outcome measure. When Spanish predictors were entered into a new model, the coefficient for Spanish vocabulary was once again positive ( $\beta = 1.13$ ,  $p < 0.001$ ) such that a larger Spanish vocabulary predicted worse reading efficiency (i.e., longer fixation durations, longer reading times, shorter saccades and less word skipping). In contrast, Spanish cloze in Model 4.2 displayed a negative coefficient ( $\beta = -1.20$ ,  $p < 0.001$ ), indicating that higher standardized scores in Spanish cloze were associated with shorter fixations, faster reading times, longer saccades, and more word skipping. Furthermore, in the final model, Model 4.3 illustrated in **Figure 4**, there was a trending, but non-significant, positive interaction between the anaphoric reference measure and Spanish cloze ( $\beta = 0.18$ ,  $p = 0.08$ ), such that in passages with more anaphoric references, higher levels of Spanish cloze were trending toward, but not significantly associated with, lower decrements in reading efficiency, as compared to passages with fewer anaphoric references.

## DISCUSSION

This study examined the contributions of cross-linguistic sentence integration skills as well as the lexical skills of decoding and vocabulary knowledge to the processing and comprehension of naturalistic English texts in adolescents with either English or Spanish as their L1. In particular, it focused on the modulation of these skills by syntactic structures that have been demonstrated to challenge middle-school readers, namely, complex noun phrases, complex clauses, and anaphoric references. Middle-schoolers who spoke English or Spanish as an L1 read nonfiction passages in English that were leveled from grades 5 to 8 while eye-tracking measures were collected to assess reading efficiency (fluency and speed of reading). Participants also answered post-reading questions about the passages as a measure of comprehension.

## RQ1: Are Greater L1 and L2 Proficiency in Sentence Integration Associated With Higher Levels of L2 Text Reading Efficiency and Comprehension?

We hypothesized that greater L1 and L2 proficiency in sentence integration would be associated with higher levels of L2 text comprehension and that phrase complexity, clause complexity, and anaphoric references would all negatively modulate the expected positive effect of cross-linguistic sentence integration skills on offline text comprehension. English sentence integration and vocabulary scores were positive predictors of English text comprehension, and this relationship was similar for both L1 English and L1 Spanish speakers. This result is consistent with prior studies of English learners' reading development (e.g., Lesaux et al., 2007; Lesaux and Harris, 2017) and with most reading comprehension models, including the RSF (Perfetti and Stafura, 2014), which generally highlight the important role of skills in the language of the text in both reading fluency and comprehension.

We expected that L1 sentence integration abilities would support L2 text comprehension, and this was indeed the case when controlling for the effect of Spanish L1 vocabulary. Prior literature has found weak or nonexistent associations of L1 sentence integration with L2 reading among students in L2 dominant school environments but more consistent positive associations among students in bilingual education settings where students receive academic instruction in both languages. Although the L1 Spanish speakers in our sample, except for two in bilingual English-Spanish schools, attended almost entirely English-speaking schools, they still retained relatively balanced or Spanish-dominant language skills when we tested them in middle school. Our biliterate sample was also reported by parents to consistently engage in reading in Spanish at home, with an average of 34% of daily reading reported to take place in Spanish. Thus, in spite of educational settings similar to those found in prior studies with no CLI effects in text reading, our sample possessed proficiency in L1 Spanish reading that may have led to outcomes more similar to students in bilingual education than ones in English-only instruction. If the transfer of a common underlying proficiency, as in Cummins' (1979) proposal, requires some minimal level of L1 abilities, the home language and literacy experiences in our sample appear to have sustained at least a level of Spanish proficiency that supported the transfer of sentence integration skills to English reading.

The positive association of L1 sentence integration skills with reading was not limited to offline, post-reading comprehension but also seen in reading efficiency, where higher levels of Spanish sentence integration were associated with faster and more efficient online text processing. The parallel findings in both online behavior (seen in eye-tracking) and offline behavior (seen in the post-reading comprehension task) suggests that CLI of L1 Spanish sentence integration is not, or is not solely, a post-reading process of reasoning or strategizing about the comprehension questions asked, or re-creating a representation of the text meaning in relation to the comprehension questions asked of students after reading. Instead, these triangulated results suggest that CLI occurs during the reading process as readers are

integrating words into larger sentence and text representations. For monolingual models of text processing and comprehension such as the RSF, these findings indicate that linguistic resources beyond the language of the text can support this integration process even when lexical knowledge in that language, such as reflected in our L1 Spanish speakers' English vocabulary scores, may not provide the high-quality lexical representations of meaning that are called for in the RSF to support sentence- and text-level meaning integration. While sentence integration in the RSF model is generally assumed to be a within-language skill supporting word-to-text integration, results from both the behavioral and eye-movement analyses thus indicate that a cross-language sentence integration competency beyond vocabulary knowledge may also support text processing and comprehension in an L2.

## **RQ2: Does Syntactic Complexity of the Text Modulate the Association of L1 Sentence Integration Skills With L2 Text Reading Efficiency and Comprehension?**

Because syntactically complex texts present greater difficulty in both L1 and L2 reading, we had hypothesized that syntactic complexity of the text would modulate the association between L1 sentence integration skills and both text processing and comprehension. Our results demonstrate a significant negative main effect of syntactic complexity on paragraph comprehension. In other words, more complex texts, in terms of clause complexity, were more difficult to comprehend both for the full sample and for L1 Spanish speakers when examined separately. Further, clause complexity positively moderated the relationship of Spanish sentence integration skills with offline comprehension such that the positive association of Spanish sentence integration with L2 comprehension was more pronounced in texts with greater clause complexity. In the least complex passages, there was no evident relationship between L1 sentence integration and paragraph comprehension while in the most complex passages, higher L1 sentence integration skills were associated with better paragraph comprehension. L2 comprehension thus appears to draw more heavily on sentence integration skills developed in the L1 when sentences are more structurally complex and difficult to understand.

