JSLHR

Research Article

Code-Switching and Language Proficiency in Bilingual Children With and Without Developmental Language Disorder

Maria Kapantzoglou,[a] Julie Esparza Brown,[a] Lauren M. Cycyk,[b] and Gerasimos Fergadiotis[a]

Purpose: This study examined the frequency of code-switching by Spanish–English-speaking children as a function of language proficiency in each language and diagnosis (developmental language disorder [DLD] or typical language development [TLD]).
Method: Sixty-two Spanish–English-speaking children, 5–7 years of age, participated in this study (24 with DLD and 38 with TLD). Language samples were used to determine the level of language proficiency in each language as a continuum of performance. Correlational analyses were conducted to evaluate the relationship between Spanish and English language proficiency and the frequency of code-switching considering the total number of code-switched words, intrasentential code-switching (i.e., number of sentences/C-units with code-switched parts), and intersentential code-switching (i.e., code-switched sentences/C-units). Negative binomial and zero-inflated Poisson regressions were conducted to evaluate whether the frequency of code-switched words, code-switched parts, and code-switched sentences is different across children with DLD and their peers with TLD, controlling for Spanish and English proficiency.
Results: When Spanish was the target language, lower proficiency in Spanish was associated with higher frequency of code-switches to English for the group with DLD, but not for their peers with TLD who code-switched to English regardless of their Spanish and English proficiency. There were no statistically significant effects of diagnosis on the frequency of code-switches.
Conclusion: Results indicate that code-switching occurs similarly across children with DLD and their peers with TLD; therefore, the frequency and type of code-switches should not be used as an indicator of DLD.

Language alternation or code-switching[1] is a common phenomenon that occurs among bilingual children with typical language development (TLD) and children with developmental language disorder (DLD)/specific language impairment (Genesee & Nicoladis, 2006; Greene et al., 2012; Gutiérrez-Clellen et al., 2009). When assessing the language development of children who are bilingual, oftentimes, each language is assessed separately, asking the child to respond to language tasks in only one of their two languages (Kohnert, 2010). However, it has been well documented that bilingual children with DLD and their peers with TLD will switch between the two languages in formal and informal language testing conditions as they do in conversational language (Greene et al., 2012; Gutiérrez-Clellen et al., 2009).

Studies suggest that language proficiency in each language is one of the factors that affect the frequency of code-switching (Becker, 1997; Raichlin et al., 2019; Reyes, 2004; Ribot & Hoff, 2014). Although there is some agreement that code-switching indicates language skills in both languages, the relationship between language proficiency in each language and code-switching seems complex in both children with DLD and their peers with TLD, and findings are mixed (Gutiérrez-Clellen et al., 2009; Iluz-Cohen & Walters, 2012; Ribot & Hoff, 2014). This study examines the frequency of code-switching by Spanish–English-speaking children as a function of language proficiency in each language and diagnosis (DLD or TLD). Although

[a]Portland State University, OR
[b]University of Oregon, Eugene
Correspondence to Maria Kapantzoglou: mkapa2@pdx.edu

Editor-in-Chief: Stephen M. Camarata
Editor: Sudha Arunachalam
Received April 16, 2020
Revision received July 28, 2020
Accepted December 19, 2020
https://doi.org/10.1044/2020_JSLHR-20-00182

[1]The terms *code-switching* and *code-mixing* have been used to indicate intra- and intersentential language alternation but are oftentimes used interchangeably. In this study, we will be using the term *codes-switching* to indicate any word-, phrase-, and sentence-level language change in relation to the target language during tasks.

Disclosure: *The authors have declared that no competing interests existed at the time of publication.*

code-switching is a typical phenomenon in bilingual communities and in and of itself would not be an indicator of DLD (Restrepo, 1998; Simon-Cereijido & Gutiérrez-Clellen, 2007), understanding such relationships is important for diagnostic purposes. Code-switches in one or both languages may raise concerns as an indicator of an impaired mechanism for learning or monitoring language when there is an expected language (Dollaghan, 1998; Lanvers, 2001; Miccio et al., 2009; Spaulding, 2010). It is particularly challenging to distinguish between different and impaired communication skills in developing bilinguals, especially when the second language has not yet developed enough and the first is being gradually lost versus when both languages are highly valued and supported (Kohnert, 2010).

Determining children's proficiency in each language is a complex task, and it affects the study of the relationship between language proficiency and code-switching during testing (Bedore et al., 2012; Smyk et al., 2013; Treffers-Daller, 2019). Parent and/or teacher reports are oftentimes used to determine the child's exposure to and use of each language during the day and to make inferences about how strong each language is (i.e., language proficiency; Gutiérrez-Clellen et al., 2009) and the relative strength of each language with respect to the other (i.e., language dominance; Greene et al., 2012; Gutiérrez-Clellen et al., 2009; Ribot & Hoff, 2014). However, parent and/or teacher reports are indirect measures and substantially affected by the frequency with which parents and teachers interact with children in one particular language and in specific contexts. Also, given that language characteristics change across contexts and conditions (Grosjean, 1997; Rojas et al., 2016), parent and teacher reports about children's communication may not always accurately reflect the child's proficiency in each language and language dominance in the testing context. Language sample analysis has been one of the informative and reliable methods of language assessment for bilingual children to identify DLD (Dollaghan & Horner, 2011; Restrepo, 1998) and to determine language proficiency specifically (MacSwan & Rolstad, 2006). The goals of this study were to examine (a) the relationship between language proficiency (as measured directly by sentence length and complexity, grammaticality, vocabulary, and fluency) and code-switching in language samples elicited through story-retelling tasks in 5- to 7-year-old Spanish–English-speaking children (kindergarteners and first graders) with DLD and their peers with TLD and (b) the effects of diagnosis (DLD, TLD) on the frequency and type of code-switching in each language, controlling for the Spanish and English proficiency of the child.

Factors Affecting Code-Switching in Bilinguals

Studies indicate that there are various factors affecting the frequency of code-switching (Becker, 1997; Reyes, 2004; Ribot & Hoff, 2014). Becker (1997) identifies three main groups of factors that impact the frequency of code-switching, including the structural linguistic characteristics of the two languages, internal psycholinguistic factors, and external social factors. For example, the frequency of code-switching increases when two languages share many equivalent linguistic structures and decreases when two languages are structurally very different (Becker, 1997; Woolford, 1983). Internal psycholinguistic factors that may play a role include using code-switching to emphasize information or to convey a more refined message (Becker, 1997; Zentella, 1990). For instance, a child who is bilingual may code-switch to emphasize a particular characteristic, "The beach we went to yesterday fue muy hermosa (was very beautiful)," or to more precisely express meanings that do not translate exactly between the two languages, "Do you like my new shoes? Me los estoy estrenando (I am wearing them for the first time)." Such internal factors may be either conscious or nonconscious; therefore, they may affect the frequency of code-switching even when bilingual students are instructed to speak in only one language. Finally, bilingual children attend to various external social cues that affect the frequency of code-switching. These may include, for example, the physical characteristics of the listener (age, gender, or ethnicity), the language proficiency and preference of the interlocutor, the topic of conversation, and the setting in which a particular interaction occurs (Bailey & Huang, 2011; Becker, 1997; Grosjean, 1997, 2001; Raichlin et al., 2019; Zentella, 1990).

Language Proficiency and Code-Switching

The relationship between a child's proficiency in each language and code-switching has primarily been studied by examining differences in code-switching between groups of children as determined by the relative strength of each of their languages, such as balanced Spanish–English bilinguals and first language (L1)–dominant and second language (L2)–dominant groups. Therefore, the review that follows discusses the general findings of this body of work. An important issue when studying dominance groups is the complexity in determining the strongest language and across-study differences regarding the operationalization of the construct (Treffers-Daller, 2019). For instance, some studies might operationalize dominance using measures of language exposure and use rather than proficiency measures in each language (e.g., Greene et al., 2012). Although there is relationship between language use and proficiency, children who hear or use a language more frequently do not always have higher proficiency in that language (Bedore et al., 2012). Language proficiency is typically conceptualized as a continuum of performance versus a categorical characteristic; however, there are few studies that have examined code-switching and treated language proficiency as a continuum (e.g., Gutiérrez-Clellen et al., 2009; Montanari et al., 2019). For example, Gutiérrez-Clellen et al. (2009) determined proficiency and dominance based on both indirect (parent and/or teacher report) and direct measures (children's scores on a prepublication version of the Morphosyntax subtest of the Bilingual English–Spanish Assessment [BESA]; Peña et al., 2014). There are few studies that have relied only on direct, observational measures of language proficiency in more naturalistic contexts (e.g., Montanari et al., 2019, used utterance length and lexical diversity in language samples as indicators of proficiency).

