ASHA American Speech-Language-Hearing Association

LSHSS

Research Article

La Rana or *El Rana*: Dual Language Learners' Grammatical Variability in Narrative Retells

Svenja Gusewski[a] and **Raúl Rojas**[b]

[a] Department of Communication Disorders, Southern Connecticut State University, New Haven, [b] Department of Speech-Language-Hearing: Sciences and Disorders, The University of Kansas, Lawrence

ARTICLE INFO

Article History:
Received December 31, 2023
Revision received March 1, 2024
Accepted April 3, 2024

Editor-in-Chief: Kelly Farquharson
Editor: Karla Nadine Washington

https://doi.org/10.1044/2024_LSHSS-23-00202

ABSTRACT

Purpose: This longitudinal study investigated the trajectory of Spanish article accuracy in Spanish–English dual language learners (DLLs) from preschool to first grade, addressing the need for longitudinal data on the variability of Spanish grammatical skills in DLLs in English immersion classrooms.
Method: Language sample analysis was conducted on 336 Spanish and English narrative retells elicited from 31 Spanish–English DLLs (range: 45–85 months). Growth curve models captured within- and between-individual change in article accuracy from the beginning of preschool to the end of first grade.
Results: As a group, DLLs did not exhibit significant positive or negative growth in Spanish article accuracy over time. On average, article accuracy remained stable at 76% from preschool throughout first grade. Participants exhibited significant variability in article accuracy that was partly explained by changes in Spanish proficiency. Spanish article accuracy was lower for DLLs with lower Spanish proficiency indexed by measures from the Spanish language samples, while English proficiency indexed by the English language samples did not affect Spanish article accuracy.
Conclusions: These findings suggest that expectations for Spanish grammatical performance in DLLs need to be adjusted to account for the possible impact of not receiving Spanish support in English immersion school settings. DLLs in these instructional programs do not exhibit article accuracy at a level expected for monolingual Spanish speakers. Significant individual differences in both individual status and growth rates of Spanish article accuracy highlight the broad variability in Spanish language skills of DLLs in the United States.

The development of language is a dynamic process, irrespective of the language or languages being acquired. For monolingual children, these dynamic changes occur within linguistic subdomains, while for dual language learners (DLLs) these dynamic changes also occur across languages. Specifically in DLLs, each of their languages undergo changes in order to meet the varying communicative demands and needs of the child, which are largely dependent upon environmental factors. Language dominance is operationalized as relative level of proficiency or frequency of use. Traditional frameworks determined language dominance based on the chronological order of acquisition, thus assuming a hierarchical relationship between first language (L1) and second language (L2). Yet, when a child's exposure to the L1 decreases and L2 exposure increases, a gradual shift to L2 dominance can occur. Consequently, the L1 might not continue to be the child's more proficient, and therefore dominant language. For example, Bedore et al. (2012) found that measures of current language use were better indicators of language dominance than age of first exposure. Their study also suggested that the classification of dominance depends on the linguistic domain assessed due to differences in semantic and morphosyntactic development. In sum, dual language development includes both acquisition and attrition processes since growth of one language can result in continued growth of the other language, or alternatively in its

Correspondence to Svenja Gusewski: gusewskis1@southernct.edu.
Disclosure: *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

decline (Anderson, 1999; Castilla-Earls et al., 2016; Davison et al., 2011; Herdina & Jessner, 2013; Hiebert & Rojas, 2021; Montrul, 2008).

Traditional methods in dual language development research were based on static measures for the first language (L1) and second language (L2; e.g., Genesee, 1989; Weinreich, 1968). These methods oftentimes relied on static measures at two points in time, which failed to account for the considerable fluctuation and variation that occurs during dual language development. Furthermore, traditional methods regarded native-like proficiency as the yardstick for all languages of a dual language learner (see Grosjean, 1985, for a discussion). However, recent work has shown that DLLs differ from monolinguals both in their language development and in other cognitive skills (e.g., executive functions; Bialystok, 2001; Bialystok et al., 2010; Cook, 2003).

In order to more precisely characterize early dual language growth and contribute important developmental data that address the broad variability in DLLs' language skills, the current study aims to characterize dual language development from a functional, least biased, and longitudinal perspective. Moreover, this study captures a crucial developmental and transitional period by examining dual language development from the start of systematic academic exposure to English in preschool and into the early years of elementary school. Thus, the current study aims to identify typical grammatical acquisition trajectories in DLLs who are simultaneously (a) in the process of acquiring English, (b) in the process of still acquiring Spanish, (c) restructuring their dual language systems, and (d) potentially beginning to undergo gradual attrition of some of their Spanish skills due to the subtractive nature of being academically instructed exclusively in English. This last point relates to research findings suggesting that systematic, academic exposure to an L2 in school without L1 language support may lead to protracted development, incomplete acquisition, or even attrition of the first language (Herdina & Jessner, 2013; Jacobson, 2012; Montrul, 2008).

This study focuses on grammatical skills because specific grammatical difficulties have been documented in typical dual language acquisition, yet are also considered a robust indicator of language impairment (LI; Bedore & Leonard, 1998; Bedore et al., 2010; Jackson-Maldonado & Maldonado, 2017; Leonard, 2016; Paradis, 2005). Furthermore, language dominance shifts in the morphosyntactic domain seem to occur later in development compared to shifts in semantic skills (Bedore et al., 2012) and are, therefore, more indicative of a complete dominance shift. It is important to note that while the current study aims to characterize developmental grammatical patterns in English and Spanish in young DLLs, its focus is not to develop clinical markers for LI in this population. Rather, this study aims to provide clinical researchers and clinicians with information on the expected range of performance and typical variability as a first necessary step to distinguish language disorders from expected variability within dual language acquisition. We capture variability in dual language production over time rather than relying on estimates based on language input through pre-established language dominance groups.

Several studies have suggested that articles are one of the grammatical elements during Spanish language acquisition that are often produced incorrectly by monolingual Spanish-speaking children with LI as well as Spanish–English speaking DLLs (Anderson & Márquez, 2009; Anderson & Souto, 2005; Bedore & Leonard, 2001, 2005; Castilla-Earls et al., 2016; Castilla-Earls, Auza, et al., 2020; Gutiérrez-Clellen et al., 2006; Jackson-Maldonado & Maldonado, 2017; Jacobson & Walden, 2013; Morgan et al., 2009, 2013; Restrepo, 1998; Restrepo & Gutiérrez-Clellen, 2001; Simon-Cereijido & Gutiérrez-Clellen, 2007). Furthermore, article use seems to be challenging for children with LI across a wide range of languages, including English and Spanish (Leonard, 2016). For example, several studies have shown difficulties in article use in children with LI who are learning Italian (e.g., Bortolini et al., 1997, 2002; Bottari et al., 1998, 2001; Leonard & Bortolini, 1998; Leonard et al., 1992). Italian, like Spanish, is a highly inflected Romance language and requires article–noun agreement in gender and number.

In Spanish, articles are marked for gender and number to agree with the noun (see Table 1 for masculine and feminine singular and plural articles in Spanish). Almost all masculine nouns end in *–o* (such as *el perro* [the dog], *el niño* [the boy]) and most feminine nouns end in *–a* (e.g., *la falda* [the skirt], *la niña* [the girl]). Nouns that end in *–e* can be either feminine or masculine (e.g., *el coche* [the car], *la noche* [the night]). Definite articles are high-frequency words in Spanish, according to EsPal, a Spanish lexical database (Duchon et al., 2013). Despite their high frequency and their obligatory status for most noun

Table 1. Indefinite and definite articles in Spanish.

Article	Masculine singular	Feminine singular	Masculine plural	Feminine plural
Indefinite	*Un*	*Una*	*Unos*	*Unas*
Definite	*El*	*La*	*Los*	*Las*

phrases (NPs), their saliency is low because they are unstressed monosyllabic words (Mariscal, 2009).

Articles emerge early in monolingual and dual language acquisition: monolingual Spanish-speaking children produce articles correctly by the age of 2 or 3 years (Baron et al., 2022), whereas Spanish–English DLLs have been shown to accurately produce articles by the age of 5 years (Bedore & Leonard, 2005; Castilla-Earls et al., 2016; Gutiérrez-Clellen et al., 2006; Hernández-Pina, 1984; Kernan & Blount, 1966; Pérez-Leroux et al., 2012; Pérez-Pereira, 1989; Simon-Cereijido & Gutiérrez-Clellen, 2007).

However, expectations for mastery of article production in DLLs are unclear due to (a) the lack of longitudinal studies that capture dynamic changes in dual language development and (b) the impact of changing communicative needs. Stated differently, systematic academic exposure to English without Spanish language support can lead to growth in one language and decline in the other (Herdina & Jessner, 2013). Specifically, grammatical elements such as articles are vulnerable to language attrition (Anderson, 1999; Maher, 1991) and bilingual effects related to shifts in language proficiency (Castilla-Earls et al., 2019). Furthermore, DLLs have been found to show differences in the acquisition of grammatical gender when compared to monolingual Spanish speakers not only just in terms of delayed acquisition but also in terms of differences in internal representations of gender features (Cuza & Pérez-Tattam, 2016). For example, a recent eye-tracking study found that Spanish–English bilingual school-aged children (ages 5;6 to 8;6 [years; months]) did not use grammatical gender information to actively anticipate the upcoming noun (Baron et al., 2022). Monolingual child and adult speakers from a variety of first language backgrounds, however, use grammatical gender to facilitate syntactic processing of the upcoming noun (e.g., Akhutina et al., 1999; Bates et al., 1996; Dahan et al., 2000; Dussias et al., 2013; Lew-Williams & Fernald, 2007, 2010; van Heugten & Shi, 2009). In sum, bilingual Spanish–English children show differences from monolingual children both in the processing and production of grammatical gender.

Article Accuracy in Monolingual Spanish-Speaking Children

Three studies have examined article use in monolingual Spanish-speaking children, and one study compared monolingual Spanish-speaking children to Spanish-speaking DLLs, contrasting children with and without LI using cross-sectional designs (see Table 2 for an overview). Anderson and Souto (2005) examined the article use of monolingual Spanish-speaking children with LI (M_{age} = 58 months) and age-matched children with typical language development (TD) across three combined language samples (picture description, narrative retell, and play sample). They found a significant difference in article accuracy between groups (TD = 98.4%, LI = 85.4%), with the LI group's scores ranging from 56.4% to 98.4% and the TD group's scores ranging from 95.3% to 100%. Similarly, Castilla-Earls, Auza, et al. (2020) found significant differences in article accuracy comparing monolingual Spanish-speaking children with LI (M_{age} = 65 months) to monolingual Spanish speakers with TD (M_{age} = 68 months). Spanish speakers with typical language produced 96% of articles correctly, while

Table 2. Means and standard deviations for percent accuracy of article use across monolingual Spanish-speaking children and DLLs with TD and LI.

Study	*n*	Age (years; months)	Monolingual Spanish			Spanish DLL		
			TD	LI	Effect size *d*	TD	LI	Effect size *d*
Anderson & Souto (2005)	11	4;3–5;4	98% (1.5)	84% (15)	1.3			
Baron et al. (2018)	228	4;0–7;6				33%–76%[c] 53–88%[a]		
Bedore & Leonard (2001)	15	3;1–5;6				80 (13.5)	47 (21.5)	1.8
Bedore & Leonard (2005)	15	3;1–5;6				96% (5)	84% (16)	1.1
Castilla-Earls et al. (2016)	16	4;7–8;4	86% (21)[a]	38% (22)[a]	2.2	78% (16)[b]	58% (23)[b]	0.6
Castilla-Earls, Auza, et al. (2020)	66	4;0–6;11				69% (26)	26% (25)	1.7
Castilla-Earls, Auza, et al. (2020)	50	4;0–6;11	96% (11)	70% (22)	1.5			
Morgan et al. (2013)	57	5;5–6;6	81% (15)	54% (28)	1.2	69% (23)	46% (22)	1.0
Restrepo & Gutiérrez-Clellen (2001)	30	5;0–7;1				97% (nr)	78% (nr)	
Simon-Cereijido & Gutiérrez-Clellen (2007)	38	4;7				96% (6)	79% (22)	1.1

Note. DLL = dual language learner; TD = typical language development; LI = language impairment; *d* = Cohen's *d*; nr = not reported.
[a]Spanish-dominant DLLs. [b]Balanced DLLs. [c]English dominant DLLs.