In online text processing, we again found that syntactic characteristics of the text were associated with reading efficiency. However, unlike for comprehension, better reading efficiency was predicted by the greater presence of anaphoric references and not by clause complexity. In other words, passages with more anaphoric references were read faster and more efficiently than those with fewer such references. In the study passages, these text features may have provided links among concepts in the text that led to more fluent and efficient online text processing. We also found a trend, but non-significant association for the interaction of anaphoric references and L1 sentence integration skills, raising the question of whether in a larger sample or a longer paradigm, these text characteristics would modulate the

facilitative main effect of L1 sentence integration on text processing efficiency.

While the present exploratory study cannot identify mechanisms for CLI in complex text, the RSF was formulated as a bottom-up, memory-based integration model, suggesting that beyond lexical representation, individual differences in short-term and working memory capacities could explain comprehension outcomes. Results of the present study do not support a memory-based explanation of individual differences as our measures of sentence integration in both Spanish and English were written and untimed, and in group comparisons, L1 Spanish and L1 English speakers were matched on verbal working memory and nonverbal reasoning. However, the study findings do suggest that in comprehending text with complex clauses that may be likely to tax memory resources, bilinguals may be drawing upon expectation-based meaning construction skills measured by our sentence integration task. These skills may be specific to features shared by Spanish and English or shared in the academic register. For example, while Spanish and English share a canonical subject-verb-object word order in simple sentences, we might speculate that the relative flexibility of Spanish regarding word order could facilitate comprehension of variation in word order in academic language marked by complex clauses. For Spanish speakers, a bias toward disregarding canonical word order in complex clauses, or toward attending to lexical cues in anaphoric references, could conceivably thus facilitate text processing and comprehension in texts containing anaphors and complex clausal structures. The crosslinguistic influence as well as the modulatory effect of text complexity on CLI demonstrated here may thus be specific to a feature such as word order, or combinatorial similarities in a particular language, Spanish, or pair of languages, Spanish and English. Alternatively, crosslinguistic influence might arise from language-general factors such as general skills with statistical learning, predictive processing, parsing, or inferencing. Future studies that employ more than one task, or more specific tasks, may disentangle these or other possibilities.

## **CONCLUSION**

Our findings extend monolingual reading models such as the RSF by providing evidence at one level, that of word to sentence and text integration, of the cross-linguistic influence of L1 sentence integration skills in L2 reading. Further, the cross-linguistic resources reflected in bilingual text processing and comprehension are dynamic and modulated by syntactic complexity of the text. Bilingual resources in integrating word meanings may thus be sensitive to text-level features not seen in studies of single word and sentence processing. The present study thus also extends word level models in bilingual research which propose a unified model of language processing and extend the implications of lexical models to sentence integration in the comprehension of connected text. Results also suggest that sentence integration may involve unified components beyond the lexicon. Future research may thus investigate how monolingual construction-integration models of reading comprehension, such as



the RSF, can incorporate bilingual processing and cross-linguistic influence, not only at the lower levels of phonology, orthography, morphology and the lexicon, but also in integrative processes that are involved in comprehending text features such as the complex clauses and anaphoric references examined in this study.

Because the study included only one global measure of sentence integration, a cloze task, one limitation is that it cannot differentiate among multiple forms of sentence integration that may have contributed to the results, such as word order sequencing, parsing, or referential association. In addition, there was substantial heterogeneity among the L1 Spanish speakers in terms of age of English acquisition and Spanish language ability. This heterogeneity reflects the linguistic and demographic mix in many United States schools and was partially accounted for using mixed effects modeling instead of simple group comparisons. However, it also demands caution in interpreting the effects of predictors that are correlated in the sample, such as English and Spanish vocabulary. Future research might attend to age of acquisition and language skill differences in a larger and longitudinal sample of L2 readers.

In summary, the current study suggests that L1 sentence integration skills can facilitate L2 reading efficiency and comprehension. L1 sentence integration skill in particular appears to support comprehension of complex clauses, even when that text is in an L2. It adds to prior educational research on cross-linguistic influence on reading outcomes using a sample representative of United States public schools, which are largely English-only and do not provide home language literacy instruction. In spite of the English-only educational background of our sample, we found a facilitatory CLI similar to that found in prior studies of children in bilingual immersion education settings, suggesting that when home language skills are developed and maintained outside of the school context, these skills can help support efficiency and comprehension of L2 reading tasks required of children at school. In supplement to bilingual education, practices and policies which support family—or community-based, out-of-school, home language development may thus also support minority language speakers' success in L2 education.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

## REFERENCES

- Aguilar, G., Uccelli, P., and Galloway, E. P. (2020). Toward Bilingualism: Unpacking the Contribution of Mid-Adolescent Dual Language Learners' Spanish and English Academic Language Skills to English Reading Comprehension. *TESOL Q.* 54 (4), 1010–1036. doi:10.1002/tesq.570
- Alonso, R. A. (2019). "Crosslinguistic Influence in Second Language Acquisition," in *The Encyclopedia of Applied Linguistics*. Editor C. A. Chapelle, 1–7. doi:10.1002/9781405198431.wbeal0292 (Ed.)