In general, results indicate an inverse relationship between language proficiency or dominance in a language and code-switching when speaking that language (Gutiérrez-Clellen et al., 2009; Poulisse & Bongaerts, 1994). Specifically, various studies indicate that *dominant* bilinguals (i.e., with one language being stronger than the other) code-switch with lower frequency when they speak in their stronger language (Greene et al., 2012; Gutiérrez-Clellen et al., 2009). For example, Greene et al. (2012) found that English-dominant participants, as indicated by reported language exposure and use, code-switched less frequently to Spanish when lexical items were presented in English, as compared to more code-switches to English when lexical items were presented in their weaker language—Spanish. Similarly, Spanish-dominant children code-switched less frequently to English when lexical items were presented in Spanish as compared to more code-switches to Spanish when lexical items were presented in children's weaker language—English. Gutiérrez-Clellen et al. (2009) also found that English-dominant children, based on proficiency measures, code-switched less frequently to Spanish when asked to retell a story in English as compared to the higher number of code-switches from Spanish to English when the story was elicited in Spanish. On the other hand, in the same study, Spanish-dominant children tested in their weaker language (English) showed a very low frequency of code-switches to their stronger language (Spanish), and they code-switched more frequently from their stronger to their weaker language. It could be that Spanish-dominant children preferred the mainstream language for sociocultural factors as the authors argued. This explanation is supported by studies that indicate no remarkable relationships between code-switching and proficiency or dominance. For example, in Greene et al.'s study, the frequency of code-switches to Spanish in English single-word naming was low across the language dominance groups. Also, Montanari et al. (2019) did not find a relationship between children's proficiency and the frequency of code-switches in 4.5-year-old Spanish–English-speaking preschoolers.

Some literature indicates that *balanced* bilinguals (i.e., roughly equally proficient in both languages) also code-switch with significant frequency and often more frequently than dominant bilinguals (Gollan & Ferreira, 2009; Greene et al., 2012). In contrast, Ribot and Hoff (2014) reported that children who did not code-switch tended to be balanced bilinguals. An important point is that Spanish-dominant, English-dominant, and balanced groups may include children that vary substantially in their levels of proficiency. In this case, highly proficient balanced bilinguals may perform differently than balanced bilinguals with low proficiency, or it may be that the level of difference in proficiency between the dominant and nondominant languages may affect the frequency of code-switching to different degrees (Costa & Santesteban, 2004; Poulisse & Bongaerts, 1994). For example, Costa and Santesteban (2004) argued that, although there might be differences in the effort needed to switch from a dominant L1 to a weaker L2 and vice versa, this might not be the case in highly proficient bilinguals or between a dominant L1 and a much weaker third language. Thus, balanced groups may perform differently across studies if a balanced group has a larger number of highly proficient children in one study than in another, and dominant groups may perform differently if they differ in the range of imbalance between the two languages across studies. Treating language proficiency performance as a continuum versus examining language proficiency groups facilitates a more accurate study of the variability of language proficiency levels and of the relationship between proficiency and the frequency of code-switches.

Finally, types of code-switching may also differ depending on language proficiency levels of each language (Montanari et al., 2019; Sankoff & Poplack, 1981; Yow et al., 2016). Both Sankoff and Poplack (1981) and Yow et al. (2016) found that less proficient bilinguals tend to use more intersentential code-switching (i.e., switching after the sentence in one language has been completed), which is easier, whereas more proficient bilinguals are more comfortable with intrasentential code-switches (i.e., combining the grammatical rules of the two languages within a sentence). In contrast, Montanari et al. (2019) obtained mixed findings in their longitudinal study of Spanish–English-speaking preschoolers observed at 3.5 and 4.5 years of age. Although the authors did not focus on differences in proficiency, children were predominately Spanish speaking when enrolled. They also attended English immersion Head Start programs focusing on academic readiness in Los Angeles, so English proficiency reasonably increased with time. Children at 4.5 years of age showed more intrasentential code-switching from English to Spanish (9%) than intersentential code-switching (5%), but no differences were identified for Spanish to English code-switches and at the younger age group. Given that the effects of proficiency level may be different for intrasentential and intersentential instances of code-switching, it is informative to examine the effects of proficiency on total counts of code-switching as well as on intra- and intersentential types in particular.

Measuring Language Proficiency in Bilinguals

The measurement of language proficiency is one of the main challenges when studying the relationship between language proficiency and code-switching (Bedore et al., 2012; MacSwan, 2000; MacSwan & Rolstad, 2006; Pray, 2005). There is general agreement that language proficiency in a language is reflected by the diversity of vocabulary used, grammatical accuracy, syntactic complexity of the utterances, and overall fluency of the speaker (Norris & Ortega, 2009; Skehan, 2009; Smyk et al., 2013). However, the characteristics of language in a sample used for assessment can vary significantly across different types of measures and sampling contexts (Bedore et al., 2012; Kapantzoglou et al., 2017). This is an issue in language evaluation for monolinguals, but more so for bilinguals, whose language proficiency can differ substantially across contexts and measures depending on individual experiences in each language (Bailey & Huang, 2011; Grosjean, 1997, 2001). As a result, the proficiency levels determined by a particular measure and in a particular context may not be an accurate

representation of the proficiency level of the speaker in a different context. For instance, many proficiency measures commonly used by state education agencies have been criticized as biased due to the academic language included in the test items (MacSwan & Rolstad, 2006; Pray, 2005). Bilingual students with high proficiency may underperform on such measures due to low familiarity with academic language, even when they are native speakers of the language being tested (Pray, 2005).

Parent reports, commonly used to capture proficiency (e.g., Gutiérrez-Clellen et al., 2009), are oftentimes based on parent–child interactions that occur in only one of the two languages and in the home context. However, the communication and language elicited through parent–child interactions at home may differ significantly from the language obtained in language testing contexts (Bailey & Huang, 2011). In addition, although parents are good reporters of vocabulary levels, they may not be as sensitive to grammatical accuracy and syntactic complexity (Marchman & Martínez-Sussmann, 2002).

Teacher reports, another commonly used proficiency measure, may yield different results, given that teachers may have more interactions with children in English in the school setting (Bedore et al., 2011; Gutiérrez-Clellen & Kreiter, 2003; Restrepo, 1998). Moreover, studies have indicated that teacher reports account for only a small percentage of children's skills as measured from language samples and other tests. For example, Gutiérrez-Clellen and Kreiter (2003) found that teacher report accounted for only 19.36% ($r^2 = 19.36$) of the grammaticality in language samples collected from 5- to 6-year-old Spanish–English-speaking children. These results were consistent with findings from Bedore et al. (2011), who found that parent reports regarding 4- to 5-year-old Spanish–English speakers accounted for only 3.24% of children's performance on semantic subtests and 5.29% of children's performance on grammatical subtests.

Language sample analysis has many advantages as a method of language assessment. It is currently one of the most valid direct methods for assessing language for diagnostic purposes (Dollaghan & Horner, 2011; Restrepo, 1998) and for determining language proficiency in bilinguals (MacSwan & Rolstad, 2006). For example, due to its ecological validity, it has been used extensively to study language development in children (Jackson-Maldonado et al., 1998; Rice et al., 1995). Also, language sample analysis provides clinicians and researchers the opportunity to measure microlinguistic constructs for diagnostic purposes, such as lexical diversity, grammaticality, and syntactic complexity during discourse, a complex cognitive–linguistic behavior (Heilmann et al., 2008; Kapantzoglou et al., 2017). Overall, analyzing discourse yields information about whether and how possible language characteristics and/or limitations may affect children's communication in daily life. Currently, there is no study to our knowledge that has used language samples to determine the level of language proficiency, as a continuum of performance, and to examine its relationship with the frequency and type of code-switches. Given the variability of language proficiency across contexts, this approach can be informative for determining the relationship between language proficiency and code-switching.