Spanish speakers with LI produced 70% of articles correctly in an elicitation task.

Jackson-Maldonado and Maldonado (2017) examined error types (substitutions, omissions, and additions) in article production in spontaneous narratives. They found that monolingual Spanish-speaking children with LI (n = 18; M_{age} = 7;3) showed significant differences in their correct use of articles compared to age-matched controls with TD. Two limitations of this study were addressed in the current study. First, Jackson-Maldonado and Maldonado (2017) examined raw counts of error types but did not control for obligatory contexts. A measure of accuracy is a more reliable estimate of a child's grammatical skills because it considers the proportion of correct use compared to the total number of obligatory contexts for the grammatical target form. Furthermore, the current study used narrative retells instead of spontaneous narratives to increase comparability across participants in terms of narrative content and length, which in turn affects the number of obligatory contexts (Gusewski et al., 2019).

Morgan et al. (2013) compared the article use of monolingual Spanish-speaking children (TD = 19, M_{age} = 72 months; LI = 9, M_{age} = 67 months) to Spanish–English DLLs (TD = 23, M_{age} = 72 months; LI = 7, M_{age} = 68 months) using an elicited task. They did not find significant group differences in article accuracy between monolingual Spanish-speaking children with LI and Spanish–English DLLs with TD using an elicited task. None of the groups reached mean article accuracy at an 85%-mastery criterion (monolingual TD = 81% accuracy, DLL TD = 69% accuracy, monolingual LI = 54% accuracy, and DLL LI = 46% accuracy). Variability was considerable, with standard deviations ranging from 15% to 28%. These results highlight that the risk of misdiagnosis of a language disorder in DLLs is heightened, especially when compared to monolingual standards. Consequently, DLLs should be compared to their DLL peers to improve diagnostic accuracy (Castilla-Earls, Bedore, et al., 2020).

Article Accuracy in Spanish–English DLLs

A number of studies have examined article use in Spanish–English DLLs by comparing children with and without LI using cross-sectional designs (see Table 2 for an overview). These studies are further characterized by multiple group comparisons, which include comparisons between language status groups (LI vs. typical development), and/or additional comparisons by language proficiency (Spanish dominant vs. English dominant; Spanish monolingual vs. Spanish–English DLL). The sample size per group tends to be relatively small due to multiple preestablished group comparisons.

Castilla-Earls, Auza, et al. (2020) found significant differences in article accuracy comparing Spanish–English DLLs with LI (M_{age} = 58 months) to DLLs with TD (M_{age} = 65 months). DLLs with typical language produced 69% of articles correctly, while DLLs with LI produced 26% of articles correctly in an elicited task. Restrepo and Gutiérrez-Clellen (2001) also found significant group differences in article accuracy in Spanish–English DLLs with LI compared to DLLs with TD (TD = 97% accuracy, LI = 78% accuracy) obtained in a picture description, conversation, and narrative retell tasks from a sample of 30 Spanish-speaking DLLs (TD = 15, M_{age} = 6;1 in years; months; LI = 15; M_{age} = 5;9). Similarly, Simon-Cereijido and Gutiérrez-Clellen (2007) found that article accuracy in a narrative retell and a spontaneous narrative was significantly higher in DLLs with TD (96%) compared to DLLs with LI (79%) in a sample of 38 Spanish-speaking DLLs (LI = 19, M_{age} = 56 months; TD = 19, M_{age} = 54 months). Bedore and Leonard (2001) also found that Spanish-dominant with LI exhibited significantly lower accuracy scores for both indefinite and definite articles (46% and 48% accuracy) relative to their age-matched peers (78% and 83% accuracy) in an elicitation task. In a follow-up study, Bedore and Leonard (2005) analyzed the same grammatical elements in spontaneous speech samples elicited during play from the same participants. In spontaneous speech, DLLs with LI still exhibited lower accuracy on article use (84% compared to 96% for age-matched peers).

Castilla-Earls et al. (2016) compared article accuracy elicited from Spanish-dominant DLLs (TD = 8, LI = 8; M_{age} = 67 months) to balanced DLLs (TD = 8, LI= 8; M_{age} = 83 months). Their data revealed that differences in article accuracy rates between DLLs with LI and TD are more pronounced in Spanish-dominant DLLs (TD = 86% vs. LI = 38% accuracy) relative to balanced DLLs (TD = 78% vs. LI = 58% accuracy). On average, participants in the balanced group had similar proficiency levels in English and Spanish. However, it is unclear if some of these participants might have undergone attrition of their Spanish skills since the group classification was solely based on English proficiency levels. Furthermore, the balanced DLL group was on average 16 months older than the Spanish-dominant group, which might be related to shifting dominance profiles. Thus, the impact of language proficiency levels on article accuracy might be even more pronounced when comparing DLLs who are undergoing dramatic changes in their Spanish proficiency due to systematic academic exposure to English without Spanish support.

Gaps in the Current Evidence-Base

Several studies have shown that article use in Spanish is challenging, especially for children with LI. However, the

degree of difficulty of article use in DLLs remains unclear due to methodological differences between studies: half of the studies have used elicitation tasks (Baron et al., 2018; Bedore & Leonard, 2001; Castilla-Earls et al., 2016; Morgan et al., 2013), whereas the other half used more naturalistic approaches of functional language use such as language sample analysis (Bedore & Leonard, 2001; Restrepo & Gutiérrez-Clellen, 2001; Simon-Cereijido & Gutiérrez-Clellen, 2007). When eliciting plural articles in Spanish with closed-format questions such as, "*¿Qué animales tienen las orejas largas?*" [Which animals have long ears?] or "*¿Con qué se abre la puerta?*" [What is used to unlock the door?], it is pragmatically appropriate and grammatically correct to answer with a plural noun without the article. Thus, using approaches that require functional language use such as narrative retells can more precisely capture true plural omission errors, as well as serving as a key component of the converging evidence framework (Castilla-Earls, Bedore, et al., 2020) to improve diagnostic accuracy with DLLs.

One major limitation of the existing literature is that studies have specifically selected Spanish-speaking DLLs with very limited English proficiency (e.g., Bedore & Leonard, 2001, 2005; Simon-Cereijido & Gutiérrez-Clellen, 2007); however, dual language skills are not static and are likely to change with increased exposure to English in school settings (Anderson, 1999). In English, articles are not marked for gender and number, which might render article use in Spanish more challenging as children's communicative needs and skills in each language change (Baron et al., 2022). It is important to consider that with increased academic instruction time in English, DLLs might become more proficient in English and use Spanish less frequently, which might result in a gradual decline of Spanish language skills (Hiebert & Rojas, 2021).

Furthermore, most studies that attempted to address the impact of language proficiency on article accuracy in DLLs were conducted with pre-established group comparisons, which oftentimes fit participants into static language dominance and ability groups, and where sample sizes per group were relatively small. These cross-sectional studies have found differences in article accuracy across individuals with different levels of language dominance, yet the impact of changes in language proficiency on article accuracy within the same individuals has not been examined longitudinally. None of the existing studies have addressed the issue of article use from either a longitudinal framework or by taking changes in the children's developing English language skills into account. These studies fail to acknowledge that dual language abilities reside on a dynamic continuum and that dual language skills are highly variable (Su et al., 2022).

Current Study

The current study addressed these limitations by examining dual language skills in a sample of typically developing DLLs who span the bilingual continuum. The purpose of this study was to capture the variability of typical dual language skills through growth curve modeling (GCM), a data analysis technique that is adept at capturing and interpreting individual differences. The current study modeled growth trajectories of Spanish article accuracy, while controlling for changes in English and Spanish proficiency longitudinally. The following two research questions were examined: First, how does Spanish article accuracy change from the beginning of preschool to the end of first grade in a sample of Spanish–English DLLs enrolled in an English immersion program? Second, how do longitudinal changes in Spanish article accuracy differ across DLLs with different levels of English and Spanish proficiency?

Spanish article accuracy is expected to demonstrate relatively steep and significant growth at the beginning of preschool, which may plateau or even decrease over time due to the impact of English-only instruction on Spanish skills (Hiebert & Rojas, 2021; Morgan et al., 2013). Furthermore, changes in English proficiency, indexed by growth in English tense marking accuracy, are expected to have a significant and negative impact on the growth trajectory of Spanish article accuracy, which is in line with research findings highlighting the dependence of Spanish article accuracy on language dominance (Baron et al., 2018; Castilla-Earls et al., 2016; Castilla-Earls, Auza, et al., 2020; Morgan et al., 2013).

Method

Participants

The present longitudinal study modeled growth trajectories of Spanish article accuracy and English tense marking accuracy in the narrative retells of 31 Spanish–English DLLs with typical language (12 boys, 19 girls) from preschool to first grade. The participants were part of a larger ongoing longitudinal project that has been reported on previously (Gusewski & Rojas, 2017). Eight participants were excluded from the original sample of 39 participants (18 boys, 21 girls) because they did not meet the inclusionary criteria of the present study (see Analytic Approach section for more details on inclusionary criteria).

The participants' mean age at the beginning of the study was 45 months. The participants were recruited from a laboratory school with a high enrollment of children from low-income and ethnically diverse households (86% of students qualify for free or reduced lunch; 94% Latine; 4% African American). In this school, DLLs were

instructed in English exclusively, however, preschool classrooms were staffed with a bilingual aid or teacher as support if needed. According to a parent report about the children's home language environment (adapted from Francis et al., 2005), the participants represented the bilingual continuum: 52% were exposed to mostly Spanish at home, 29% were exposed to Spanish and English equally at home, and 19% were exposed to mostly English at home. It is important to note that all the participants were exposed to Spanish at home, as none of the parents reported using English exclusively. Eighty-seven percent of the mothers had completed high school. The children's typical language status was verified through the administration of a screener in English and Spanish in kindergarten (Bilingual English Spanish Oral Screener; Peña et al., 2008) as well as parent report.

Procedure

Language samples were elicited in a narrative retell task in which participants retold counterbalanced stories from the "A Boy, a Dog, and a Frog" series (Mayer, 1969, 1974, 1975a, 1975b) in English and Spanish each semester from the beginning of preschool (PreK3) to first grade. The examiner first told the child a scripted version of the story using the corresponding wordless picture storybook, and then asked the child to retell the same story using the same wordless picture storybook. For a more detailed overview of the narrative retell language sample procedure consult Miller et al. (2015), Gusewski and Rojas (2017), or Heilmann et al. (2016). To elicit a language sample in the target language, the examiner spoke only in the target language to the child from the onset of the retell task in that language. In other words, if the target language was Spanish, then the examiner conducted the retell task with the child entirely in Spanish. All examiners and trained transcribers possessed native to near-native oral and written language proficiency in English and Spanish. The retell samples were digitally recorded, orthographically transcribed, and coded using the Systematic Analysis of Language Transcripts 20 software (Miller & Iglesias, 2020). Twenty-five percent of the total number of transcripts were randomly selected for transcription agreement (word-by-word) and coding agreement (point-by-point) checks (Heilmann et al., 2008). Reliability was calculated by dividing the number of agreements by the sum of agreements and disagreements and then multiplying by 100. Interrater reliability was high for both transcription and coding accuracy in each target language (English $M_{TranscriptionAcc.}$ = 97%, SD = 3%; English $M_{CodingAcc.}$ = 93%, SD = 5%; Spanish $M_{TranscriptionAcc.}$ = 94%, SD = 9%; Spanish $M_{CodingAcc.}$ = 91%, SD = 10%). Additionally, 25% of narrative retells were selected for reliability checks of the study-specific coding system for article accuracy in Spanish and tense marking accuracy in English. Coding agreement for each category (correct use, incorrect use, and omissions) was calculated for English tense markers and Spanish articles. Coding agreement across raters was high ($M_{Articles\_Spanish}$ = 93%, SD = 16%; $M_{TenseMarkers\_English}$ = 90%, SD = 23%).