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Harvard University Committee on the Use of Human Subjects (IRB16-0866). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

SL contributed to conceptualization, data collection and curation, designed and conducted the statistical analysis, and wrote the manuscript. VW contributed to conceptualization, data curation, and reviewed and edited the manuscript. LM contributed to conceptualization, data collection, and reviewed and edited the manuscript. GL contributed to conceptualization, funding acquisition, reviewed and edited the manuscript, and supervised the project. All authors contributed to the article and approved the submitted version.

## FUNDING

We would like to thank Dr. James Ryan, former Dean of the Harvard Graduate School of Education, as well as the Natural Sciences and Engineering Research Council of Canada (RGPIN-2020-05052), for providing financial support for the study (both awarded to GL). The study was carried out at the Harvard Center for Brain Science and involved the use of instrumentation supported by the NIH Shared Instrumentation Grant Program (S10OD020039).

## ACKNOWLEDGMENTS

We would like to express our sincerest gratitude to the children and families who participated in the research. We also thank the dedicated BEE lab members who assisted in recruiting and interacting with the study participants.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcomm.2021.651769/full#supplementary-material>

- Anderson, J. A. E., Mak, L., Keyvani Chahi, A., and Bialystok, E. (2018). The Language and Social Background Questionnaire: Assessing Degree of Bilingualism in a Diverse Population. *Behav. Res.* 50 (1), 250–263. doi:10.3758/s13428-017-0867-9
- Auerbach, J. G., Zilberman-Hayun, Y., Berger, A., and Atzaba-Poria, N. (2019). Longitudinal Pathways to reading Achievement at Age 7 years: Child and Environmental Influences. *Reading Psychol.* 40 (3), 269–292. doi:10.1080/02702711.2019.1614128
- Baayen, H., Vasishth, S., Kliegl, R., and Bates, D. (2017). The Cave of Shadows: Addressing the Human Factor with Generalized Additive Mixed Models. *J. Mem. Lang.* 94, 206–234. doi:10.1016/j.jml.2016.11.006