Code-Switching in Children With DLD

Children with DLD have language deficits, which may include difficulties in one or more language domains—phonology, semantics, grammar, and/or pragmatics (Bishop & Leonard, 2000; Kapantzoglou et al., 2015)—in the absence of cognitive, motor, neurological, or hearing problems (Leonard, 2014). In bilinguals, there is an increased likelihood of misdiagnosis of DLD because differences in language performance may reflect typical variability in language experiences as opposed to deficits in the underlying mechanism of language development (Barragan et al., 2018; Kohnert, 2010; Morgan et al., 2013). Along these lines, a high number of code-switches might be interpreted as an indicator of DLD because children with DLD are expected to have more lexical gaps in each language, and therefore, they may be expected to code-switch with higher frequency than their peers with TLD (Dollaghan, 1998; Lanvers, 2001; Walters, 2005). Additionally, children with DLD may code-switch more frequently due to poor inhibitory control, which might yield differences in the frequency with which they suppress nontarget information activated, compared to their peers with TLD (Spaulding, 2010). Alternatively, it could be that children with DLD code-switch similarly to children with TLD, and for both groups of children, the frequency and types of code-switches are primarily determined by factors other than children's diagnosis.

Studies that have assessed the frequency of code-switching in children with DLD and their peers with TLD are limited, and the findings are mixed. For instance, Gutiérrez-Clellen et al. (2009) focused on intrasentential code-switching and found that children with DLD did not code-switch more than their peers with TLD in language samples elicited through story-retelling and storytelling tasks. The study focused on 5- to 6-year-old Spanish–English-speaking children who code-switched, and comparisons were made based on the percentage of sentences with intrasentential code-switching. All children had approximately two years of exposure to English, and the DLD and TLD groups had similar levels of proficiency in the two languages. Greene et al.'s (2012) findings, in contrast, indicated the opposite pattern. Five-year-old balanced bilingual preschoolers at risk for DLD code-switched more frequently than their balanced bilingual peers with TLD and their Spanish- and English-dominant peers with TLD and DLD in structured elicitation tasks administered in English. Specifically, single-word code-switching was examined in this study using 15 combined expressive semantic items from a prepublication version of the Bilingual English–Spanish Oral Screener (Peña et al., 2010), administered in English and Spanish. Results were consistent with Iluz-Cohen and Walters' (2012) study on English–Hebrew-speaking preschoolers living in English-speaking homes in Israel and attending Hebrew-speaking preschool programs. Overall, children with DLD code-switched more frequently than their peers with TLD

in both directions and across home, preschool, and bilingual settings, with the only exception being the preschool setting where children with TLD code-switched to L2 more frequently than their peers with DLD. The latter study did not examine language proficiency groups, while Gutiérrez-Clellen et al. used a combination of direct structured tasks and indirect measures to determine language proficiency and dominance, and Greene et al. used indirect measures to determine dominance groups based on language exposure and use. Also, none of the previous studies compared intra- and intersentential code-switching across children with DLD and children with TLD. This study measures language proficiency directly, using children's language samples, and controls for it when assessing differences in code-switching between children with DLD and their peers with TLD to more accurately examine the effects of diagnosis on the frequency and types of code-switching. Specifically, this study examines the following research questions:

1. What is the relationship between Spanish and English language proficiency and the frequency of code-switching considering the total number of code-switched words (CSW), intrasentential code-switching (i.e., code-switched parts of the sentence [CSP]), and intersentential code-switching (i.e., code-switched sentences/C-units [CSS])? We hypothesize that, for children attending Spanish dual immersion programs in the United States, there is a negative relationship between language proficiency and the frequency of code-switching (Greene et al., 2012; Gutiérrez-Clellen et al., 2009; Poulisse & Bongaerts, 1994). That is, higher levels of Spanish proficiency will be associated with lower levels of code-switching to English, and higher levels of English proficiency will be associated with lower levels of code-switching to Spanish. Considering that intrasentential code-switching might be more challenging than intersentential code-switching (Sankoff & Poplack, 1981), we anticipate that children with higher Spanish proficiency will have a lower number of intersentential code-switching to English and an even lower number of intrasentential code-switching. Children with lower Spanish proficiency are expected to have a relatively high number of intrasentential code-switching and an even higher number of intersentential code-switching, which might be easier.
2. Is the frequency of CSW, CSP, and CSS different across children with DLD and with TLD, controlling for Spanish and English proficiency? We hypothesize that if there are differences after controlling for children's language proficiency, these differences are going to be found for intrasentential code-switching, which requires combining the rules of the two languages.

Method

Participants

Sixty-two Spanish–English-speaking children, 5–7 years of age, participated in this study (TLD: $n = 38$, $M_{age} = 74.5$, $SD = 7.66$; DLD: $n = 24$, $M_{age} = 72.58$, $SD = 7.71$). Table 1 includes participant characteristics. The age difference between the groups with DLD and TLD was not statistically significant. There were 34 boys and 28 girls. All children were recruited from kindergarten and first-grade Spanish dual immersion programs in public schools in a metropolitan area in the northwestern United States. The programs followed a 90/10 model for dual immersion. Instruction in kindergarten was delivered 90% in Spanish and 10% in English. For each subsequent year, the percentage of Spanish was decreased by 10% and increased in English. By fourth grade, instruction

Table 1. Participant characteristics based on parent report.

Characteristic	TLD (*N* = 38)	DLD (*N* = 24)
Age		
M (months)	74.5	72.58
SD	7.66	7.71
Gender	%*N* = 100	%*N* = 100
Male	20	14
Female	18	10
Language dominance	%*N* = 100	%*N* = 100
Stronger Spanish	8	5
Stronger English	7	4
Balanced	23	15
Spanish language exposure	%*N* = 97.4	%*N* = 95.8
Always/frequently	32	21
Sometimes	3	1
Infrequently	2	1
Missing	1	1
English language exposure	%*N* = 92.1	%*N* = 91.7
Always/frequently	22	15
Sometimes	9	5
Infrequently/none	4	2
Missing	3	2
Spanish language use	%*N* = 94.7	%*N* = 100
Always/frequently	28	19
Sometimes	6	4
Infrequently	2	1
Missing	2	0
English language use	%*N* = 97.4	%*N* = 95.8
Always/frequently	19	13
Sometimes	12	6
Infrequently/none	6	4
Missing	> 1	> 1
Type of Spanish	%*N* = 89.5	%*N* = 95.8
Mexican	29	17
Guatemalan	1	3
Cuban	2	1
Other	2	2
Missing	4	1
Maternal education	%*N* = 92.1	%*N* = 100
College/university	6	3
Technical school	3	1
High school	16	10
Elementary	6	6
Other	3	2
No diploma	1	2
Missing	3	0
Free or reduced-price lunch	%*N* = 89.5	%*N* = 91.7
Yes	34	22
Missing	4	2

Note. TLD = typical language development; DLD = developmental language disorder.

in the two languages is a balanced 50%–50%. In this model, literacy was taught in Spanish first. Children with DLD were recruited from the same schools and classrooms as children with TLD. Regarding participants' language dominance, as measured by the Spanish–English Language Proficiency Scale (SELPS; Smyk et al., 2013) based on language samples, 13 children had stronger Spanish (a difference of 0.5 points or greater on the 5-point scale of the SELPS; TLD = 8, DLD = 5), 38 were balanced (TLD = 23, DLD = 15), and 11 had stronger English (TLD = 7, DLD = 4). All children were exposed to Spanish from birth and came from Spanish-speaking or bilingual Spanish–English homes with at least one of their parents speaking Spanish. The great majority of children spoke Mexican Spanish and were from lower socioeconomic backgrounds determined on the basis of eligibility for free or reduced-price lunch from parent report.

All child participants met the following criteria:

1. no history of hearing loss, sensorimotor or neurological problems, severe psychological disorders, or health problems, according to parent report, and
2. passed a hearing screening at 500 Hz at 25 dB and at 1000, 2000, and 4000 Hz at 20 dB in both ears (ASHA Panel on Audiologic Assessment, 1997).

All children classified as having TLD also met three of the following four criteria for at least one of their two languages: (a) Parent report indicated no concern of language impairment (Pratt et al., 2020; Restrepo, 1998); (b) the number of grammatical errors per C-unit in the language sample was below 20% (Restrepo, 1998); (c) the mean length of utterance in words was age appropriate (Simon-Cereijido & Gutiérrez-Clellen, 2007); and (d) 5- and 6-year-old children scored at or above the cut-score on the Morphosyntax subtest of the BESA, and 7-year-old children scored within the normal range (cut-score = 1 *SD* below the mean) on two grammatical subtests of the Clinical Evaluation of Language Fundamentals–Fourth Edition (CELF-4) Spanish (Word Structure and Recalling Sentences; Wiig et al., 2006) and CELF-5 in English (Word Structure and Recalling Sentences; Wiig et al., 2006).