Coding System for Article Accuracy in Spanish

To systematically capture the accuracy of article use in the Spanish language samples, all NPs in which article use was obligatory were identified. Obligatory contexts for NPs include both complete utterances ("*El niño estaba durmiendo*" [The(masc.) boy(masc.) was sleeping]) and other utterances that contain an NP ("*Tiene un perro*" [He/she has a(masc.) dog(masc.)]), including isolated NPs in instances of labeling ("*Una rana*" [A(fem.) frog(fem.)]). NPs in obligatory contexts were coded as correct when the gender and number of the article and noun matched (e.g., "*El* niño" [The(masc.) boy(masc.)]), or coded as incorrect when the gender and/ or number of the article and noun did not match (e.g., "*El rana*" [The(masc.) frog(fem.)]).

Gender agreement errors where phonotactic constraints of Spanish obligate the use of a masculine article with a feminine noun (feminine nouns that begin with a stressed "*a*" are used with the masculine article "*el*" or "*un*," such as "*el agua*" [the(masc.) water(fem.)]), were not counted as incorrect but excluded from analysis instead, because they are not true gender agreement errors but errors of not applying exemptions to the rule.

Furthermore, instances of article omissions were coded as omitted (e.g., "*Agarró* ***la*** *rana*" [He/she grabbed ***the** frog]). Solely grammatical contexts where an article was obligatory were coded as omissions; in Spanish, plural forms can be used without articles while singular forms require an article ("*Había flores*" [There were flowers] vs. "*había una flor*" [there was one flower]).

NPs that contained an English article or noun (e.g., "*the* rana," "la *frog*") were excluded from the analysis of Spanish article accuracy. These code-switched NPs do not provide obligatory contexts for gender agreement between article and noun, which is required in Spanish, but not in English. However, these code-switched NPs were coded and analyzed as a separate variable because they represent the dynamic linguistic repertoire of bilinguals.

Instances when articles were produced with speech sound errors (omissions of consonants or vowel, and distortions of vowels) were excluded from analysis since the intended article production was not identifiable through the child's speech. For instances, if a child said "*a rana*," this

production could have been interpreted in at least two different ways and therefore would be excluded from analysis:

C A[CS] rana. Instance where the article is code-switched.

C la[EP:A] rana. Initial consonant deletion, child intended to produce *"La rana."*

Note. [EP] marks phonological errors, [CS] marks code-switching.

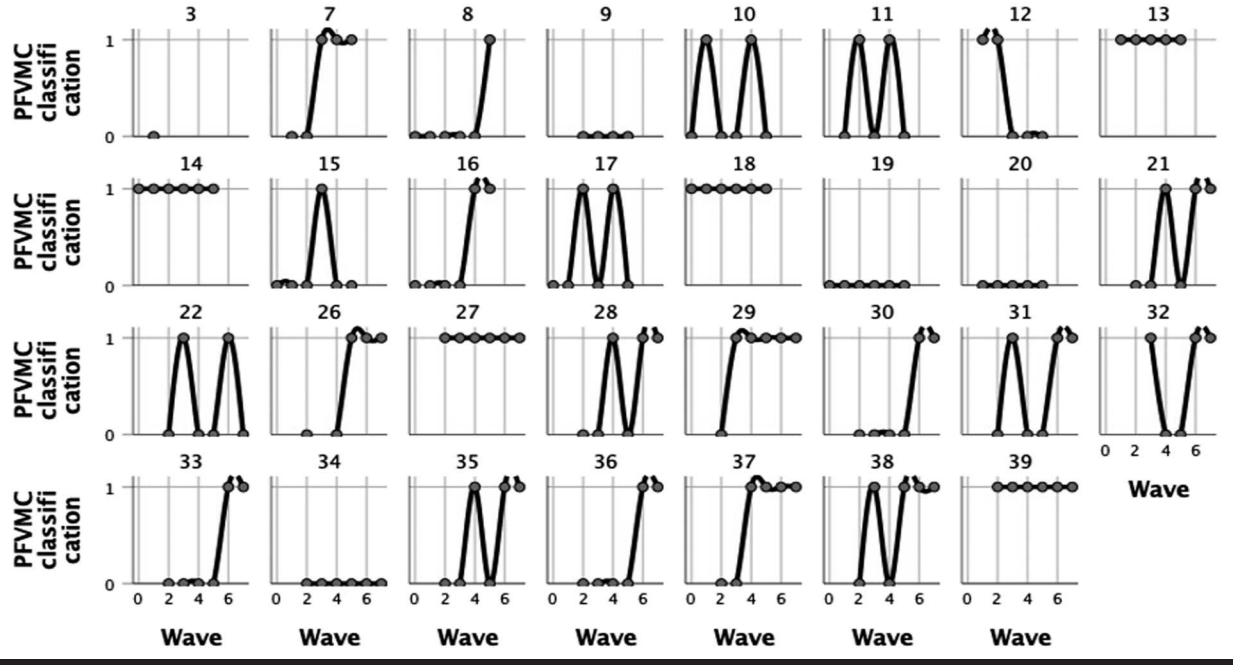
Outcome Measures

This study examined article accuracy growth while controlling for changes in English and Spanish proficiency. Thus, the variable of interest was article accuracy and the control variables were English tense-marking accuracy, the total number of obligatory contexts for articles in Spanish, and the proportion of code-switching within an NP in Spanish. The article accuracy composite score was calculated by dividing the total number of correctly used articles by the total number of obligatory contexts (correct, incorrect, and omissions) for articles. The following section provides operationalized definitions of all time-varying and time-invariant variables that were included in the growth curve models.

The Productive Finite Verb Morphology Composite (Productive FVMC; Gusewski & Rojas, 2017) measures tense-marking accuracy and was used to account for English proficiency as morphosyntactic skills have been found to index shifts in language dominance (Bedore et al., 2012). Morphosyntactic skills are more representative of attaining English proficiency because they tend to develop at a slower rate than vocabulary skills in DLLs (Bedore et al., 2012; Paradis, 2010). English tense-marking accuracy was used as a time-varying covariate to account for the impact of English proficiency on the growth of Spanish grammatical skills over time. To create a time-varying classification of high versus low English tense-marking skills, the longitudinal estimates of tense-marking accuracy provided in Gusewski and Rojas (2017) were used. For example, participants with a Productive FVMC of 65% at the beginning of preschool or a Productive FVMC of 90% at the beginning of kindergarten were classified with high English tense-marking skills at that specific time point. Figure 1 demonstrates that participants in this study changed in their classification of high versus low tense-marking accuracy. Thus, operationalizing English tense marking as a time-varying covariate captures dynamic changes in English proficiency.

Two variables were used to account for Spanish proficiency. The total number of obligatory contexts for articles was one of two variables used to account for Spanish proficiency by monitoring facility or difficulty in producing sufficient NPs in Spanish. This variable was treated as a time-invariant covariate at the onset of kindergarten.

Figure 1. Time-varying classification of low (0) versus high (1) English tense-marking accuracy Productive Finite Verb Morphology Composite (PFVMC) for individual participants from PK3 (Wave 0) to first grade (Wave 7). PK3 = participants who were enrolled in classrooms for 3-year-olds.



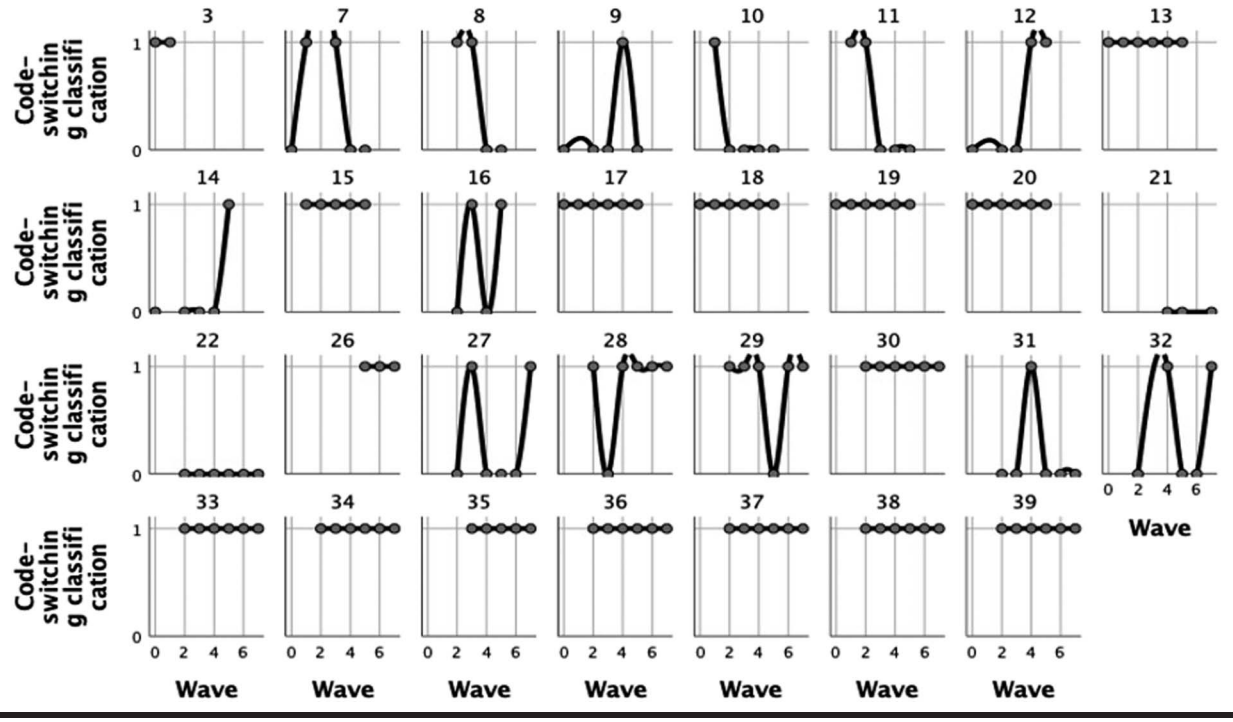
Specifically, participants who produced less than the grand mean of obligatory contexts (M = 26, across all semesters) at the beginning of kindergarten were classified as producing low obligatory contexts for articles in Spanish throughout any semester in this study (n_{low} = 14; n_{high} = 17). To illustrate, the story script for *Frog Goes to Dinner* in Spanish contains 91 NPs with articles, and an average of 4 NPs per picture. Thus, children who produced less than the group's grand mean of obligatory contexts for NPs (M = 26) demonstrated difficulty retelling the story by not sufficiently mentioning the characters and objects.