- Babayigit, S., and Shapiro, L. (2020). Component Skills that Underpin Listening Comprehension and reading Comprehension in Learners with English as First and Additional Language. *J. Res. Reading* 43 (1), 78–97.
- Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Soft.* 67 (1), 1–48. doi:10.18637/jss.v067.i01
- Bialystok, E., Luk, G., and Kwan, E. (2005). Bilingualism, Biliteracy, and Learning to Read: Interactions Among Languages and Writing Systems. *Scientific Stud. Reading* 9 (1), 43–61. doi:10.1207/s1532799xssr0901\_4
- Bieber, G., Hulac, D. M., and Schweinle, W. (2015). An Independent Evaluation of the Technical Features of the Basic Reading Inventory. *J. Psychoeducational Assess.* 33 (3), 199–209. doi:10.1177/0734282914550808
- Blythe, H. I., and Joseph, H. S. S. L. (2011). “Children’s Eye Movements During Reading” in *The Oxford Handbook of Eye Movements*. Oxford University Press, 643–662.
- Borg, I., Groenen, P. J., and Mair, P. (2012). *Applied Multidimensional Scaling*. Springer.
- Bosch, L. J., Segers, E., and Verhoeven, L. (2020). First and Second Language Vocabulary Affect Early Second Language reading Comprehension Development. *J. Res. Reading* 43 (3), 290–308. doi:10.1111/1467-9817.12304
- Brimo, D., Apel, K., and Fountain, T. (2017). Examining the Contributions of Syntactic Awareness and Syntactic Knowledge to Reading Comprehension. *J. Res. Read* 40(1), 57–74. doi:10.1111/1467-9817.12050
- Brysaert, M., and Biemiller, A. (2017). Test-based Age-Of-Acquisition Norms for 44 Thousand English Word Meanings. *Behav. Res.* 49 (4), 1520–1523. doi:10.3758/s13428-016-0811-4
- Brysaert, M., and New, B. (2009). Moving beyond Kučera and Francis: A Critical Evaluation of Current Word Frequency Norms and the Introduction of a New and Improved Word Frequency Measure for American English. *Behav. Res. Methods* 41 (4), 977–990. doi:10.3758/BRM.41.4.977
- Burchinal, M. R., Peisner-Feinberg, E., Pianta, R., and Howes, C. (2002). Development of Academic Skills from Preschool through Second Grade: Family and Classroom Predictors of Developmental Trajectories. *J. Sch. Psychol.* 40 (5), 415–436. doi:10.1016/S0022-4405(02)00107-3
- Cain, K., and Oakhill, J. (2004). “Reading Comprehension Difficulties,” in *Handbook of Children’s Literacy*. Editors T. Nunes and P. Bryant (Springer Netherlands), 313–338. doi:10.1007/978-94-017-1731-1\_18
- Canty, A. J. (2002). Resampling Methods in R: the Boot Package. *The Newsl. R Project* 2, 3. Available at: [http://cran.fhcr.org/doc/Rnews/Rnews\\_2002-3.pdf](http://cran.fhcr.org/doc/Rnews/Rnews_2002-3.pdf).
- Chung, S. C., Chen, X., and Geva, E. (2019). Deconstructing and Reconstructing Cross-Language Transfer in Bilingual reading Development: An Interactive Framework. *J. Neurolinguist.* 50, 149–161. doi:10.1016/j.jneuroling.2018.01.003
- Clark, M. K., and Kamhi, A. G. (2014). Influence of Prior Knowledge and Interest on Fourth- and Fifth-Grade Passage Comprehension on the Qualitative Reading Inventory-4. *Lshss* 45 (4), 291–301. doi:10.1044/2014\_LSHSS-13-0074
- Clifton, C., Staub, A., and Rayner, K. (2007). “Eye Movements in Reading Words and Sentences” in *Eye Movements*. Elsevier, 341–371.
- Collins, A. A., Lindström, E. R., and Compton, D. L. (2018). Comparing Students with and without reading Difficulties on reading Comprehension Assessments: A Meta-Analysis. *J. Learn. Disabil.* 51 (2), 108–123. doi:10.1177/0022219417704636
- Cop, U., Keuleers, E., Drieghe, D., and Duyck, W. (2015). Frequency Effects in Monolingual and Bilingual Natural reading. *Psychon. Bull. Rev.* 22 (5), 1216–1234. doi:10.3758/s13423-015-0819-2
- Crossley, S. A., Kyle, K., and Dascalu, M. (2019). The Tool for the Automatic Analysis of Cohesion 2.0: Integrating Semantic Similarity and Text Overlap. *Behav. Res.* 51 (1), 14–27. doi:10.3758/s13428-018-1142-4
- Cummins, J. (1979). *Cognitive/academic Language Proficiency, Linguistic Interdependence, the Optimum Age Question and Some Other Matters*. Ontario Institute for Studies in Education.
- Cummins, J. (2000). *Language, Power, and Pedagogy: Bilingual Children in the Crossfire*. Multilingual Matters. doi:10.21832/9781853596773
- Cunnings, I. (2017). Parsing and Working Memory in Bilingual Sentence Processing. *Biling.: Lang. Cogn.* 20 (04), 659–678. doi:10.1017/S1366728916000675