All children with DLD met two of the following four criteria for both languages: (a) Parent report indicated concern of language impairment (Restrepo, 1998); (b) the number of grammatical errors per C-unit in the language sample was 20% or above; (c) the mean length of utterance in words was low for their age (Simon-Cereijido & Gutiérrez-Clellen, 2007); and (d) 5- and 6-year-old children scored below the cut-score on the Morphosyntax subtest of the BESA (Peña et al., 2018), and 7-year-old children scored below the cut-score of 1 *SD* below the mean on two grammatical subtests of the CELF-4 Spanish (Word Structure and Recalling Sentences; Wiig et al., 2006) and CELF-5 in English (Word Structure and Recalling Sentences; Wiig et al., 2013). Cloze and sentence repetition tasks, including the Word Structure and Recalling Sentences subtests, respectively, are commonly used grammatical tasks for evaluating language skills in bilingual Spanish–English-speaking children (Bedore & Leonard, 2001; Kapantzoglou et al., 2016; Morgan et al., 2013; Simon-Cereijido & Gutiérrez-Clellen, 2007).

General Procedure

Parent questionnaires and consent forms were distributed and collected by teachers. Parents completed the questionnaires at home and returned them to the teachers. Children whose parents agreed to participate in the study and returned the parent questionnaires were further evaluated and classified into the groups with TLD and DLD. The remaining qualification measures were administered in two separate sessions of 40–60 min each within a 2-week interval. Trained bilingual examiners blind to the children's language status assessed the children during the school day in a quiet area in the school. To maintain the same spoken language throughout each session, in the first session, children completed the hearing screening and the Spanish diagnostic measures. In the second session, children completed the English diagnostic measures. Because children's assignment to examiners was random in each session, many children were tested by different examiners in each language, while others had the same examiner in both sessions. Examiners were female, bilingual Spanish–English research assistants, and considering the school and broader social context, examiners presented themselves to children as bilinguals as well. They established rapport with the children in the beginning of the session in the target language and explicitly shared with them the plan to focus on one language in each of the two sessions. The examiners gave occasional reminders to use the target language during each task if children switched to using only the nontarget language.

Measures

Parent Report

All parents filled out an experimenter-created questionnaire requesting demographic information, parents' and child's education, child's language development, child's medical history, family history related to language and learning skills, and child's exposure to and use of each language.

BESA

The BESA is a standardized, norm-referenced test designed as a diagnostic tool for children with potential DLD who speak Spanish. The Morphosyntax subtest was used because it is considered to be accurate between the ages of 5 years and 6 years 11 months in bilingual children. According to the technical manual, for Spanish–English-speaking children between 5 years and 6 years 11 months, the sensitivity of the Morphosyntax subtest is 88.6–88.9, and the specificity is 81.6–88.2.

CELF-4 Spanish

Seven-year-old children were evaluated for DLD in Spanish using CELF-4 Spanish. The test manual reports sensitivity of 96% and specificity of 87% for the core language score at 1 *SD* below the mean.

CELF-5

Seven-year-old children were evaluated for DLD in English using CELF-5. The test manual reports sensitivity of 100% and specificity of 91% for the core language score at 1 *SD* below the mean.

Language Sample Analyses

A language sample in the form of a story retell was collected from each child in each language to assess their language abilities based on the number of grammatical errors in the language sample (Restrepo, 1998). The examiner read the script of two different wordless picture books, *A Boy, a Dog, a Frog, and a Friend* (Mayer & Mayer, 1967) and *Frog on His Own* (Mayer, 1973), in the target language and then asked the child to retell the story to the examiner (Systematic Analysis of Language Transcripts; Miller & Iglesias, 2012). The two stories were randomly assigned to the two languages for each student. Narratives were transcribed and coded for grammatical errors by trained research assistants using the Systematic Analysis of Language Transcripts (Miller & Iglesias, 2012) computer program. Semantic, phonological, or cohesive errors were not counted as grammatical errors. Instead, omissions, such as *Jugaron *prep el barco (They played *prep the boat)*; substitutions, such as *el rana (the-Masculine frog)*, instead of *la rana (the-Feminine frog)*; additions, such as *el niño volvió con[Addition] a casa (the boy returned with[Addition] at home)*; and word order mixes were considered grammatical errors.

The frequency of code-switches was also measured based on the Spanish and English language samples. Three different measures were estimated for each target language: (a) total number of CSW in the sample (Spanish sample: "El dog buscó a la rana in a hole" = 4 CSW), (b) intrasentential code-switching as measured by the number of sentences/C-units with CSP (one or more words within a sentence, but not the entire sentence; Spanish sample: "El dog buscó a la rana in a hole" = 1 CSP), and (c) intersentential code-switching as measured by the number of CSS (a complete C-unit; Spanish sample: "The dog looked for the frog in a hole" = 1 CSS).

Reliability for language sample analyses. Two trained fluent bilingual examiners scored 32% of the samples independently for C-units and grammatical errors. Interrater reliability values were 96% for 1-point-off agreement for number of C-units and 87% for grammatical errors. Any discrepancies were resolved by consensus with a third research assistant.

SELPS

Children's Spanish and English language samples, based on the "frog stories," were used to assess each child's language proficiency using the SELPS. The SELPS measures sentence length and complexity, grammaticality, vocabulary, and fluency. An overall proficiency level was determined ranging from 1 to 5 (1 = *silent/observer*, 2 = *a few words or formulaic phrases*, 3 = *short sentences and phrases with multiple grammatical errors*, 4 = *full sentences with a few grammatical errors*, 5 = *native-like productions*). The weighted *k* estimate for the scale is .81, indicating high interrater reliability. Language samples were rated after they were transcribed, having as a reference both the language sample transcription and the audio file. For this study, raters focused on utterances in the target language and did not penalize for code-switching. Raters were bilingual Spanish–English research assistants who underwent training for using the scale.

Reliability for the SELPS. Two trained fluent bilingual examiners scored all language samples independently. There was 92% 0.5-point-off agreement. For differences of 0.5 points or smaller, the two scores were averaged. For differences greater than 0.5, a third fluent bilingual examiner scored the language samples independently. The two scores with a difference of 0.5 points or smaller were averaged. Larger differences were solved by consensus among the three examiners.

Analyses

Data were prepared for statistical analysis following Kline (2010) and Tabachnick and Fidell (2007). After importing data in SPSS (Version 23, IBM Corp.), data were screened for missing values and univariate outliers defined as data points 4 *SD*s from the mean (Kline, 2010; Stevens, 2002). The percentage of missing data ranged from 0% to 11.29% across individual variables, with a total of 18 missing values out of 806 data points (2.23%). The reasons for missing scores included inability to test a child at a given time, recording equipment failures, and data entry concerns. These reasons are unrelated to the target ability; therefore, data were assumed to be missing completely at random (Little & Rubin, 1989; Rubin, 1976).

Distributions were visually inspected, and Kolmogorov–Smirnov tests of normality were conducted to assess the normality assumption (see Table 2). Due to high frequency of zero values in the dependent variables (CSW, CSP, and CSS in Spanish and English samples) and skewed data, the assumption of normality was violated for all dependent variables (see Table 2). Levene's tests were used to assess the assumption of equality of variances for each variable across the TLD and DLD groups. Equality of variances was not assumed for the analyses due to unequal sample sizes across groups and statistically significant results on Levene's tests for four out of six of the dependent measures (CSW, CSP, and CSS in Spanish samples and CSS in English samples) and for the total number of words in Spanish samples (see Table 2).

Given the above-described characteristics of the data, Spearman's rho nonparametric correlations were conducted to evaluate the relationships between Spanish and English language proficiency and the frequency of code-switching considering the total number of CSW, intrasentential code-switching (i.e., CSP), and intersentential code-switching (i.e., CSS; see Table 3). For descriptive statistics and correlations, to control for variations in language sample length, and to maximize the interpretability of the outcomes, all measures were converted into percentages. Specifically, CSW was divided by the number of total words in the language sample, regardless of language, and multiplied by 100; CSP and CSS were divided by the total number of sentences in the sample (i.e., C-units), regardless of language, and multiplied by 100.