To account for children's full use of their dynamic linguistic repertoire, we did not set a specific criterion of words in the target language or exclude sentences that contained code-switching from analysis (see Analytic Approach section for more details). Rather, we attempted to capture code-switching as a natural process in DLLs that may be driven by language preference, or language proficiency (Gross & Castilla-Earls, 2023). As we captured code-switching specifically within an NP, meaning NPs that contained an English article or noun (e.g., "*the* rana," "la *frog"*), we operationalized these instances as indicative of language proficiency as they capture typical patterns of use when a child does not quickly access either the lexical item in the target language (e.g., "la *frog*"; Nicoladis & Secco, 2000), or the gender of the noun (e.g., "*the* rana"). Thus, the proportion of code-switching within an NP was used as a time-varying covariate to account for Spanish proficiency, as DLLs who trend toward English dominance are more likely to produce English words in their Spanish NPs, but not the inverse (Anderson, 1999; Guiberson et al., 2006). Specifically, DLLs with lower morphosyntactic skills in Spanish have been shown to use English words within Spanish utterances when retelling a story (Gross & Castilla-Earls, 2023). The classification cutoff of low versus high proportions of code-switching was set at 20%. The percent of code-switching within NPs was determined by dividing the number of code-switched NPs by the sum of the code-switched NPs and the total number of obligatory contexts for article accuracy. Figure 2 demonstrates that participants in this study changed in their classification of high versus low code-switching in NPs. Thus, operationalizing code-switching as a time-varying covariate captures dynamic changes in dual language proficiency.

Analytic Approach

Before analysis, the following inclusionary criteria were established in line with Bedore and Leonard (2005): accuracy rates were not calculated for time points at which children produced fewer than five obligatory contexts for article use in Spanish. Thus, when a participant did not produce sufficient

Figure 2. Time-varying classification of low (1) versus high (0) code-switching in noun phrases for individual participants from PK3 (Wave 0) to first grade (Wave 7). PK3 = participants who were enrolled in classrooms for 3-year-olds.



obligatory contexts to calculate reliable accuracy rates at a specific time point, this data point was set to a missing value. The same criterion was used for the English samples, where at least five obligatory contexts for tense marking were required for a sample to be analyzed. A total of eight out of 39 participants from the larger original data set did not meet this inclusion criteria because they did not provide at least two Spanish language samples with at least five obligatory contexts for Spanish articles during the six semesters. Therefore, the final sample size for this study was 31.

An exclusionary criterion based on the proportion of words produced in the target language was not established because we aimed at monitoring children's dynamic dual language skills; a stringent criterion such as 70% of words produced in the target language would have excluded children who produced enough obligatory contexts for article accuracy in Spanish while producing, for instance, 50% of words in English. To illustrate using a specific example: 50% of words produced in Spanish were deemed sufficient to analyze Spanish skills as long as the child produced at least five instances of articles in Spanish, which was defined as the presence of a Spanish noun that requires an article. After controlling for the presence of at least five obligatory contexts for articles, participants who did not have at least two language samples that provided sufficient obligatory contexts to calculate article accuracy scores were excluded from analysis.

This project used an accelerated cohort design by collecting language samples from 15 participants who were enrolled in classrooms for 3-year-olds (PK3) and 16 participants who were enrolled in classrooms for 4-year-olds (PK4) at the beginning of the project. This design allows for a planned missing data design protocol that spans eight total academic semesters (see Table 3); each of the 31 participants included in this current study completed narrative retell tasks in English and Spanish over six academic semesters. In other words, the children who began the study when they were in PK3 provided language samples from the beginning of PK3 to the end of kindergarten, while the slightly older cohort that started the study in PK4 provided language samples from the beginning of PK4 to the end of the first grade. The study analyzed a total of 172 English and 164 Spanish language samples over eight semesters, highlighting that some children may not have provided language samples in one or both languages during every wave.

As a preliminary step for GCM, unconditional growth curve models were estimated to find the best fitting growth curve shape for the conditional models. These models estimated the growth trajectory of Spanish article accuracy over eight semesters ranging from the beginning of preschool (PK3) to the end of first grade. Conditional GCM allows for the examination of the relative contributions of covariates on an outcome measure over time. Time-invariant covariates remain static while contributing to change in the outcome measure; time-varying covariates may change themselves, while contributing to change in the outcome measure. For the current study, the total number of obligatory contexts for articles in Spanish was used as a time-invariant predictor to account for Spanish proficiency. English tense-marking accuracy and the proportion of code-switching within an NP in Spanish were used as time-varying covariates to control for the impact of changes in English and Spanish proficiency on the growth of Spanish morphosyntactic skills over time. All statistical analyses were conducted in SPSS 26.0 (IBM Corp., 2018).

Results

The current study modeled growth trajectories of Spanish article accuracy, while controlling for changes in English and Spanish proficiency longitudinally. The first research question, "How does Spanish article accuracy change from the beginning of preschool to the end of first grade in a sample of Spanish–English DLLs enrolled in an English immersion program?", was addressed with unconditional growth curve models that estimated the growth trajectory of Spanish article accuracy from the beginning of preschool (PK3) to the end of the first grade. The second research question, "How do longitudinal changes in article accuracy differ across DLLs with different levels of English and Spanish proficiency?", was addressed in conditional growth curve models that estimated the growth trajectory of Spanish article accuracy, while controlling for English proficiency (indexed by English tense-marking accuracy in English language samples) as well as for Spanish proficiency (indexed by the proportion of code-switching within NPs and the total number of obligatory contexts for articles in the Spanish language samples). Descriptive information on participant and language sample characteristics can be found in Table 4.

Table 3. Number of participants per semester organized by grade (planned missing design).

Semester and grade	Fall PK3	Spring PK3	Fall PK4	Spring PK4	Fall K	Spring K	Fall 1st	Spring 1st
Number of participants	15	15	31	31	31	31	16	16

Note. PK3 = participants who were enrolled in classrooms for 3-year-olds; PK4 = participants who were enrolled in classrooms for 4-year-olds; K = kindergarten; 1st = first grade.

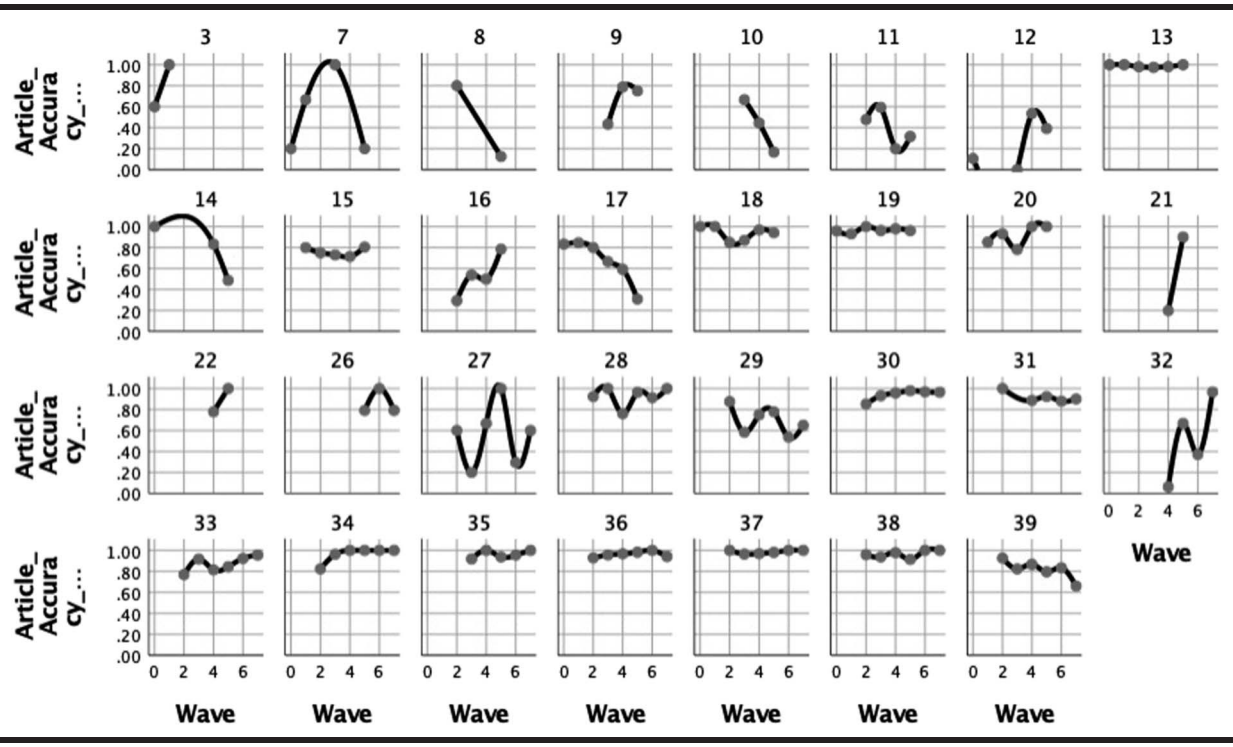
Table 4. Participant and language sample descriptive statistics.

Variable	PK3 – Fall	PK3 – Spring	PK4 – Fall	PK4 – Spring	K – Fall	K – Spring	1st – Fall	1st – Spring
Participants	10	10	31	31	31	31	21	21
Samples	8	11	28	27	29	30	15	16
Age (months)	45.2 (4.6)	51.27 (3.7)	56.9 (3.7)	61.4 (3.7)	68.7 (3.7)	73.2 (3.9)	80.8 (5.7)	85.3 (3.4)
English LSA measures								
PFVMC	53.9% (25.6)	67.8% (17.6)	71.0% (15.4)	82.7% (10.4)	85.4% (11.5)	87.7% (7.9)	92.4% (5.7)	93.6% (8.5)
Spanish LSA measures								
Article accuracy	71.2% (37.2)	88.7% (11.9)	82.7% (18.4)	75.7% (26.8)	74.8% (27.0)	75.6% (28.5)	83.4% (24.5)	88.8% (14.7)
OCs	9.9 (8.4)	13.2 (11.2)	18.6 (16.1)	21.7 (15.3)	28.1 (17.1)	30.8 (16.2)	32.8 (14.0)	41.6 (19.8)
%NPCS	19.3% (26.3)	1.1% (2.4)	25.9% (34.7)	21.4% (29.3)	20.3% (27.1)	21.6% (26.9)	13.8% (25.4)	19.0% (31.1)
NDW	38.0 (22.8)	47.7 (21.5)	60.5 (23.2)	63.0 (18.9)	68.8 (27.1)	62.9 (23.6)	74.1 (17.7)	85.2 (21.4)
MLUw	3.76 (1.36)	3.89 (1.23)	4.73 (1.30)	5.35 (1.08)	5.29 (1.71)	5.45 (1.51)	6.24 (0.97)	6.85 (0.71)
%G	63.1% (21.4)	73.5% (16.6)	72.8% (13.3)	68.1% (18.1)	64.9% (17.5)	70.5% (18.3)	68.3% (17.5)	66.8% (24.1)
Utt.	21.9 (16.0)	23.6 (9.2)	30.8 (9.5)	29.15 (9.6)	36.5 (13.6)	31.8 (8.4)	33.2 (6.6)	37.8 (9.4)

Note. All LSA measures are based on complete and intelligible utterances. PK3 = participants who were enrolled in classrooms for 3-year-olds; PK4 = participants who were enrolled in classrooms for 4-year-olds; K = kindergarten; 1st = first grade; LSA = language sample analysis; PFVMC = Productive Finite Verb Morphology Composite; OCs = obligatory contexts for articles; %NPCS = percent code-switching within noun phrases; NDW = number of different words; MLUw = mean length of utterances in words; %G = percent grammatical utterances; Utt. = total number of utterances.

Visual inspection of individual growth plots indicated a range of linear and curvilinear trajectories for article accuracy (see Figure 3), with comparable numbers of individual participants displaying patterns of positive growth (n = 6), and negative growth (n = 12). A subset of children seemed to display asymptotic growth patterns (n = 13), characterized by a high intercept that constrained how much growth was possible. In this subset, the trajectories of three children (#18, #31, and #36) seemed to decline toward the last semesters, while the other 10 trajectories remained relatively stable.

Figure 3. Smooth nonparametric trajectories superimposed on empirical growth plots for article accuracy from PK3 (Wave 0) to first grade (Wave 7); PK3 = participants who were enrolled in classrooms for 3-year-olds.