- Cutting, L. E., and Scarborough, H. S. (2006). Prediction of reading Comprehension: Relative Contributions of Word Recognition, Language Proficiency, and Other Cognitive Skills Can Depend on How Comprehension Is Measured. *Scientific Stud. Reading* 10 (3), 277–299. doi:10.1207/s1532799xssr1003\_5
- De Groot, H. (2018). *An Eye-Tracking Investigation of L2 Readers of a Novel by Agatha Christie*. dissertation. Ghent University. doi:10.21468/scipost.report.674
- de Leeuw, L., Segers, E., and Verhoeven, L. (2016). The Effect of Student-Related and Text-Related Characteristics on Student’s Reading Behaviour and Text Comprehension: An Eye Movement Study. *Scientific Stud. Reading* 20 (3), 248–263. doi:10.1080/10888438.2016.1146285
- Deacon, S. H., and Kieffer, M. (2018). Understanding How Syntactic Awareness Contributes to reading Comprehension: Evidence from Mediation and Longitudinal Models. *J. Educ. Psychol.* 110 (1), 72–86. doi:10.1037/edu0000198
- Durgunoglu, A. Y., Nagy, W. E., and Hancin-Bhatt, B. J. (1993). Cross-language Transfer of Phonological Awareness. *J. Educ. Psychol.* 85 (3), 453.
- Dussias, P. E. (2010). Uses of Eye-Tracking Data in Second Language Sentence Processing Research. *Ann. Rev. Appl. Linguist* 30, 149–166. doi:10.1017/S026719051000005X
- Ehrlich, S. F., and Rayner, K. (1981). Contextual Effects on Word Perception and Eye Movements during reading. *J. Verbal Learn. Verbal Behav.* 20 (6), 641–655. doi:10.1016/s0022-5371(81)90220-6
- Francis, D. J., Snow, C. E., August, D., Carlson, C. D., Miller, J., and Iglesias, A. (2006). Measures of reading Comprehension: A Latent Variable Analysis of the Diagnostic Assessment of reading Comprehension. *Scientific Stud. Reading* 10 (3), 301–322. doi:10.1207/s1532799xssr1003\_6
- Gabriele, A., Troseth, E., Martohardjono, G., and Otheguy, R. (2009). Emergent Literacy Skills in Bilingual Children: Evidence for the Role of L1 Syntactic Comprehension. *Int. J. Bilingual Education Bilingualism* 12 (5), 533–547. doi:10.1080/13670050802331816
- Garrison-Fletcher, L., Martohardjono, G., and Chodorow, M. (2019). Impact of home Language reading Skills’ on L2 reading Comprehension: a Study of Newcomer Emergent Bilinguals. *Int. J. Bilingual Education Bilingualism* 1, 1–14. doi:10.1080/13670050.2019.1671307
- Gellert, A. S., and Elbro, C. (2013). Cloze Tests May Be Quick, but Are They Dirty? Development and Preliminary Validation of a Cloze Test of reading Comprehension. *J. Psychoeducational Assess.* 31 (1), 16–28. doi:10.1177/0734282912451971
- Genesee, F., and Geva, E. (2006). “Cross-Linguistic Relationships in Working Memory, Phonological Processes, and Oral Language,” in *Developing Literacy in Second-Language Learners: Report of the National Literacy Panel on Language-Minority Children and Youth* (Lawrence Erlbaum Associates Publishers), 175–183.
- Geva, E., and Farnia, F. (2012). Developmental Changes in the Nature of Language Proficiency and reading Fluency Paint a More Complex View of reading Comprehension in ELL and EL1. *Read. Writ* 25 (8), 1819–1845. doi:10.1007/s11145-011-9333-8
- Geva, E., and Siegel, L. S. (2000). Orthographic and Cognitive Factors in the Concurrent Development of Basic Reading Skills in Two Languages. *Read Writ.* 12 (1), 1–30. doi:10.1023/A:1008017710115
- Godfroid, A. (2019). *Eye Tracking in Second Language Acquisition and Bilingualism: A Research Synthesis and Methodological Guide*. Routledge. doi:10.4324/9781315775616
- Gollan, T. H., Montoya, R. I., Cera, C., and Sandoval, T. C. (2008). More Use Almost Always Means a Smaller Frequency Effect: Aging, Bilingualism, and the Weaker Links Hypothesis☆. *J. Mem. Lang.* 58 (3), 787–814. doi:10.1016/j.jml.2007.07.001
- Good, R., and Kaminsky, R. (2002). *Dynamic Indicators of Basic Early Literacy Skills*. Institute for the Development of Educational Achievement. <https://dibels.uoregon.edu/>.
- Gottardo, A., Mirza, A., Koh, P. W., Ferreira, A., and Javier, C. (2018). Unpacking Listening Comprehension: The Role of Vocabulary, Morphological Awareness, and Syntactic Knowledge in reading Comprehension. *Read. Writ* 31 (8), 1741–1764. doi:10.1007/s11145-017-9736-2
- Hoff, E. (2013). Interpreting the Early Language Trajectories of Children from Low-Ses and Language Minority Homes: Implications for Closing Achievement Gaps. *Developmental Psychol.* 49 (1), 4–14. doi:10.1037/a0027238