Table 2. Summary of means, variances, standard deviations, skewness, and kurtosis for language proficiency and code-switching scores for participants with developmental language disorder (DLD) and their peers with typical language development (TLD).

Measure	LPSp		LPEng		CSWSp		CSPSp		CSSSp		CSWEng		CSPEng		CSSEng	
DLD	TLD	DLD	TLD	DLD	TLD	DLD	TLD	DLD	TLD	DLD	TLD	DLD	TLD	DLD	TLD	DLD
M	3.7	3.1	3.6	3.17	4.65	12.11	4.38	19.09	0.52	1.58	0.12	0.27	0.44	0.92	0	0.09
Mdn	4	3.1	4	3.13	0.34	3.39	0	5.77	0	0	0	0	0	0	0	0
Variance	0.54	0.6	0.52	0.54	314	474.64	132	796.61	5.23	9.67	0.3	1.18	2.62	12.77	0	0.2
SD	0.73	0.77	0.72	0.73	17.75	21.79	11.49	28.22	2.29	3.11	0.55	1.09	1.61	3.57	0	0.44
Skewness	–1.7	–0.56	–0.31	0.78	4.96	3.16	3.83	1.68	4.8	2.02	5.39	4.65	4.2	4.64	0	4.9
Kurtosis	4.8	1.52	0.73	2.6	26.2	11.74	14.3	1.8	24	3.01	30.3	22.14	18.8	22.06	0	24

Note. TLD, $n = 38$; DLD, $n = 24$. LPSp = Spanish language proficiency; LPEng = English language proficiency; CSWSp = code-switched words / total words × 100 in the Spanish samples; CSPSp = sentences with code-switched parts / total sentences × 100 in the Spanish samples; CSSSp = code-switched sentences / total sentences × 100 in the Spanish samples; CSWEng = code-switched words / total words × 100 in the English samples; CSPEng = sentences with code-switched parts / total sentences × 100 in the English samples; CSSEng = code-switched sentences / total sentences × 100 in English language samples.

To evaluate whether the frequency of CSW, CSP, and CSS is different across the TLD and DLD groups, controlling for language proficiency in each language, Poisson, zero-inflated Poisson, and/or negative binomial regressions were conducted using the binomTools package in R (Haubo et al., 2014). Poisson regression is typically used as the starting model for count-dependent variables. Zero-inflated Poisson regression is used to deal with high number of zeros in the outcome variable when there is no overdispersion (i.e., variance is similar to the mean), whereas when there is overdispersion (i.e., variance is substantially greater than the mean), negative binomial regression models are a better fit for the data. Six dependent variables—the count of code-switches in CSW, CSP, and CSS in Spanish and English samples—were modeled as a function of children's diagnosis (TLD and DLD), with both English and Spanish language proficiency scores on the SELPS used as covariates. To control for variations in language sample length across participants, the total number of words in the Spanish or English samples was used as the setoff variable for the Spanish and English dependent variables, respectively.

Table 3. Summary of intercorrelations for language proficiency and code-switching scores for participants with developmental language disorder (DLD) and their peers with typical language development (TLD).

Measure	1	2	3	4	5	6	7
TLD ($n = 38$)							
1. LPSp							
2. LPEng	.17						
3. CSWSp	–.06	–.04					
4. CSPSp	.02	–.15	.77**				
5. CSSSp	–.24	.06	.39*	.42**			
6. CSWEng	–.11	–.27	.15	.15	–.07		
7. CSPEng	–.11	–.27	.15	.15	–.07	1**	
8. CSSEng	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DLD ($n = 24$)							
1. LPSp							
2 LPEng	.47*						
3. CSWSp	–.70**	–.11					
4. CSPSp	–.42*	–.09	.77**				
5. CSSSp	–.41*	.01	.42*	.26			
6. CSWEng	.19	–.05	–.05	.09	–.24		
7. CSPEng	.19	–.05	–.05	.09	–.24	1**	
8. CSSEng	.05	–.28	–.26	–.22	–.13	.60**	.60**

Note. LPSp = Spanish language proficiency; LPEng = English language proficiency; CSWSp = code-switched words / total words × 100 in the Spanish samples; CSPSp = sentences with code-switched parts / total sentences × 100 in the Spanish samples; CSSSp = code-switched sentences / total sentences × 100 in the Spanish samples; CSWEng = code-switched words / total words × 100 in the English samples; CSPEng = sentences with code-switched parts / total sentences × 100 in the English samples; CSSEng = code-switched sentences / total sentences × 100 in English language samples; N/A = not applicable.

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

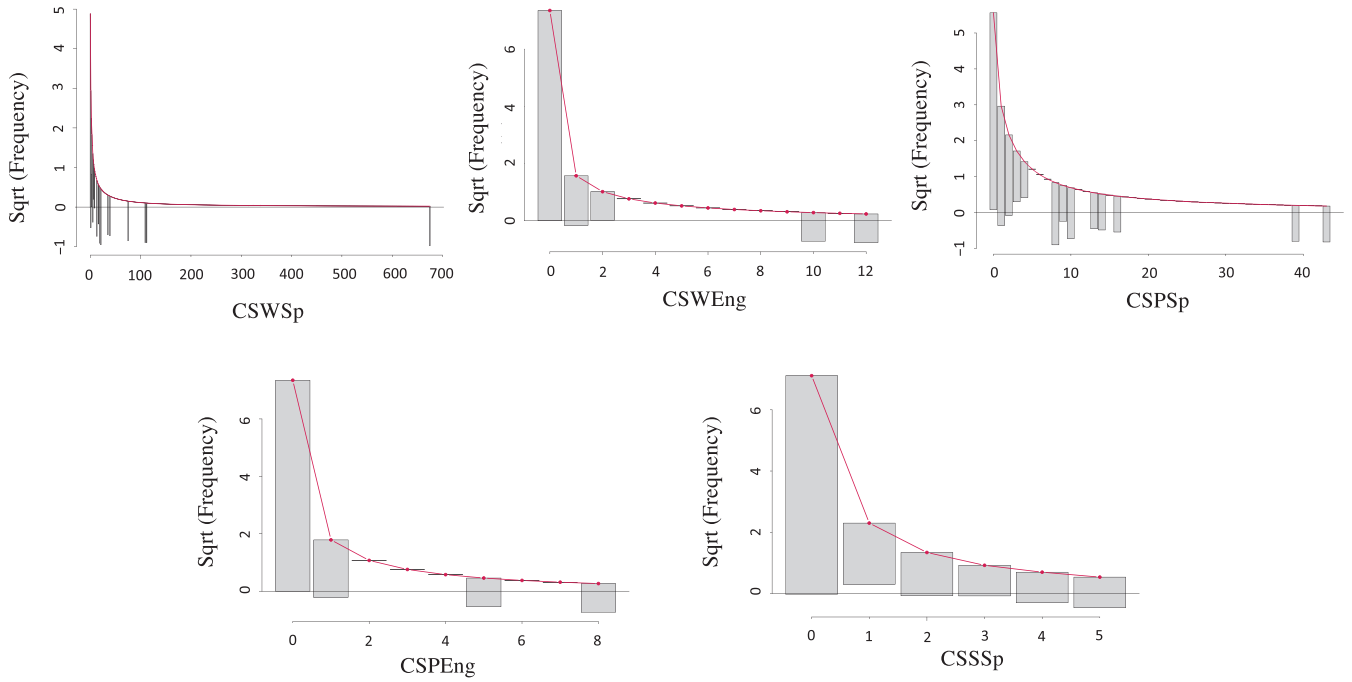
Preferred models were selected on the basis of a set of statistics. A χ^2 statistic with $p > .05$ suggested statistically good fit (i.e., there was not a statistically significant difference between observed and model-implied data). To compare goodness of fit between pairs of models, the Akaike information criterion (Akaike, 1974) was used, according to which smaller Akaike information criterion values indicate better model fit. For comparisons on the basis of significant improvement in fit for nested models, the χ^2 test, based on the difference of log-likelihoods, was used (χ^2 ΔLog L), and for nonnested models, the Vuong test was used, with $p < .05$, indicating statistically significant improvement in fit for both tests. Interpretability of the results was also considered in addition to fit indices (e.g., Marsh et al., 2005). Visual inspection of rootogram plots of the negative binomial models (see Figure 1) was conducted, and the number of zeros predicted by each model was estimated and compared to the number of zeros observed in the data set.