The first model fit to the data was an unconditional means model, which does not include the effect of time (see Model A in Table 5). It is used as a baseline model to partition outcome variance across two levels: within people over time (σ_ε^2) or between people (σ_0^2). This unconditional means model also reveals whether there is systematic variation in the outcome that is worth exploring. Table 5 (Model A) shows that participants' article accuracy significantly varied within individuals over time ($\sigma_\varepsilon^2 = .028$) and that participants' article accuracy differed significantly from each other ($\sigma_0^2 = .037$). Because these two variance components were significantly different from zero, both within-person and between-person differences in article accuracy can be linked to predictors in the conditional growth models. The unconditional means model also quantifies the amount of between-person variance through the intraclass correlation coefficient. In this case, 56% of the total outcome variance in article accuracy was located between people.

The unconditional growth model added time as a predictor (see Model B1 in Table 5). A series of unconditional growth curve models that differed in their specification of randomly varying slopes and in their specification of the functional form of the trajectory was estimated to find the best fitting unconditional growth curve model (see Table A1, Appendix). Assessing model fit for nested unconditional and conditional models is based on comparing the deviance statistic (−2LL), and or the change in deviance (Δ −2LL) between models (Singer & Willet, 2003). The linear model (Model B1) was the best fitting functional form for the growth trajectory of article accuracy because adding a quadratic term (Model C1) and a cubic term (Model D1) did not result in a significant change in deviance. Thus, Model B1 is displayed as unconditional growth model in Table 5 and was used as baseline model to build the conditional growth curve models.

Adding time as a predictor in the unconditional model did not improve model fit (Δ −2LL = 0.02, Δdf =

Table 5. Growth curve model parameter estimates: Spanish article accuracy.

Parameter	Term	Model A unconditional means	Model B1 unconditional growth	Model E1 conditional growth	Model E2 conditional growth	Model E3 conditional growth	Model F1 conditional growth
Fixed effects							
Intercept	γ00	76.18***	75.72***	66.21***	63.85***	79.37***	**52.15*** **
Linear slope	γ10	—	0.12	0.15	0.07	−0.47	**0.22**
High OC Spanish	γ01	—	—	16.69*			**12.84***
Low %NPCS Spanish	γ02	—	—	—	17.12***		**21.94****
High PFVMC English	γ03	—	—	—	—	−0.69	**−0.72**
Variance components							
L1: Within-person variance	σ_ε^2	.028***	.028***	.028***	.028***	.027***	**.020*** **
L2: Between-person intercept	σ_0^2	.037***	.037***	.029***	.024**	.034***	**.085***
L1: Within-person high PFVMC	σ_1^2	—	—	—	—	—	**.010**
L1: Within-person Low %NPCS	σ_2^2	—	—	—	—	—	**.066***
Covariance	σ_{01}	—	—	—	—	—	**−.025**
Covariance	σ_{02}	—	—	—	—	—	**−.063***
Covariance	σ_{12}	—	—	—	—	—	**.014**
Proportion variance reduction							
L1: Within-person variance	R_ε^2	—	0	0	0	.04	**.29**
L2: B/w-person intercept	R_0^2	—	—	.22	.35	.08	**N/A**
Goodness-of-fit	ICC = .56						
Deviance (−2LL)		−46.030	−46.050	−51.369	−58.589	−51.016	**−80.179**
AIC		−40.030	−38.050	−41.369	−48.589	−41.016	**−56.179**
BIC		−31.121	−26.171	−26.520	−33.740	−26.308	**−20.880**

Note. Models E1–E3 estimate the uncontrolled effects of obligatory contexts of articles, code-switching within noun phrases, and English tense-marking accuracy on Spanish article accuracy growth; Model F1 estimates the effect of high English tense-marking accuracy on Spanish article accuracy growth controlling for obligatory contexts of articles and code-switching in noun phrases. Bold formatting indicates best fitting models. OCs = obligatory contexts for articles; %NPCS = percent code-switching within noun phrases; PFVMC = Productive Finite Verb Morphology Composite; L1 = Level-1 variance; L2 = Level-2 variance; −2LL = −2 log-likelihood deviance statistic; ICC = intraclass correlation coefficient; AIC = Akaike information criterion; BIC = Bayesian information criterion.

*$p < .05$. **$p < .01$. ***$p < .001$.

1, p = .89), which is emphasized by the nonsignificant slope parameter (γ_{10} = 0.12, p = .89). Comparing the Level 1 residual variance (σ_ε^2) of model A to model B1 shows no reduction in variance, meaning that variance in article accuracy within individuals was not systematically associated with time. Thus, the prototypical growth trajectory for article accuracy from preschool to first grade was estimated as a relatively flat trajectory with an intercept at 75% (γ_{00} = 75.72, p < .001) and a nonsignificant slope (γ_{10} = 0.12, p = .888), indicating that for this group of DLLs, article accuracy remained relatively stable at 75% from preschool to first grade (see Figure 4).

Even though intraindividual variation was not significantly associated with time, both the intraindividual variance (σ_ε^2 = .028) and interindividual variance (σ_0^2 = .037) remained significant, indicating that time-varying and time-invariant predictors in the conditional growth curve models were warranted in order to explain these intra- and interindividual differences.

The first set of conditional growth curve models estimated the uncontrolled effects of individual predictors on the intercept of article accuracy. The set of time-varying predictors (low code-switching within Spanish NPs; high English tense-marking) aimed to examine the effect of changes in English and Spanish proficiency on Spanish article accuracy.

Model E1 estimated the time-invariant effect of producing a high number of obligatory contexts for articles on the intercept of article accuracy. Participants with a high number of obligatory contexts for articles were producing articles 17% more accurately (γ_{01} = 16.69, p = .022) than their peers who produced low levels of obligatory contexts.

Model E2 estimated the time-varying effect of producing a low proportion of code-switching within NPs on the intercept of article accuracy. Participants with a low proportion of code-switching within NPs were producing articles 17% more accurately (γ_{02} = 17.12, p < .001) than their peers who produced high levels of code-switching.

Model E3 estimated the time-varying effect of high English tense-marking accuracy on the intercept of Spanish article accuracy. Article accuracy in Spanish did not significantly differ between participants with high tense-marking accuracy in English compared to their peers who exhibited low tense-marking accuracy in English (γ_{03} = −0.69, p = .86).

The final model, Model F1, estimated the controlled effects of English tense-marking accuracy on the intercept of Spanish article accuracy. Controlling for the effect of producing a low proportion of code-switching within NPs and the effect of producing a high number of obligatory contexts for articles, article accuracy in Spanish did not differ between participants with high tense-marking accuracy in English compared to their peers who exhibited low tense-marking accuracy in English (γ_{03} = −0.72, p = .86).

However, article accuracy differed for DLLs who produced low proportions of code-switching within NPs, controlling for the effect of high tense-marking accuracy in English and the effect of producing a high number of obligatory contexts for articles. Participants with a low proportion of code-switching within NPs were producing articles 22% more accurately (γ_{02} = 21.94, p = .009) than their peers who produced high levels of code-switching (see Figure 5).

Article accuracy also differed for DLLs who produced a high number of obligatory contexts for articles, controlling for the effect high tense-marking accuracy in English and the effect of producing a low proportion of code-switching within NPs. Participants with a high number of obligatory contexts for articles were producing articles 13% more accurately (γ_{01} = 12.84, p = .032) than their peers who produced low levels of obligatory contexts.

Figure 4. Prototypical growth trajectories for article accuracy. PK3 = participants who were enrolled in classrooms for 3-year-olds; PK4 = participants who were enrolled in classrooms for 4-year-olds; K = kindergarten; 1 = first grade.

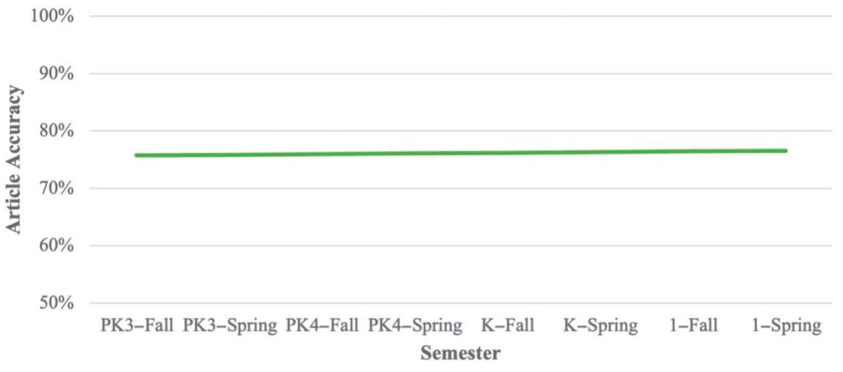
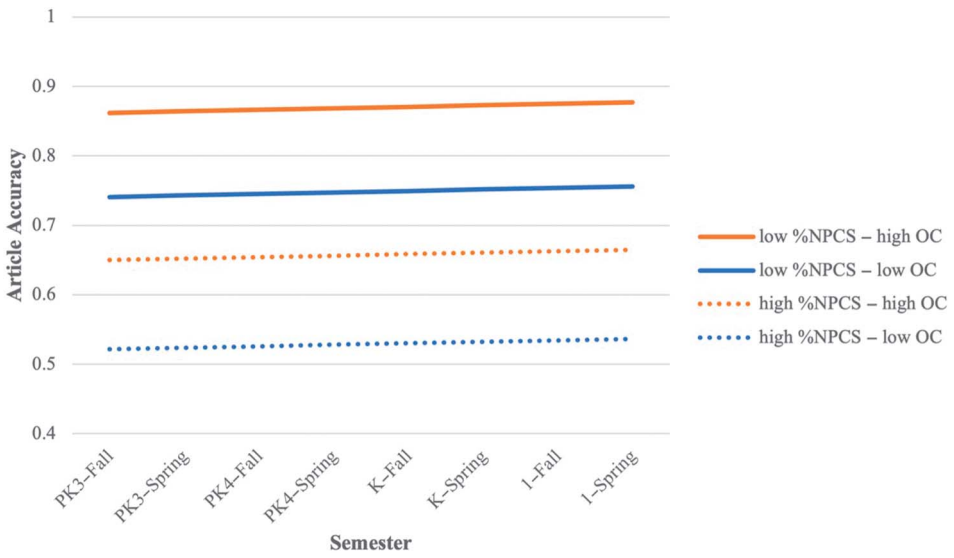


Figure 5. Growth trajectories of article accuracy controlling for English tense marking, code-switching within Spanish noun phrases (%NPCS), and obligatory contexts for articles in Spanish (OC). Tense marking has been omitted from the figure due to nonsignificance. PK3 = participants who were enrolled in classrooms for 3-year-olds; PK4 = participants who were enrolled in classrooms for 4-year-olds; K = kindergarten; 1 = first grade; OC = obligatory contexts for article; %NPCS = percent code-switching within noun phrases.



Compared to the unconditional growth model (Model B1), this conditional model (Model F1) provided a significantly better fit to the data (Δ −2LL = 34.13, Δdf = 8, $p < .001$). However, the final model still postulated a flat trajectory for article accuracy from preschool to first grade with an intercept at 52% (γ_{00} = 52.15, $p < .001$, γ_{10} = 0.22, $p = .799$). Thus, for DLLs who produced a high proportion of code-switched articles in Spanish, and a low number of obligatory contexts for Spanish articles, irrespective of their English tense-marking accuracy, article accuracy remained stable at 52% from preschool to first grade (see Figure 5).

Since the effect of time did not account systematically for intraindividual variance in article accuracy, a significant amount of intraindividual variance remains unexplained. Even though the prototypical growth trajectory for the group was estimated to be flat, clear differences in the direction and rate of growth are noticeable in the observed individual trajectories of participants (see Figure 3). This highlights the high variability in Spanish language skills in DLLs.