- Jeon, E. H., and Yamashita, J. (2014). L2 reading Comprehension and its Correlates: A Meta-Analysis. *Lang. Learn.* 64 (1), 160–212. doi:10.1111/lang.12034
- Johns, J. L. (2008). *Basic Reading Inventory: Pre-primer through Grade Twelve and Early Literacy Assessments*. 10th edn. Kendall Hunt.
- Joseph, H. S. S. L., Bremner, G., Liversedge, S. P., and Nation, K. (2015). Working Memory, reading Ability and the Effects of Distance and Typicality on Anaphor Resolution in Children. *J. Cogn. Psychol.* 27 (5), 622–639. doi:10.1080/20445911.2015.1005095
- Kaufman, A. S., and Kaufman, N. L. (2004). *Manual for the Kaufman Brief Intelligence Test*. 2nd Edn. Circle Pines, MN: American Guidance Service.
- Keenan, J. M., Betjemann, R. S., and Olson, R. K. (2008). Reading Comprehension Tests Vary in the Skills They Assess: Differential Dependence on Decoding and Oral Comprehension. *Scientific Stud. Reading* 12 (3), 281–300. doi:10.1080/10888430802132279
- Keenan, J. M., and Meenan, C. E. (2014). Test Differences in Diagnosing reading Comprehension Deficits. *J. Learn. Disabil.* 47 (2), 125–135. doi:10.1177/0022219412439326
- Kieffer, M. J. (2012). Early Oral Language and Later reading Development in Spanish-speaking English Language Learners: Evidence from a Nine-Year Longitudinal Study. *J. Appl. Developmental Psychol.* 33 (3), 146–157. doi:10.1016/j.appdev.2012.02.003
- Kim, M., Crossley, S. A., and Skalicky, S. (2018). Effects of Lexical Features, Textual Properties, and Individual Differences on Word Processing Times during Second Language reading Comprehension. *Read. Writ* 31 (5), 1155–1180. doi:10.1007/s11145-018-9833-x
- Kintsch, W., and van Dijk, T. A. (1978). Toward a Model of Text Comprehension and Production. *Psychol. Rev.* 85 (5), 363–394. doi:10.1037/0033-295X.85.5.363
- Kremin, L. V., Arredondo, M. M., Hsu, L. S.-J., Satterfield, T., and Kovelman, I. (2019). The Effects of Spanish Heritage Language Literacy on English reading for Spanish-English Bilingual Children in the US. *Int. J. Bilingual Education Bilingualism* 22 (2), 192–206. doi:10.1080/13670050.2016.1239692
- Kuperman, V., Matsuki, K., and Van Dyke, J. A. (2018). Contributions of Reader- and Text-Level Characteristics to Eye-Movement Patterns during Passage reading. *J. Exp. Psychol. Learn. Mem. Cogn.* 44 (11), 1687–1713. doi:10.1037/xlm0000547
- Kyle, K., and Crossley, S. A. (2018). Measuring Syntactic Complexity in L2 Writing Using fine-grained Clausal and Phrasal Indices. *Mod. Lang. J.* 102 (2), 333–349. doi:10.1111/modl.12468
- Leider, C. M., Proctor, C. P., Silverman, R. D., and Harring, 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. *Read. Writ* 26 (9), 1459–1485. doi:10.1007/s11145-013-9427-6
- Lervåg, A., Hulme, C., and Melby-Lervåg, M. (2018). Unpicking the Developmental Relationship Between Oral Language Skills and Reading Comprehension: It's Simple, But Complex. *Child Dev.* 89 (5), 1821–1838. doi:10.1111/cdev.12861
- Lesaux, N. K., Crosson, A. C., Kieffer, M. J., and Pierce, M. (2010). Uneven Profiles: Language Minority Learners' Word reading, Vocabulary, and reading Comprehension Skills. *J. Appl. Developmental Psychol.* 31 (6), 475–483. doi:10.1016/j.appdev.2010.09.004
- Lesaux, N. K., and Harris, J. R. (2017). An Investigation of Comprehension Processes Among Adolescent English Learners with reading Difficulties. *Top. Lang. Disord.* 37 (2), 182–203. doi:10.1097/tld.0000000000000120
- Lesaux, N. K., Lipka, O., and Siegel, L. S. (2006). Investigating Cognitive and Linguistic Abilities that Influence the reading Comprehension Skills of Children from Diverse Linguistic Backgrounds. *Read. Writ* 19 (1), 99–131. doi:10.1007/s11145-005-4713-6
- Lesaux, N. K., Rupp, A. A., and Siegel, L. S. (2007). Growth in reading Skills of Children from Diverse Linguistic Backgrounds: Findings from a 5-year Longitudinal Study. *J. Educ. Psychol.* 99 (4), 821–834. doi:10.1037/0022-0663.99.4.821
- Lesaux, N. K., and Siegel, L. S. (2003). The Development of reading in Children Who Speak English as a Second Language. *Developmental Psychol.* 39 (6), 1005–1019. doi:10.1037/0012-1649.39.6.1005
- Levy, R. (2008). Expectation-based Syntactic Comprehension. *Cognition* 106 (3), 1126–1177. doi:10.1016/j.cognition.2007.05.006
- Lindsey, K. A., Manis, F. R., and Bailey, C. E. (2003). Prediction of First-Grade reading in Spanish-speaking English-language Learners. *J. Educ. Psychol.* 95 (3), 482–494. doi:10.1037/0022-0663.95.3.482
- Low, P. B., and Siegel, L. S. (2005). A Comparison of the Cognitive Processes Underlying reading Comprehension in Native English and ESL Speakers. *WL&L* 8 (2), 131–155. doi:10.1075/wll.8.2.09low
- Lu, X. (2010). Automatic Analysis of Syntactic Complexity in Second Language Writing. *Ijcl* 15 (4), 474–496. doi:10.1075/ijcl.15.4.02lu
- Lubliner, S., and Hiebert, E. H. (2011). An Analysis of English-Spanish Cognates as a Source of General Academic Language. *Biling. Res. J.* 34 (1), 76–93. doi:10.1080/15235882.2011.568589
- Luk, G., and Bialystok, E. (2013). Bilingualism Is Not a Categorical Variable: Interaction between Language Proficiency and Usage. *J. Cogn. Psychol.* 25 (5), 605–621. doi:10.1080/20445911.2013.795574
- Luk, G., and Christodoulou, J. (2016). “Assessing and Understanding the Needs of Dual-Language Learners,” in *The Leading Edge of Early Childhood Education: Linking Science to Policy for a New Generation*. Editors N. Lesaux and S. Jones (Harvard Education Press), 67–90.
- Mair, P., De Leeuw, J., and Groenen, P. (2017). *Gifi: Multivariate Analysis with Optimal Scaling. R Package-Version 0.3-7/r266*.
- Mair, Patrick., and Wilcox, R. (2016). Robust Statistical Methods in R Using the Wrs2 Package. *Behav. Res. Methods*, 1–25.
- Melby-Lervåg, M., and Lervåg, A. (2011). Cross-linguistic Transfer of Oral Language, Decoding, Phonological Awareness and reading Comprehension: A Meta-Analysis of the Correlational Evidence. *J. Res. Reading* 34 (1), 114–135. doi:10.1111/j.1467-9817.2010.01477.x
- Melby-Lervåg, M., and Lervåg, A. (2014). Reading Comprehension and its Underlying Components in Second-Language Learners: A Meta-Analysis of Studies Comparing First- and Second-Language Learners. *Psychol. Bull.* 140 (2), 409–433. doi:10.1037/a0033890
- Mikolov, T., Yih, W., and Zweig, G. (2013). “Linguistic Regularities in Continuous Space Word Representations,” in Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Atlanta, GA: Association for Computational Linguistics), 746–751. Available at: <https://www.aclweb.org/anthology/N13-1090>
- Mulder, E., van de Ven, M., Segers, E., Krepel, A., de Bree, E. H., de Jong, P. F., et al. (2020). Word-to-text Integration in English as a Second Language reading Comprehension. *Read. Writ* 34, 1049–1087. doi:10.1007/s11145-020-10097-3
- Nakamoto, J., Lindsey, K. A., and Manis, F. R. (2012). Development of Reading Skills from K-3 in Spanish-Speaking English Language Learners Following Three Programs of Instruction. *Read. Writ.* 25 (2), 537–567. doi:10.1007/s11145-010-9285-4
- Nation, K., and Snowling, M. J. (2000). Factors Influencing Syntactic Awareness Skills in Normal Readers and Poor Comprehenders. *Appl. Psycholinguist.* 21 (2), 229–241. doi:10.1017/S0142716400002046
- Ordóñez, C., Carlo, M., Snow, C., and McLaughlin, B. (2002). Depth and Breadth of Vocabulary in Two Languages: Which Vocabulary Skills Transfer? *J. Edu. Psychol.* 94 (4), 719–728. doi:10.1037/0022-0663.94.4.719
- Pasquarella, A., Gottardo, A., and Grant, A. (2012). Comparing Factors Related to reading Comprehension in Adolescents Who Speak English as a First (L1) or Second (L2) Language. *Scientific Stud. Reading* 16 (6), 475–503. doi:10.1080/10888438.2011.593066
- Perfetti, C., and Stafura, J. (2014). Word Knowledge in a Theory of reading Comprehension. *Scientific Stud. Reading* 18 (1), 22–37. doi:10.1080/10888438.2013.827687
- Peterson, R. L., Boada, R., McGrath, L. M., Willcutt, E. G., Olson, R. K., and Pennington, B. F. (2017). Cognitive Prediction of reading, Math, and Attention: Shared and Unique Influences. *J. Learn. Disabil.* 50 (4), 408–421. doi:10.1177/0022219415618500
- Phillips Galloway, E., Uccelli, P., Aguilar, G., and Barr, C. D. (2020). Exploring the Cross-Linguistic Contribution of Spanish and English Academic Language Skills to English Text Comprehension for Middle-Grade Dual Language Learners. *AERA Open* 6 (1), 233285841989257. doi:10.1177/2332858419892575
- Phillips Galloway, E., and Uccelli, P. (2019). Examining Developmental Relations between Core Academic Language Skills and reading Comprehension for