Results

Descriptive Statistics

Spearman's rho (r_s) nonparametric correlations, means, medians, standard deviations, variance, skewness, and kurtosis for all study variables for children with DLD

Figure 1. Rootograms of observed and predicted code-switching counts for each negative binomial model and the zero-inflated Poisson model for code-switched sentences in Spanish. The square-root frequency of the counts (*y*-axis) is plotted as a function of binned untransformed counts. The red line represents the predicted counts. The gray bar represents the observed counts. Bars above zero on the *y*-axis represent overprediction. Bars below zero on the *x*-axis represent underprediction. CSPEng = sentences with code-switched parts in the English samples; CSPSp = sentences with code-switched parts in the Spanish samples; CSSSp = code-switched sentences in the Spanish samples; CSWEng = code-switched words in the English samples; CSWSp = code-switched words in the Spanish samples.



and their peers with TLD are shown in Tables 1 and 2. Furthermore, the total number of words per language sample across the TLD and DLD groups was estimated and compared. No significant differences were noted at $a = .05$ across groups in the Spanish samples (TLD: $M = 193.7$, $SD = 74.96$; DLD: $M = 201.46$, $SD = 132.1$) and in the English samples (TLD: $M = 242.95$, $SD = 77.78$; DLD: $M = 202.46$, $SD = 76.4$). The group with TLD showed overall higher language proficiency than the group with DLD in Spanish (TLD: $M = 3.7$, $SD = 0.77$; DLD: $M = 3.13$, $SD = 0.77$) and in English (TLD: $M = 3.6$, $SD = 0.70$; DLD: $M = 3.17$, $SD = 0.74$). As anticipated, there were Spanish and English proficiency associations in the DLD group, but not for the TLD group. Children with DLD are expected to have difficulties with language learning, which typically shows as low performance in both languages, whereas for children with TLD, variability in language proficiency patterns across languages is expected.

Regarding the prevalence of children who code-switched, 52.6% (20/38) of children with TLD and 70.8% (17/24) of children with DLD code-switched at least once across the Spanish and English samples; this difference was not statistically significant, $\chi^2(1, N = 62) = 2.03$, $p = .15$. Regarding the prevalence of children who code-switched depending on the target language, in the Spanish samples, a total of 59.68% of children code-switched (37/62)—52.6% (20/38) of children with TLD and 70.8% (17/24) of children with DLD. In the English samples, 9.68% (6/62) of children code-switched—7.9% (3/38) of children with TLD and 12.5% (3/24) of children with DLD. Overall, more children code-switched in the Spanish samples into English than in the English samples into Spanish, and the difference was statistically significant, $\chi^2(1, N = 62) = 34.21$, $p < .01$. The same pattern was observed for the group with DLD, $\chi^2(1, N = 24) = 16.8$, $p < .01$, and their peers with TLD, $\chi^2(1, N = 38) = 18.02$, $p < .01$; for both groups, more children code-switched into English when Spanish was the target language than into Spanish when English was the target language. Finally, paired-samples *t* tests indicated that both children with DLD and their peers with TLD used higher proportions of CSP compared to CSS ($p = .013$ and $p = .006$, respectively) in Spanish language samples. In English samples, the frequency of code-switches to Spanish was very low for both intra- and intersentential types (see Table 2), and the differences were not statistically significant ($p = .17$ and $p = .21$, for the groups with DLD and TLD, respectively).

Language Proficiency and Code-Switching

With respect to the relationship between Spanish and English language proficiency and the frequency of code-switches, for the Spanish language samples, Spanish language proficiency was significantly and negatively associated with

the percentage of CSW ($r_s = -.70$, $p < .01$), intrasentential code-switching as measured by the percentage of CSP ($r_s = -.42$, $p = .04$), and intersentential code-switching as measured by the percentage of CSS ($r_s = -.41$, $p = .05$) into English for the group with DLD. So, lower proficiency in Spanish was associated with higher frequency of code-switches into English regardless of the type of code-switch. Spanish proficiency was not associated with code-switches into English in the TLD group. English language proficiency was not associated with the frequency of code-switches to English in either of the groups (TLD or DLD). Thus, in the Spanish samples, children with DLD code-switched to English, depending on their Spanish proficiency, whereas children with TLD switched to English regardless of their Spanish and English language proficiency. Regarding the English language samples, Spanish and English language proficiency was not associated with code-switching for either group. The frequency of code-switches to Spanish from English was low, regardless of language proficiency, across types of code-switches.

Effects of Diagnosis (DLD and TLD) on Frequency of Code-Switches

With respect to differences in the frequency of CSW, CSP, and CSS across groups with TLD and DLD, Poisson regression was used as a starting model of all six dependent variables (CSW, CSP, CSS in Spanish and English samples). Next, negative binomial regression models were conducted to account for overdispersion in the data (i.e., variance substantially greater than the mean) for all variables, except for CSS in Spanish for which the variance was similar to the mean, and therefore, a zero-inflated Poisson model was used as a second step.

The Poisson models did not adequately fit the data, as they failed to account for the overdispersion of the observed counts; however, the negative binomial models and the zero-inflated Poisson regression for CSS in Spanish fit reliably better than the Poisson models and demonstrated adequate goodness of fit (see Table 4). Visual inspection of rootogram plots of the negative binomial models (see Figure 1) indicated mild-to-moderate underprediction and overprediction across variables, as anticipated, given that this study focused on the relationship between diagnosis and frequency of code-switches and not all factors affecting code-switches. Unstandardized regression coefficients (see Table 5) were statistically significant for Spanish language proficiency in the Spanish samples and English language proficiency in the English samples. Furthermore, exponentiation of the models with significant coefficients showed dependent variables to interact with language proficiency at differing magnitudes. For example, for every one unit decrease in Spanish language proficiency, the number of CSW to English in Spanish samples (CSWSp) was predicted to increase at the steepest rate by approximately 83% ($e^{b_{\text{LPSpa}}} = 0.17$, 95% CI [0.09, 0.29]). Predicted rate change was smaller for the remaining dependent variables. Also, upon inspection, the uncertainty reflected in the width of the confidence intervals was considerably different across models. Regression coefficients were not statistically significant for diagnosis (TLD and DLD). Children's diagnosis had no effect on the frequency of code-switches, after controlling for language proficiency and language sample length. There were also no statistically significant cross-linguistic influences. English proficiency did not affect the frequency of code-switches in the Spanish samples, and Spanish proficiency did not affect the frequency of code-switches in the English samples.

Discussion

This study examined the prevalence of children with DLD and their peers with TLD who code-switch the relationship between language proficiency and frequency of code-switches in these groups and between group differences in the frequency of code-switching, controlling for language proficiency and variations in language sample length. Results indicated that substantially more children code-switched to English when Spanish was the target language—52.6% (20/38) of children with TLD and 70.8% (17/24) of children with DLD—than to Spanish when English was the target language—7.9% (3/38) of children with TLD and 12.5% (3/24) of children with DLD. Correlational analyses indicated that, when Spanish was the target language, lower proficiency in Spanish was associated with higher frequency of code-switches (CSW, CSP, and CSS) to English for the group with DLD, but not for their peers with TLD who code-switched to English regardless of their Spanish and English proficiency. When English was the target language, the frequency of code-switches was very low across the groups with DLD and TLD, and Spanish and English language proficiency was not associated with code-switching within groups. Also, there were no statistically significant effects of children's diagnosis on the frequency of code-switches, when controlling for Spanish and English language proficiency and accounting for variability in language sample length across children. Spanish proficiency was the only significant predictor for code-switches to English, and English proficiency was the only significant predictor for code-switches to Spanish for all measures of code-switching (CSW, CSP, CSS). Nevertheless, in the English samples, only 9.68% (6/62) of participants code-switched to Spanish versus 59.68% (37/62) of participants code-switching in the Spanish samples to English, and the frequency of code-switches to Spanish was low regardless of English proficiency levels and diagnosis (DLD, %CSW to Spanish: $M = 0.27$, $SD = 1.07$; %CSW to English: $M = 12.11$, $SD = 21.79$; TLD, %CSW to Spanish: $M = 0.12$, $SD = 0.55$; %CSW to English: $M = 4.65$, $SD = 17.75$).