Discussion

The current study modeled growth trajectories of Spanish article accuracy while controlling for changes in English and Spanish proficiency longitudinally. The following research questions were addressed: (a) How does Spanish article accuracy change from the beginning of preschool to the end of first grade in a sample of Spanish–English DLLs enrolled in an English immersion program? (b) How do longitudinal changes in article accuracy differ across DLLs with different levels of English and Spanish proficiency?

The results indicated that article accuracy remained stable at 52% from preschool to the first grade for DLLs who produced low numbers of obligatory contexts for articles and high proportions of code-switching within NPs. Controlling for the impact of changes in English and Spanish proficiency on Spanish article accuracy showed that DLLs with high tense-marking accuracy in English did not differ in their Spanish article accuracy compared to their peers with low English tense-marking accuracy. However, DLLs who, in their Spanish language samples, exhibited low proportions of code-switching within NPs produced articles with 22% more accuracy than their peers with high proportions of code-switching within NPs. Similarly, DLLs with a high number of obligatory contexts for Spanish articles were producing articles with 13% more accuracy than their peers who produced low levels of obligatory contexts.

The fact that the effect of time did not account systematically for intraindividual variance in article accuracy, as demonstrated by flat growth trajectories, aligns with current research highlighting the maintenance of native language skills rather than true language loss. Thus, dual language learning is characterized by a gradual shift in language dominance linked to changes in language environments (i.e., bilingual effects; Castilla-Earls et al., 2019) rather than a zero-sum game in which growth in one language causes loss of skills in the other (Oppenheim et al., 2020). Consequently, the current study did not find a direct and significant

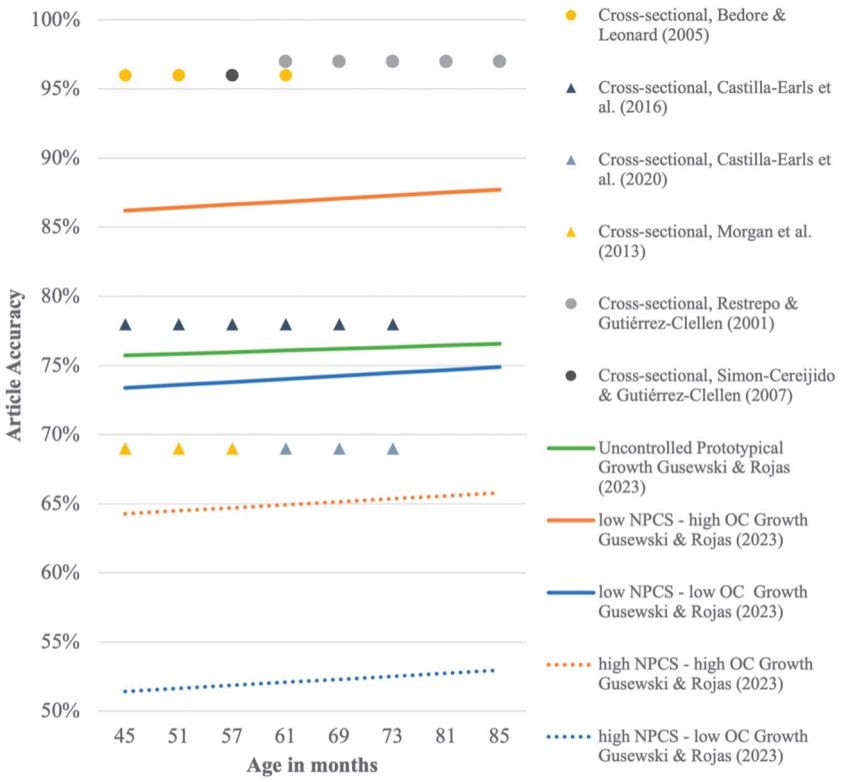
relationship between English grammatical skills (indexed by English tense marking) and Spanish grammatical skills (indexed by Spanish article accuracy).

Similarly, our findings highlighted a shift toward English dominance; DLLs who produced high proportions of code-switching within NPs produced Spanish articles with less accuracy. The frequent use of English in Spanish NPs indicates that these DLLs seem to be English-dominant and consequently Spanish is their less proficient language. These finding indicate that DLLs who seem to shift in their language dominance profiles toward English dominance might be showing signs of that dominance shift in their Spanish language samples by producing articles with less accuracy, frequently producing code-switching within an NP, and overall producing fewer obligatory contexts for Spanish articles. It is important to note that the lack of a direct relationship between English and Spanish morphosyntactic skills is in line with current research explaining dual language acquisition through an incremental learning account that is driven by experience-based changes (Oppenheim et al., 2020). The participants in this study exhibited steady growth in their English morphosyntactic skills (as reported in Gusewski & Rojas, 2017), which did not directly impact their Spanish morphosyntactic skills, but rather led to English dominance as indexed by the frequent use of English words in Spanish language samples (i.e., code-switching). Previous research has shown general increases of code-switching in the nondominant language as children get older and transition into an English-dominated environment (Sheng et al., 2013). Furthermore, DLLs also show more frequent within utterance code-switching to English in their Spanish narrative retells (but not the inverse) when their Spanish morphosyntactic skills are lower (Gross & Castilla-Earls, 2023). Thus, code-switching is a communicative strategy that bilingual children employ to meet changing environmental demands.

Clinical Implications

From a clinical perspective, these findings highlight two fundamental characteristics of typical dual language development for Spanish–English bilinguals in the United States: (a) a bilingual is not two monolinguals in one person (Grosjean, 1989), and (b) skills, especially in the native language, vary considerably between children as well as within the same child when measured at different times. Consequently, expectations about grammatical skills should be based on comparisons to other DLLs with similar language environments rather than based on norms from monolingual or Spanish-dominant children as highlighted in Figure 6.

Figure 6. Unconditional article accuracy trajectory compared to current literature. NPCS = code-switching within a noun phrase; OC = obligatory contexts for articles.



The Spanish article accuracy values displayed by the DLLs in this study are lower than expected based on findings from other studies (see Figure 6). This longitudinal study captured an ages range from 3 years, 9 months to 7 years 1 month, during which article accuracy remained relatively stable around 75%, without controlling for differences in English and Spanish proficiency. The most Spanish-proficient DLLs (who produced high obligatory contexts and low code-switching within NPs) in this study produced articles with 86% accuracy. Previous cross-sectional studies have found monolingual Spanish-speaking children's accuracy on articles to range from 95% to 100% between the ages of 4 years, 3 months and 5 years 3 months (Anderson & Souto, 2005). For Spanish–English DLLs, previous cross-sectional studies have found averages of article accuracy to range from 96% to 97% in language samples of DLLs between 3 years, 11 months up to 6 years 1 month (Bedore & Leonard, 2005; Restrepo & Gutiérrez-Clellen, 2001; Simon-Cereijido & Gutiérrez-Clellen, 2007).

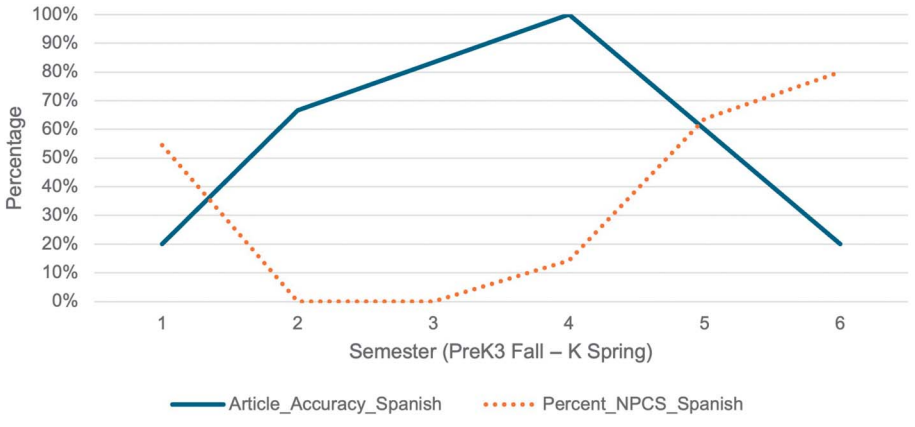
Apart from highlighting the impact of differences in Spanish proficiency driven by a gradual shift toward English dominance, the current study highlights that DLLs in an English immersion program might not reach "true" mastery of article production, considering that mastery is often defined as the ability to produce the target with at least 80% accuracy (Baron et al., 2018). Monolingual Spanish-speaking children have been found to reach mastery by age 4 years, while DLLs have been suggested to reach mastery around age 5 years (Bedore & Leonard, 2005; Castilla-Earls, Auza, et al., 2020; Hernández-Pina, 1984; Kernan & Blount, 1966; Pérez-Leroux et al., 2012; Pérez-Pereira, 1989; Simon-Cereijido & Gutiérrez-Clellen, 2007). As this current study shows, this expectation will have to be adjusted according to the children's language dominance profiles, which are highly dependent upon language exposure and use (Bedore et al., 2012).

To illustrate, the trajectory of Jaime (pseudonym) is offered as a case study to exemplify how dynamic DLL's dominance profiles are (see Figure 7). During semesters where he frequently used English within Spanish NPs, his Spanish article accuracy was 20% in Spanish NPs when he produced both the noun and article in Spanish (Semesters 1 and 6). During semesters where he used English within Spanish NPs less than 20% of the time, his article accuracy ranged from 70% to 100% (Semesters 2–4). Consequently, if a speech-language therapist collected a Spanish language sample as part of a comprehensive assessment, Jaime's performance should be interpreted considering his language dominance profile. Exhibiting low Spanish article accuracy while using strategies such as producing English articles or nouns in a Spanish NP should not be cause for concern, especially when using a converging evidence framework where other findings also point to English dominance (Castilla-Earls, Bedore, et al., 2020).

The findings of this study further highlight that bilinguals are not two monolinguals in one person. DLLs seem to process grammatical gender differently due to differences in internal representation of gender features (Cuza & Pérez-Tattam, 2016). While monolingual children actively use grammatical gender to facilitate processing (e.g., Akhutina et al., 1999; Bates et al., 1996; Dahan et al., 2000; Dussias et al., 2013; Lew-Williams & Fernald, 2007, 2010; van Heugten & Shi, 2009), bilingual school-aged children do not use grammatical gender information to actively anticipate the upcoming noun (Baron et al., 2022). In sum, DLLs process grammatical gender information differently, which may explain the high variability in their production of grammatical gender.

These lower-than-expected article accuracy values are in line with recent research that found differences in article accuracy based on language dominance. For example, Baron et al. (2018) reported article accuracy values

Figure 7. Case study highlighting the impact of high code-switching within noun phrases on article accuracy. PreK3 = preschool; K = kindergarten.



ranging from 33% to 76% for English-dominant DLLs and 53% to 88% for Spanish-dominant DLLs (age range: 4;0 to 7;6). Castilla-Earls et al. (2016) reported article accuracy values of 78% for balanced DLLs, while Spanish-dominant DLLs exhibited article accuracy at 86%. The balanced DLLs in Castilla-Earls et al. were on average 16 months older than their Spanish-dominant peers, indicating that dominance shifts are likely to occur as DLLs experience more exposure to English. However, cross-sectional studies cannot provide evidence for these longitudinal phenomena. Furthermore, cross-sectional studies cannot account for changes in language proficiency within the same individuals.

Limitations and Future Directions

Instead of grouping DLLs into static language dominance groups based on reported language use and exposure, the current study provided insights into language dominance shifts by examining performance-based measures of language proficiency. This study addressed a gap in the current literature by examining Spanish article accuracy from a longitudinal perspective, while controlling for dynamic changes in dual language development. DLLs in this study exhibited maintenance of Spanish grammatical skills, instead of growth or decline, which provided further evidence of an incremental learning account of dual language development that is driven by experience-based changes (Oppenheim et al., 2020). DLLs who exhibited maintenance of Spanish grammatical skills below an 80% mastery criterion were most likely most vulnerable to the effect of increased exposure to English in the academic setting without Spanish language support. Thus, for DLLs who were still in the process of developing Spanish grammatical skills and had not reached mastery yet, changes in the language environment seemed to impact Spanish grammatical development as their Spanish grammatical skills did not continue to grow as would be expected developmentally.