- English Learners and Their Peers. *J. Educ. Psychol.* 111 (1), 15–31. doi:10.1037/edu0000276
- Prevo, M. J. L., Malda, M., Mesman, J., and van IJzendoorn, M. H. (2016). Within- and Cross-Language Relations between Oral Language Proficiency and School Outcomes in Bilingual Children with an Immigrant Background. *Rev. Educ. Res.* 86 (1), 237–276. doi:10.3102/0034654315584685
- Proctor, C. P., August, D., Snow, C., and Barr, C. D. (2010). The Interdependence Continuum: A Perspective on the Nature of Spanish-English Bilingual Reading Comprehension. *Bilingual Res. J.* 33 (1), 5–20. doi:10.1080/15235881003733209
- Proctor, C. P., Carlo, M., August, D., and Snow, C. (2005). Native Spanish-speaking Children reading in English: Toward a Model of Comprehension. *J. Educ. Psychol.* 97 (2), 246–256. doi:10.1037/0022-0663.97.2.246
- Proctor, C. P., Harring, J. R., and Silverman, R. D. (2017). Linguistic Interdependence between Spanish Language and English Language and reading: A Longitudinal Exploration from Second through Fifth Grade. *Bilingual Res. J.* 40 (4), 372–391. doi:10.1080/15235882.2017.1383949
- Proctor, C. P., Silverman, R. D., Harring, J. R., and Montecillo, C. (2012). The Role of Vocabulary Depth in Predicting reading Comprehension Among English Monolingual and Spanish-English Bilingual Children in Elementary School. *Read. Writ* 25 (7), 1635–1664. doi:10.1007/s11145-011-9336-5
- Quinn, J. M., and Wagner, R. K. (2018). Using Meta-Analytic Structural Equation Modeling to Study Developmental Change in Relations between Language and Literacy. *Child. Dev.* 89 (6), 1956–1969. doi:10.1111/cdev.13049
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*; 2015.
- Rayner, Keith. (2009). Eye Movements in reading: Models and Data. *J. Eye Movement Res.* 2 (5), 1. doi:10.16910/jemr.2.5.2
- Rayner, Keith. (2012). *Eye Movements in reading: Perceptual and Language Processes*. Elsevier.
- Rayner, K. (1998). Eye Movements in reading and Information Processing: 20 Years of Research. *Psychol. Bull.* 124 (3), 372–422. doi:10.1037/0033-2909.124.3.372
- Rayner, K., and Raney, G. E. (1996). Eye Movement Control in reading and Visual Search: Effects of Word Frequency. *Psychon. Bull. Rev.* 3 (2), 245–248. doi:10.3758/bf03212426
- Reichle, E. D., Liversedge, S. P., Drieghe, D., Blythe, H. I., Joseph, H. S. S. L., White, S. J., et al. (2013). Using E-Z Reader to Examine the Concurrent Development of Eye-Movement Control and Reading Skill. *Dev. Rev.* 33 (2), 110–149. doi:10.1016/j.dr.2013.03.001
- Relyea, J. E., and Amendum, S. J. (2020). English Reading Growth in Spanish-Speaking Bilingual Students: Moderating Effect of English Proficiency on Cross-Linguistic Influence. *Child Dev.* 91 (4), 1150–1165. doi:10.1111/cdev.13288
- Reynolds, C. R., and Voress, J. K. (2007). *TOMAL-2 Test of Memory and Learning*. Pro-Ed.& Pro-ed (Firm).
- Rigby, R. A., and Stasinopoulos, D. M. (2005). Generalized Additive Models for Location, Scale and Shape (With Discussion). *J. R. Stat. Soc C* 54, 507–554. doi:10.1111/j.1467-9876.2005.00510.x
- Rossi, E., Krass, K., and Kootstra, G. J. (2019). Psycholinguistic Methods in Multilingual Research. *The Handbook Neurosci. Multilingualism*, 75–99. doi:10.1002/9781119387725.ch4
- Scott, C. M., and Balthazar, C. H. (2010). The Grammar of Information: Challenges for Older Students with Language Impairments. *Topics in Language Disorders* 30 (4), 288–307. doi:10.1097/TLD.0b013e3181f90878
- Siu, C. T.-S., and Ho, C. S.-H. (2015). Cross-language Transfer of Syntactic Skills and reading Comprehension Among Young Cantonese-English Bilingual Students. *Read. Res. Q.* 50 (3), 313–336. doi:10.1002/rrq.101
- Siu, C. T.-S., and Ho, S.-H. (2020). A Longitudinal Investigation of Syntactic Awareness and reading Comprehension in Chinese-English Bilingual Children. *Learn. Instruction* 67, 101327. doi:10.1016/j.learninstruc.2020.101327
- Snell, J., and Theeuwes, J. (2020). A story about Statistical Learning in a story: Regularities Impact Eye Movements during Book reading. *J. Mem. Lang.* 113, 104127. doi:10.1016/j.jml.2020.104127
- Sohail, J. (2015). The within- and Cross-Language Role of Syntactic Awareness in reading Comprehension Among French Immersion Students. PhD Thesis. Available at: <http://search.proquest.com.ezp-prod1.hul.harvard.edu/docview/1753920158?accountid=11311>.
- Sorenson Duncan, T., and Paradis, J. (2020). How Does Maternal Education Influence the Linguistic Environment Supporting Bilingual Language Development in Child Second Language Learners of English?. *Int. J. Bilingualism* 24 (1), 46–61. doi:10.1177/1367006918768366
- SR Research Ltd (). *EyeLink and Experiment Builder*. Ontario, Canada: SR Research Ltd. Mississauga.
- Stasinopoulos, M. D., Rigby, R. A., Heller, G. Z., Voudouris, V., and De Bastiani, F. (2017). *Flexible Regression and Smoothing: Using GAMLSS in R*. Chapman and Hall/CRC. doi:10.1201/b21973
- Swanson, H. L., Rosston, K., Gerber, M., and Solari, E. (2008). Influence of Oral Language and Phonological Awareness on Children’s Bilingual reading. *J. Sch. Psychol.* 46 (4), 413–429. doi:10.1016/j.jsp.2007.07.002
- Taylor, W. L. (1953). “Cloze Procedure”: A New Tool for Measuring Readability. *Journalism Q.* 30 (4), 415–433. doi:10.1177/107769905303000401
- Tiffin-Richards, S. P., and Schroeder, S. (2018). The Development of Wrap-Up Processes in Text reading: A Study of Children’s Eye Movements. *J. Exp. Psychol. Learn. Mem. Cogn.* 44 (7), 1051–1063. doi:10.1037/xlm0000506
- Torgesen, J., Wagner, R., and Rashotte, C. (2011). in *TOWRE-2*. Pro-Ed (Austin, TX: Test of Word Reading Efficiency).
- Trace, J. (2020). Clozing the gap: How Far Do Cloze Items Measure?. *Lang. Test.* 37 (2), 235–253. doi:10.1177/0265532219888617
- Tremblay, A. (2011). Proficiency Assessment Standards in Second Language Acquisition Research. *Stud. Second Lang. Acquis* 33 (3), 339–372. doi:10.1017/S0272263111000015
- Uccelli, P., Barr, C. D., Dobbs, C. L., Galloway, E. P., Meneses, A., and Sánchez, E. (2015a). Core Academic Language Skills: An Expanded Operational Construct and a Novel Instrument to Chart School-Relevant Language Proficiency in Preadolescent and Adolescent Learners. *Appl. Psycholinguistics* 36 (5), 1077–1109. doi:10.1017/S014271641400006X
- Uccelli, P., Galloway, E. P., Barr, C. D., Meneses, A., and Dobbs, C. L. (2015b). Beyond Vocabulary: Exploring Cross-Disciplinary Academic-Language Proficiency and its Association with reading Comprehension. *Read. Res. Q.* 50 (3), 337–356. doi:10.1002/rrq.104
- van den Bosch, L. J., Segers, E., and Verhoeven, L. (2018). Online Processing of Causal Relations in Beginning First and Second Language Readers. *Learn. Individual Differences* 61, 59–67. doi:10.1016/j.lindif.2017.11.007
- Van der Loo, M. P. J. (2014). The Stringdist Package for Approximate String Matching. *R. J.* 6 (1), 111–122. doi:10.32614/rj-2014-011
- Vanderwal, T., Eilbott, J., and Castellanos, F. X. (2019). Movies in the Magnet: Naturalistic Paradigms in Developmental Functional Neuroimaging. *Dev. Cogn. Neurosci.* 36, 100600. doi:10.1016/j.dcn.2018.10.004
- Van Dijk, T. A., and Kintsch, W. (1983). *Strategies of Discourse Comprehension*. Academic Press.
- Verhoeven, L., and van Leeuwe, J. (2009). Modeling the Growth of Word-Decoding Skills: Evidence from Dutch. *Scientific Stud. Reading* 13 (3), 205–223. doi:10.1080/10888430902851356
- von der Malsburg, T., and Angele, B. (2017). False Positives and Other Statistical Errors in Standard Analyses of Eye Movements in reading. *J. Mem. Lang.* 94, 119–133. doi:10.1016/j.jml.2016.10.003
- Whitford, V., and Joannis, M. F. (2018). Do eye Movements Reveal Differences between Monolingual and Bilingual Children’s First-Language and Second-Language reading? A Focus on Word Frequency Effects. *J. Exp. Child Psychol.* 173, 318–337. doi:10.1016/j.jecp.2018.03.014
- Whitford, V., and Titone, D. (2015). Second-language Experience Modulates Eye Movements during First- and Second-Language Sentence reading: Evidence from a Gaze-Contingent Moving Window Paradigm. *J. Exp. Psychol. Learn. Mem. Cogn.* 41 (4), 1118–1129. doi:10.1037/xlm0000093
- Whitford, V., and Titone, D. (2012). Second-language Experience Modulates First- and Second-Language Word Frequency Effects: Evidence from Eye Movement