Code-Switching and Language Proficiency

The finding that children—both those with DLD and their peers with TLD—demonstrated a low frequency of code-switches to Spanish regardless of their proficiency

Table 4. Comparison of fit statistics for the Poisson and negative binomial or zero-inflated (ZI) Poisson models.

Model	AIC	Log *L*	Predicted no. of zeros	Observed no. of zeros	Goodness of fit χ^2 (*df* = 105)	Goodness of fit *p*	Relative fit χ^2 ΔLog *L* (*df* = 1)	Relative fit *p*	Relative fit Vuong *p*
CSWSp									
Poisson	702	–347.36	5		565	< .001	401.45	< .001	
Negative binomial	303	–146.63	25	25	60.23	.32			
CSPSp									
Poisson	557	–274.83	13	31	456.53	< .001	322.78	< .001	
Negative binomial	237	–113.43	31		54.13	.55			
CSSSp									
Poisson	98	–45.18	47	53	67.39	.14			.046
ZI Poisson	76	–30.08	52						
CSWEng									
Poisson	147	–69.68	40	56	122.25	< .001	77.9	< .001	
Negative binomial	71	–30.73	54		15.198	.99			
CSPEng									
Poisson	92	–42.2	47	56	68.97	.11	30.39	< .001	
Negative binomial	64	–27	54		16.455	.99			

Note. Goodness of fit was calculated using χ^2 goodness of fit of the residual deviance. For nested models, relative fit was calculated taking the χ^2 goodness of fit for the difference in log-likelihoods between a given Poisson model and its negative binomial model counterpart. Relative fit for the Poisson and ZI Poisson nonnested models was estimated using the Vuong *p* value. AIC = Akaike information criterion; Log *L* = log likelihood; Δ = difference; CSWSp = code-switched words in the Spanish samples; CSPSp = code-switched parts of sentences/C-units in the Spanish samples; CSSSp = code-switched sentences/C-units in the Spanish samples; CSWEng = code-switched words in the English samples; CSPEng = code-switched parts of sentences/C-units in the English samples. There were no or very few instances of code-switched sentences in the English samples; therefore, analyses could not be conducted.

in English is consistent with Gutiérrez-Clellen et al.'s (2009) findings that the Spanish-dominant children in their study showed a very low frequency of code-switches to Spanish when tested in their nondominant English. This study included Spanish-dominant, English-dominant, and balanced bilinguals, and the results strengthen the postulation that, at least in some contexts, sociocultural factors might be stronger than psycholinguistic factors. In Greene et al.'s (2012) study as well, the frequency of code-switches to Spanish was low across the different language dominance groups based on language exposure and use, and Montanari et al. (2019) did not find a relationship between proficiency measures and code-switching in Spanish–English-speaking preschoolers. As Gutiérrez-Clellen et al. argued, the frequency of code-switches may be highly affected by children's awareness of the language prescribed by the majority culture. This study indicates that this might be the case even for children in Spanish dual immersion programs (90% or 80% Spanish

Table 5. Results summary of the negative binomial and zero-inflated models.

Parameters	CSWSp	CSPSp	CSSSp	CSWEng	CSPEng
Unstandardized coefficients					
Intercept *a*	0.58 (–2.81, 3.97)	–1.59 (–6.14, 2.69)	–3.31 (–12.00, 5.39)	–6.48 (–17.00, 1.63)	–6.075 (–15.27, 3.06)
b_{LI}	0.66 (–0.20, 1.55)	0.90 (–0.14, 2.01)	–0.01 (–2.34, 2.32)	0.48 (–1.87, 3.035)	0.73 (–1.42, 2.90)
b_{LPEng}	0.26 (–0.38, 0.91)	0.09 (–0.66, 0.86)	1.34 (–0.22, 2.90)	–1.99 (–3.73, –.24)*	–1.95 (–4.28, –0.22)*
b_{LPSpa}	–1.79 (–2.44, –1.23)*	–1.28 (–2.12, –0.46)*	–2.19 (–3.49, –0.90)*	1.66 (–1.07, 5.66)	1.27 (–0.82, 4.19)
Incident rate ratios					
e^a					
$e^{b_{LI}}$					
$e^{b_{LPEng}}$				0.13 (0.02, 0.78)	0.14 (0.01, 0.80)
$e^{b_{LPSpa}}$	0.17 (0.09, 0.29)	0.28 (0.12, 0.64)	0.11 (0.03, 0.41)		

Note. Confidence intervals in parentheses are based on the likelihood profile approach. Incident rate ratios are exponentiated coefficients, where the value represents the mean rate change given one unit increase in the predictor. Incident rate ratios for nonsignificant coefficients were not calculated. Count model coefficients (Poisson with log link) are presented. To control for variations in language sample length across participants, the number of total words in the Spanish or English samples was used as the setoff variable for the Spanish and English dependent variables, respectively. CSWSp = code-switched words in the Spanish samples; CSPSp = sentences with code-switched parts in the Spanish samples; CSSSp = code-switched sentences in the Spanish samples; CSWEng = code-switched words in the English samples; CSPEng = sentences with code-switched parts in the English samples; LPEng = English language proficiency; LPSp = Spanish language proficiency.

*$p < .001$.

and 10% or 20% English in kindergarten and first grade, respectively), when the broader sociopolitical context does not support the home language (Comeau et al., 2003; Garcia et al., 1983).

The findings of this study are not in line with the direction of code-switches predicted by the inhibitory control model for bilingual production (Meuter & Allport, 1999). According to this model, it is more difficult to code-switch when speaking the weaker language (i.e., asymmetrical switching costs) because more effort is needed to return to the strong suppressed language due to inertia—persistence of active suppression of the stronger language. Thus, fewer switches are expected when speaking the weaker language. Nevertheless, in the current study, both children with DLD and their peers with TLD demonstrated low frequency of code-switches to Spanish regardless of their proficiency in English. Our findings are consistent with studies indicating no remarkable relationships between code-switching and proficiency or dominance (Greene et al., 2012; Montanari et al., 2019). Discrepancies across studies could be due to task characteristics and other methodological differences.

Prevalence of Children Who Code-Switched

In the current study, we replicated Greene et al.'s (2012) finding that a large percentage of children code-switched to English—more than half of children with DLD and their peers with TLD. Findings indicate that code-switching is a common phenomenon in testing conditions even when children are instructed to use one of their two languages, in both standardized testing and language sample analysis assessment contexts. In both studies, more children code-switched from Spanish to English than vice versa, but in this study, the differences were larger. In this study, in the Spanish samples, 59.68% (37/62) of children code-switched to English, and in the English samples, only 9.68 (6/62) of children code-switched to Spanish. In Greene et al., in the Spanish subtest, 25.1% of participants code-switched to English, and 18.6% of participants code-switched to Spanish. Differences might be due to different language elicitation tasks across studies. This study used a story retell task, whereas Greene et al.'s study used a single-word naming task. This is in line with Gutiérrez-Clellen et al.'s (2009) results that the method of elicitation could affect the frequency of code-switches, with more participants code-switching in less restricted contexts compared to a decontextualized, repetitive, and highly structured task such as confrontation picture naming. Reyes (2004) also found that children tend to code-switch more when changing topics.

The higher prevalence of children who code-switched to English in this study than was found in Greene et al.'s (2012) study may also be due to different participant characteristics with respect to language proficiency, exposure and use across the two studies, and different sociocultural characteristics in the areas where data were collected. Although children in this study were in Spanish dual immersion programs, this study was conducted in the northwestern United States, whereas Greene et al.'s study included participants from Central Texas and northern Utah. Differences in sociocultural perceptions of minority languages in the families' broader communities may significantly affect language preference and code-switching (Artiles et al., 2010; Corona et al., 2017; Paradis & Nicoladis, 2007).

The code-switching patterns between the Spanish minority and English mainstream languages in this study for Spanish–English-speaking children in the United States are consistent with patterns observed in other cultures (e.g., Raichlin et al., 2019; Yip & Matthews, 2006). For example, Raichlin et al. (2019) studied 5- to 7-year-old Russian–Hebrew-speaking sequential bilinguals in Israel and found that children code-switched significantly more to the mainstream Hebrew language (85% of code-switched instances) than to their L1/Russian (15% of code-switched instances) when they retold a story to a bilingual puppet. These authors identified both psycholinguistic (e.g., difficulties in lexical access) and sociolinguistic factors (e.g., switching to accommodate to a listener's identity or language preferences) as responsible for the code-switches. In the same study, children code-switched more frequently when they spoke to a monolingual Hebrew puppet (20.14% of CSS from Hebrew to Russian) than when they spoke to a monolingual Russian puppet (13.49% of CSS from Russian to Hebrew). Altogether, results suggest that a bilingual context might favor more the mainstream language, depending on the socioeconomic characteristics associated with the languages, language status, and the broader sociocultural environment (Artiles et al., 2010; Corona et al., 2017).