Future studies should examine whether this trend of maintenance of Spanish grammatical skills continues after first grade or whether DLLs' Spanish grammatical skills decline with increased time in an English immersion academic setting. A limitation of this current study is the high variability of Spanish skills within this relatively small sample, which impacted our ability to estimate significant growth curve parameters. For example, observed individual data points show great variability within participants and clear differences in the direction and rate of growth between individuals (see Figure 3). Future studies with larger sample sizes could extend the findings from this study in important ways by allowing the estimation of more complex growth curve models.

Another important consideration for future studies is a qualitative analysis of Spanish article production. Gender features might be represented differently in DLLs' lexicons compared to monolingual Spanish speakers (Cuza & Pérez-Tattam, 2016). Thus, when producing Spanish articles, which need to agree with the noun in an NP in gender and number, DLLs might show difficulties when their gender representation for that specific noun is weak in their lexicon. One expression of that weak representation of gender features could be the frequent use of code-switching within NPs to avoid gender marking. Another expression of this could be the use of the masculine form ("*el*") as default article: DLLs with weak gender representations might use a default article (e.g., "*el*") for every noun, independent of its gender. This behavior would lead to potentially high variability between article accuracy measured at different time points: high article accuracy for instances when most nouns were masculine, low article accuracy when most nouns were feminine. Consequently, a qualitative analysis of each individual DLL's performance would guide our understanding of individual differences in gender representation, which is especially important since previous research has shown that gender substitutions are the most common error pattern in DLLs with typical language (Morgan et al., 2013; Restrepo & Gutiérrez-Clellen, 2001).

Similarly, a qualitative analysis of article production in Spanish could shed more light on instances of code-switching within NPs and whether these instances are strategies employed to overcome lexical or morphosyntactic challenges or stylistic devices. We would hypothesize challenges of accessing the correct lexical item to be characterized by marked pauses (e.g., "*la* 0:03 frog"), repetitions of the article (e.g., "*la la la* frog"), or interjections ("*la* um ah frog"), which are typical disfluencies for DLLs (Rojas et al., 2023), while challenges of accessing the correct gender pronoun may be characterized by revisions (e.g., "*la el* the *rana*"). Code-switching as a stylistic device may be smoother and indicative of the child's understanding that the examiner understands both Spanish and English, which is a logical assumption a child may make in an English-speaking educational setting, even though the examiners only communicated in the target language (Spanish) with the child.

Data Availability Statement

The data sets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgments

This research has been supported in part by a University of Texas at Dallas Dissertation Research Award, a

Graduate Student Scholarship from the American Speech-Language-Hearing Foundation supported by the Kala Singh Memorial Fund, and a Southern Connecticut State University Reassigned Time Scholarship (awarded to Svenja Gusewski); by Grant GA-2013-016–Jerry M. Lewis, M.D. Mental Health Research Foundation (awarded to Raúl Rojas); and Grant 13180 (awarded to Raúl Rojas). The authors wish to thank the participating families, teachers, staff, and the dedication from all prior members of the Bilingual Language Laboratory.

References

Akhutina, T., Kurgansky, A., Polinsky, M., & Bates, E. (1999). Processing of grammatical gender in a three-gender system: Experimental evidence from Russian. *Journal of Psycholinguistic Research, 28*(6), 695–713. https://doi.org/10.1023/a:1023225129058

Anderson, R. (1999). Impact of first language loss on grammar in a bilingual child. *Communication Disorders Quarterly, 21*(1), 4–16. https://doi.org/10.1177/152574019902100102

Anderson, R., & Márquez, A. (2009). The article paradigm in Spanish-speaking children with SLI in language contact situations. In J. Grinstead (Ed.), *Hispanic child languages: Typical and impaired development* (pp. 29–55). John Benjamins. https://doi.org/10.1075/lald.50.03and

Anderson, R., & Souto, S. (2005). The use of articles by monolingual Puerto Rican Spanish-speaking children with specific language impairment. *Applied Psycholinguistics, 26*(4), 621–647. https://doi.org/10.1017/S0142716405050332

Baron, A., Bedore, L. M., Peña, E. D., Lovgren-Uribe, S. D., López, A. A., & Villagran, E. (2018). Production of Spanish grammatical forms in U.S. bilingual children. *American Journal of Speech-Language Pathology, 27*(3), 975–987. https://doi.org/10.1044/2018_AJSLP-17-0074

Baron, A., Connell, K., & Griffin, Z. M. (2022). Grammatical gender in spoken word recognition in school-age Spanish-English bilingual children. *Frontiers in Psychology, 13,* Article 788076. https://doi.org/10.3389/fpsyg.2022.788076

Bates, E., Devescovi, A., Hernandez, A., & Pizzamiglio, L. (1996). Gender priming in Italian. *Perception & Psychophysics, 58*(7), 992–1004. https://doi.org/10.3758/bf03206827

Bedore, L. M., & Leonard, L. B. (1998). Specific language impairment and grammatical morphology: A discriminant function analysis. *Journal of Speech, Language, and Hearing Research, 41*(5), 1185–1192. https://doi.org/10.1044/jslhr.4105.1185

Bedore, L. M., & Leonard, L. B. (2001). Grammatical morphology deficits in Spanish-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 44*(4), 905–924. https://doi.org/10.1044/1092-4388(2001/072)

Bedore, L. M., & Leonard, L. B. (2005). Verb inflections and noun phrase morphology in the spontaneous speech of Spanish-speaking children with specific language impairment. *Applied Psycholinguistics, 26*(2), 195–225. https://doi.org/10.1017/S0142716405050149

Bedore, L. M., Peña, E. D, Gillam, R. B., & Ho, T.-H. (2010). Language sample measures and language ability in Spanish-English bilingual kindergarteners. *Journal of Communication Disorders, 43*(6), 498–510. https://doi.org/10.1016/j.jcomdis.2010.05.002

Bedore, L. M., Peña, E. D., Summers, C. L., Boerger, K. M., Resendiz, M. D., Greene, K., Bohman, T. M., & Gillam, R. B. (2012). The measure matters: Language dominance profiles across measures in Spanish–English bilingual children. *Bilingualism: Language and Cognition, 15*(3), 616–629. https://doi.org/10.1017/s1366728912000090

Bialystok, E. (2001). Faces of bilingualism. In E. Bialystok (Ed.), *Bilingualism in development: Language, literacy, and cognition* (pp. 134–151). Cambridge University Press. https://doi.org/10.1017/CBO9780511605963.002

Bialystok, E., Barac, R., Blaye, A., & Poulin-Dubois, D. (2010). Word mapping and executive functioning in young monolingual and bilingual children. *Journal of Cognition and Development, 11*(4), 485–508. https://doi.org/10.1080/15248372.2010.516420

Bortolini, U., Caselli, M. C., Deevy, P., & Leonard, L. (2002). Specific language impairment in Italian: The first steps in the search for a clinical marker. *International Journal of Language & Communication Disorders, 37*(2), 77–93. https://doi.org/10.1080/13682820110116758

Bortolini, U., Caselli, M. C., & Leonard, L. (1997). Grammatical deficits in Italian-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 40*(4), 809–820. https://doi.org/10.1044/jslhr.4004.809

Bottari, P., Cipriani, P., Chilosi, A. M., & Pfanner, L. (1998). The determiner system in a group of Italian children with SLI. *Language Acquisition, 7*(2–4), 285–315. https://doi.org/10.1207/s15327817la0702-4_7

Bottari, P., Cipriani, P., Chilosi, A. M., & Pfanner, L. (2001). The Italian determiner system in normal acquisition, specific language impairment, and childhood aphasia. *Brain and Language, 77*(3), 283–293. https://doi.org/10.1006/brln.2000.2402

Castilla-Earls, A., Auza, A., Pérez-Leroux, A. T., Fulcher-Rood, K., & Barr, C. (2020). Morphological errors in monolingual Spanish-speaking children with and without developmental language disorders. *Language, Speech, and Hearing Services in Schools, 51*(2), 270–281. https://doi.org/10.1044/2019_lshss-19-00022

Castilla-Earls, A., Bedore, L., Rojas, R., Fabiano-Smith, L., Pruitt-Lord, S., Restrepo, M. A., & Peña, E. (2020). Beyond scores: Using converging evidence to determine speech and language services eligibility for dual language learners. *American Journal of Speech-Language Pathology, 29*(3), 1116–1132. https://doi.org/10.1044/2020_AJSLP-19-00179

Castilla-Earls, A., Francis, D., Iglesias, A., & Davidson, K. (2019). The impact of the Spanish-to-English proficiency shift on the grammaticality of English learners. *Journal of Speech, Language, and Hearing Research, 62*(6), 1739–1754. https://doi.org/10.1044/2018_jslhr-l-18-0324

Castilla-Earls, A. P., Restrepo, M. A., Pérez-Leroux, A. T., Gray, S., Holmes, P., Gail, D., & Chen, Z. (2016). Interactions between bilingual effects and language impairment: Exploring grammatical markers in Spanish-speaking bilingual children. *Applied Psycholinguistics, 37*(5), 1147–1173. https://doi.org/10.1017/s0142716415000521

Cook, V. (2003). Introduction: The changing L1 in the L2 user's mind. In V. Cook (Ed.), *Effects of the second language on the first* (pp. 1–18). Multilingual Matters. https://doi.org/10.21832/9781853596346-003

Cuza, A., & Pérez-Tattam, R. (2016). Grammatical gender selection and phrasal word order in child heritage Spanish: A feature re-assembly approach. *Bilingualism: Language and Cognition, 19*(1), 50–68. https://doi.org/10.1017/S1366728914000893

Dahan, D., Swingley, D., Tanenhaus, M. K., & Magnuson, J. S. (2000). Linguistic gender and spoken-word recognition in

French. *Journal of Memory and Language, 42*(4), 465–480. https://doi.org/10.1006/jmla.1999.2688

Davison, M. D., Hammer, C., & Lawrence, F. R. (2011). Associations between preschool language and first grade reading outcomes in bilingual children. *Journal of Communication Disorders, 44*(4), 444–458. https://doi.org/10.1016/j.jcomdis.2011.02.003

Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavioral Research Methods, 45*(4), 1246–1258. https://doi.org/10.3758/s13428-013-0326-1

Dussias, P. E., Valdés Kroff, J. R., Tamargo, R. E. G., & Gerfen, C. (2013). When gender and looking go hand in hand; Grammatical gender processing in L2 Spanish. *Studies in Second Language Acquisition, 35*(2), 353–387. https://doi.org/10.1017/S0272263112000915

Francis, D. J., Carlson, C. D., Fletcher, J. M., Foorman, B. R., Goldenberg, C. R., & Vaughn, S. (2005). Oracy/literacy development of Spanish-speaking children: A multi-level program of research on language minority children and the instruction, school and community contexts, and interventions that influence their academic outcomes. *Perspectives: The International Dyslexia Association, 31,* 8–12.