- Measures of Natural Paragraph reading. *Psychon. Bull. Rev.* 19 (1), 73–80. doi:10.3758/s13423-011-0179-5
- Whitford, V., and Titone, D. (2014). The Effects of Reading Comprehension and Launch Site on Frequency-Predictability Interactions during Paragraph Reading. *Q. J. Exp. Psychol.* 67 (6), 1151–1165. doi:10.1080/17470218.2013.848216
- Whitford, V., and Titone, D. (2017). The Effects of Word Frequency and Word Predictability during First- and Second-Language Paragraph reading in Bilingual Older and Younger Adults. *Psychol. Aging* 32 (2), 158–177. doi:10.1037/pag0000151
- Wolf, M., and Denckla, M. B. (2005). *RAN/RAS: Rapid Automatized Naming and Rapid Alternating Stimulus Tests*. Pro-ed. Austin, TX.
- Woodcock, R. W., Muñoz-Sandoval, A. F., Ruef, M. L., and Alvarado, C. G. (2005). *Woodcock-Muñoz Language Survey-Revised*. Itasca, IL: Riverside.
- Zeileis, A., Wiel, M. A., Hornik, K., and Hothorn, T. (2008). Implementing a Class of Permutation Tests: The coin Package. *J. Stat. Softw.* 28 (8), 1–23.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Leon Guerrero, Whitford, Mesite and Luk. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.