Genesee, F. (1989). Early bilingual development: One language or two? *Journal of Child Language, 16*(1), 161–179. https://doi.org/10.1017/s0305000900013490

Grosjean, F. (1985). The bilingual as a competent but specific speaker-hearer. *Journal of Multilingual & Multicultural Development, 6*(6), 467–477. https://doi.org/10.1080/01434632.1985.9994221

Grosjean F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language, 36*(1), 3–15. https://doi.org/10.1016/0093-934x(89)90048-5

Gross, M. C., & Castilla-Earls, A. (2023). Code-switching during narratives by bilingual children with and without developmental language disorder. *Language, Speech, and Hearing Services in Schools, 54*(3), 996–1019. https://doi.org/10.1044/2023_lshss-22-00149

Guiberson, M. M., Barrett, K. C., Jancosek, E. G., & Yoshinaga Itano, C. (2006). Language maintenance and loss in preschool-age children of Mexican immigrants: Longitudinal study. *Communication Disorders Quarterly, 28*(1), 4–17. https://doi.org/10.1177/15257401060280010601

Gusewski, S., & Rojas, R. (2017). Tense marking in the English narrative retells of dual language preschoolers. *Language, Speech, and Hearing Services in Schools, 48*(3), 183–196. https://doi.org/10.1044/2017_LSHSS-16-0093

Gusewski, S., Jenkins, K., & Rojas, R. (2019, November 21–23). *Utility of narrative retell language sampling to examine tense marking* [Paper presentation]. American Speech-Language-Hearing Association Annual Convention, Orlando, FL.

Gutiérrez-Clellen, V. F., Restrepo, M. A., & Simon-Cereijido, G. (2006). Evaluating the discriminant accuracy of a grammatical measure with Spanish-speaking children. *Journal of Speech, Language, and Hearing Research, 49*(6), 1209–1223. https://doi.org/10.1044/1092-4388(2006/087)

Heilmann, J., Miller, J. F., Iglesias, A., Fabiano-Smith, L., Nockerts, A., & Andriacchi, K. D. (2008). Narrative transcription accuracy and reliability in two languages. *Topics in Language Disorders, 28*(2), 178–188. https://doi.org/10.1097/01.TLD.0000318937.39301.76

Heilmann, J., Rojas, R., Iglesias, A., & Miller, J. F. (2016). Clinical impact of wordless picture storybooks on bilingual narrative language production: A comparison of the 'Frog' stories. *International Journal of Language & Communication Disorders, 51*(3), 339–345. https://doi.org/10.1111/1460-6984.12201

Herdina, P., & Jessner, U. (2013). The implications of language attrition for dynamic systems theory: Next steps and consequences. *International Journal of Bilingualism, 17*(6), 752–756. https://doi.org/10.1177/1367006912454625

Hernández-Pina, F. (1984). *Teorías psicosociolingüísticas y su aplicación a la adquisición del español como lengua materna* [Psychosociolinguistic theory and its application to the acquisition of Spanish as a first language]. Siglo XXI de España.

Hiebert, L., & Rojas, R. (2021). A longitudinal study of Spanish language growth and loss in young Spanish-English bilingual children. *Journal of Communication Disorders, 92,* Article 106110. https://doi.org/10.1016/j.jcomdis.2021.106110

IBM Corp. (2018). *IBM SPSS Statistics* [Version 26.0 for Mac]. https://www.ibm.com/products/spss-statistics

Jackson-Maldonado, D., & Maldonado, R. (2017). Grammaticality differences between Spanish-speaking children with specific language impairment and their typically developing peers. *International Journal of Language & Communication Disorders, 52*(6), 750–765. https://doi.org/10.1111/1460-6984.12312

Jacobson, P. F. (2012). The effects of language impairment on the use of direct object pronouns and verb inflections in heritage Spanish speakers: A look at attrition, incomplete acquisition and maintenance. *Bilingualism: Language and Cognition, 15*(1), 22–38. https://doi.org/10.1017/S1366728911000484

Jacobson, P. F., & Walden, P. R. (2013). Lexical diversity and omission errors as predictors of language ability in the narratives of sequential Spanish–English bilinguals: A cross-language comparison. *American Journal of Speech-Language Pathology, 22*(3), 554–565. https://doi.org/10.1044/1058-0360(2013/11-0055)

Kernan, K. T., & Blount, B. G. (1966). The acquisition of Spanish grammar by Mexican children. *Anthropological Linguistics, 8*(9), 1–14.

Leonard, L. (2016). Noun-related morphosyntactic difficulties in specific language impairment across languages. *First Language, 36*(1), 3–29. https://doi.org/10.1177/0142723715626070

Leonard, L., & Bortolini, U. (1998). Grammatical morphology and the role of weak syllables in the speech of Italian-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 41*(6), 1363–1374. https://doi.org/10.1044/jslhr.4106.1363

Leonard, L., Bortolini, U., Caselli, M. C., McGregor, K., & Sabbadini, L. (1992). Morphological deficits in children with specific language impairment: The status of features in the underlying grammar. *Language Acquisition, 2*(2), 151–179. https://doi.org/10.1207/s15327817la0202_2

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science, 18*(3), 193–198. https://doi.org/10.1111/j.1467-9280.2007.01871.x

Lew-Williams, C., & Fernald, A. (2010). Real-time processing of gender-marked articles by native and non-native Spanish speakers. *Journal of Memory and Language, 63*(4), 447–464. https://doi.org/10.1016/j.jml.2010.07.003

Maher, J. (1991). A crosslinguistic study of first language contact and language attrition. In H. W. Seliger & R. M. Vago (Eds.), *First language attrition* (pp. 66–86). Cambridge University Press. https://doi.org/10.1017/CBO9780511620720.005

Mariscal, S. (2009). Early acquisition of gender agreement in the Spanish noun phrase: Starting small. *Journal of Child Language, 36*(1), 143–171. https://doi.org/10.1017/s0305000908008908

Mayer, M. (1969). *Frog, where are you?* Penguin.

Mayer, M. (1974). *Frog goes to dinner*. Penguin.

Mayer, M. (1975a). *Frog on his own*. Dial.

Mayer, M. (1975b). *One frog too many*. Dial.

Miller, J., Andriacchi, K., & Nockerts, A. (Eds.). (2015). *Assessing language production using SALT software: A clinician's guide to language sample analysis*. SALT Software, LLC.

Miller, J., & Iglesias, A. (2020). *Systematic analysis of language transcripts (SALT)* [Research Version 2020]. SALT Software, LLC. https://www.saltsoftware.com/

Montrul, S. (2008). *Incomplete acquisition in bilingualism: Re-examining the age factor*. John Benjamins. https://doi.org/10.1075/sibil.39

Morgan, G., Restrepo, M. A., & Auza, A. (2009). Variability in the grammatical profiles of Spanish-speaking children with specific language impairment. In J. Grinstead (Ed.), *Hispanic child languages: Typical and impaired development* (pp. 283–302). John Benjamins.

Morgan, G. P., Restrepo, M. A., & Auza, A. (2013). Comparison of Spanish morphology in monolingual and Spanish–English bilingual children with and without language impairment. *Bilingualism: Language and Cognition, 16*(3), 578–596. https://doi.org/10.1017/S1366728912000697

Nicoladis, E., & Secco, G. (2000). The role of a child's productive vocabulary in the language choice of a bilingual family. *First Language, 20*(58), 3–28. https://doi.org/10.1177/014272370002005801

Oppenheim, G. M., Griffin, Z., Peña, E. D., & Bedore, L. M. (2020). Longitudinal evidence for simultaneous bilingual language development with shifting language dominance, and how to explain it. *Language Learning, 70*(S2), 20–44. https://doi.org/10.1111/lang.12398

Paradis, J. (2005). Grammatical morphology in children learning English as a second language: Implications of similarities with specific language impairment. *Language, Speech, and Hearing Services in Schools, 38,* 172–187. https://doi.org/10.1044/0161-1461(2005/019)

Paradis, J. (2010). The interface between bilingual development and specific language impairment. *Applied Psycholinguistics, 31*(2), 227–252. https://doi.org/10.1017/S0142716409990373

Peña, E. D., Bedore, L. M., Gutiérrez-Clellen, V. F., Iglesias, A., & Goldstein, B. A. (2008). *Bilingual English Spanish oral screener—Experimental version (BESOS)* [Unpublished instrument].

Pérez-Leroux, A. T., Castilla-Earls, A. P., & Bruner, J. (2012). General and specific effects of lexicon in grammar: Determiner and object pronoun omissions in child Spanish. *Journal of Speech, Language, and Hearing Research, 55*(2), 313–327. https://doi.org/10.1044/1092-4388(2011/10-0004)

Pérez-Pereira, M. (1989). The acquisition of morphemes: Some evidence from Spanish. *Journal of Psycholinguistic Research, 18*(3), 289–312. https://doi.org/10.1007/BF01067038

Restrepo, M. A. (1998). Identifiers of predominantly Spanish-speaking children with language impairment. *Journal of Speech, Language, and Hearing Research, 41*(6), 1398–1411. https://doi.org/10.1044/jslhr.4106.1398

Restrepo, M. A., & Gutiérrez-Clellen, V. F. (2001). Article use in Spanish-speaking children with specific language impairment. *Journal of Child Language, 28*(2), 433–452. https://doi.org/10.1017/s0305000901004706

Rojas, R., Irani, F., Gusewski, S., & Camacho, N. (2023). A cross-sectional investigation of disfluencies in typically developing Spanish-English bilingual children. *Journal of Fluency Disorders, 77,* Article 105988. https://doi.org/10.1016/j.jfludis.2023.105988

Sheng, L., Bedore, L. M., Peña, E. D., & Fiestas, C. (2013). Semantic development in Spanish–English bilingual children: Effects of age and language experience. *Child Development, 84*(3), 1034–1045. https://doi.org/10.1111/cdev.12015

Simon-Cereijido, G., & Gutiérrez-Clellen, V. F. (2007). Spontaneous language markers of Spanish language impairment. *Applied Psycholinguistics, 28*(2), 317–339. https://doi.org/10.1017/S0142716407070166

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford University Press.

Su, P. L., Rojas, R., & Iglesias, A. (2022). Dual language profiles in Spanish-speaking English learners. *Journal of Speech, Language, and Hearing Research, 65*(7), 2608–2628. https://doi.org/10.1044/2022_jslhr-21-00447

Van Heugten, M., & Shi, R. (2009). French-learning toddlers use gender information on determiners during word recognition. *Developmental Science, 12*(3), 419–425. https://doi.org/10.1111/j.1467-7687.2008.00788.x

Weinreich, U. (1968). *Languages in contact: Findings and problems*. Mouton & Co.

Appendix

Unconditional Growth Curve Models

Table A1. Comparisons of fixed and randomly varying slopes for unconditional growth (UG) models of article accuracy.

		UG-Models: Fixed and randomly varying slopes and fit statistics								
Article accuracy	**Model**	**Lin. – Rnd.**	**Qd. – Rnd.**	**Cb. – Rnd.**	***df***	**BIC**	**–2LL**	**Δ –2LL**	**Δ *df***	**Δ *p***
UG-Linear	**B1**	**–**	—	—	**4**	**–26.171**	**–46.050**			
	B2	+	—	—	6	NC	NC			
UG-Quadratic	**C1**	**–**	**–**	—	5	–22.096	–46.945	.895	1	.34
	C2	+	–	—	NC	NC	NC			
	C3	–	+	—	NC	NC	NC			
	C4	+	+	—	NC	NC	NC			
UG-Cubic	**D1**	**–**	**–**	**–**	6	–18.638	–48.457	2.407	2	.30
	D2	+	–	–	NC	NC	NC			
	D3	–	+	–	NC	NC	NC			
	D4	–	–	+	NC	NC	NC			
	D5	+	+	–	NC	NC	NC			
	D6	+	–	+	NC	NC	NC			
	D7	–	+	+	NC	NC	NC			
	D8	+	+	+	NC	NC	NC			

Note. Bold formatting indicates best fitting models. "—" indicates nonapplicable randomly varying slopes. "–" indicates fixed slope. "+" indicates randomly varying slope. NC: Nonconverging model; Lin. = linear; Qd. = quadratic; Cb. = cubic; Rnd. = random; *df* = degrees of freedom; BIC = Bayesian information criterion; –2LL = –2 log-likelihood difference; Δ difference between subsequent models.