Code-Switching and DLD

Similar to Gutiérrez-Clellen et al. (2009), the results from the current study indicated that children's diagnosis did not have an effect on the frequency of intrasentential code-switching. This study lends more support to these findings by replicating them for intersentential code-switching and the total number of CSW. Gutiérrez-Clellen et al.'s findings indicated lack of differences in frequency of code-switches across DLD and TLD groups for both narrative and conversation tasks. In Greene et al.'s (2012) study, the DLD and TLD groups also did not differ in frequency of code-switches on the single-word naming task when they were *Spanish or English dominant*. On the other hand, these findings of a lack of differences contrast with Iluz-Cohen and Walters' (2012) study on English–Hebrew-speaking preschoolers living in English-speaking homes in Israel. Like Gutiérrez-Clellen et al.'s study and this study, Iluz-Cohen and Walters also used a story retell task but found that children with DLD code-switched more frequently than their peers with TLD. The differences in findings could be because Iluz-Cohen and Walters did not control for variability in language proficiency levels. Gutiérrez-Clellen et al.'s and our findings based on story retell also contrast with Greene et al.'s finding that *balanced* bilinguals at risk for DLD code-switched more frequently than their peers with TLD on a single-word naming task. It seems that both

proficiency and task characteristics affect the frequency of code-switches (Gutiérrez-Clellen et al., 2009; Raichlin et al., 2019; Reyes, 2004). It might be that, in a highly structured task such as a single-word naming task, some children with DLD were more pressured to code-switch when they could not find the correct term in the target language compared to TLD and English- and Spanish-dominant peers. Task effects are in line with Raichlin et al.'s (2019) findings, indicating more code-switches in the more structured story retell test than in conversation. In contrast, Gutiérrez-Clellen et al. did not find any differences in the frequency of code-switches between story retell and conversation tasks, but conversational tasks may vary significantly in the alternation of topics, and the number of shifts in conversation topics may affect the frequency of code-switches (Reyes, 2004). Gutiérrez-Clellen et al. found, however, that more children code-switched in the less restricted conversational task than in a narrative task. Altogether, it might be that story retell provides a less restrictive context than a single-word naming task and sufficient alternation of topics. This could provide children enough flexibility to communicate in the language(s) they voluntarily or involuntarily choose, despite their possible language difficulties and even when instructed to use a specific language during testing. In such a context, results from this study indicated no differences in code-switches between children with DLD and their peers with TLD when controlling for proficiency.

Furthermore, lack of differences between children with DLD and their peers with TLD does not support that children with DLD code-switch primarily to fill lexical gaps (Lanvers, 2001; Montanari et al., 2019) or due to inhibition control limitations (Spaulding, 2010) in a story retell task. Results are in line with literature suggesting that language development is not the primary factor affecting code-switches, but that there are a variety of external sociocultural factors that play a significant role in language choice (Becker, 1997; Ribot & Hoff, 2014). Findings are also consistent with studies suggesting that, as children grow, social factors could become stronger in relation to lexical access and language proficiency (Montanari et al., 2019; Reyes, 2004).

The finding of low Spanish proficiency being associated with higher frequency of code-switches to English only for the group with DLD whereas their peers with TLD code-switched to English regardless of their Spanish and English proficiency could be due to lower social skills in children with DLD (Marton et al., 2005). Although both groups code-switched more to the mainstream language than to Spanish, fewer children with DLD did so (52.6%) compared to their peers with TLD (70.8%). Children with DLD may not be as sensitive to sociocultural perceptions of the majority and minority languages as their peers with TLD. Therefore, in children with DLD, possible loss or slow development of their L1 may have a significant effect on the frequency of code-switching to their L2, whereas in children with TLD, sociocultural factors may play a stronger role; thus, children might code-switch to the majority language regardless of their proficiency levels.

Finally, in this study, Spanish proficiency levels in the DLD group had similar associations with the two types of code-switching examined—CSPSp ($r_s = -.42$) and CSSSp ($r_s = -.41$)—and a stronger negative association with CSWSp (−.70). Also, children used more intrasentential than intersentential code-switching in both DLD and TLD groups (see Table 2) in the Spanish samples, where the greatest frequency of code-switches was observed. There were no differences in English samples. Such findings, in combination with the lack of differences between the DLD and TLD groups in the number of intrasentential code-switches, contrast with the postulation that intrasentential code-switching is more difficult (Sankoff & Poplack, 1981). Alternatively, it might be that the possibility of committing errors when talking may not matter as much. It might be that switching to the language that is easier to access at any point might be more cost effective than suppressing the first linguistic elements activated to use a sentence in a single language when speaking to a bilingual interlocuter (Gollan & Ferreira, 2009).

Clinical Implications

In preschool and early school–age children, there is general consistency in findings across studies that grammatical deficits are highly prevalent in monolingual and bilingual children with DLD (e.g., Kapantzoglou et al., 2015; Leonard, 2014; Restrepo, 1998; Simon-Cereijido & Gutiérrez-Clellen, 2007). This study supports that code-switching is not an indicator of DLD in bilinguals because children with DLD code-switched with similar frequency when compared to their peers with TLD. There was variability in the frequency of code-switches across children in both groups, and the frequency of code-switching appeared to depend on factors unrelated to children's language skills.

The large percentage of children that code-switched despite the monolingual models provided in the story retell and the occasional reminders of the target language indicates that code-switching is an essential part of bilinguals' communication in a bilingual context (Genesee & Nicoladis, 2006). Therefore, restrictions to a monolingual mode may have negative effects on children's overall communication and language performance during assessments and instruction (Grosjean, 2001; Meuter & Allport, 1999). Mixed language should be accepted, and clinical and educational decisions should be made considering possible negative effects when restricting bilinguals to one language.

In this study, many more children code-switched to the mainstream language than to their home language, although they were attending Spanish dual immersion programs. This indicates that broader language support might be needed if the goal of the school is to support both languages—perhaps by fostering a broader bilingual community among the families in the community (Delgado-Gaitan, 2001). Lastly, if the goal is bilingualism, children with DLD would benefit from such societal and/or other additional support in their home language based on the outcomes that lower Spanish proficiency in this group was associated with more code-switches to English.

Limitations and Future Directions

This study assessed the prevalence of children who code-switch, the relationship between children's Spanish and English language proficiency and the frequency of code-switches, and differences in the frequency of code-switches between children with DLD and their peers with TLD, controlling for language proficiency levels and some external factors. For example, all children were tested in the school setting, in Spanish dual immersion programs, by bilingual examiners fluent in the target language, of the same gender (females), and rapport between the examiner and the child was established in the beginning of the session in the target language. Nevertheless, the study did not control for the ethnicity and other cultural characteristics of the examiners. Such characteristics may also play a role in the frequency with which children code-switch. Future studies could examine the degree to which such external factors affect the prevalence and frequency of code-switches in the target population, particularly given the small percentage of bilingual speech-language pathologists serving children that speak the same language.

This study was conducted with children who were attending Spanish dual immersion programs in public schools but did not control for the quality of the programs or for the community/neighborhood in which children lived. Given differences across studies that are conducted in different parts of the United States regarding the prevalence of children who code-switch to the majority language, it would be informative to examine how the prevalence of code-switching to the majority language may vary depending on the quality of immersion programs children attend as well as the characteristics of their neighborhood.

Acknowledgments

This project was supported by an internal Faculty Enhancement Grant from Portland State University (Principal Investigator: Maria Kapantzoglou) and the EXITO (Enhancing Cross-Disciplinary Infrastructure and Training at Oregon) Pilot Project Grant Program, an interinstitutional National Institutes of Health BUILD (Building University Infrastructure Leading to Diversity) EXITO grant to Portland State University (Principal Investigator: Carlos Crespo; 2UL1GM118964-06).

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Artiles, A. J., Kozleski, E. B., Trent, S. C., Osher, D., & Ortiz, A. (2010). Justifying and explaining disproportionality, 1968–2008: A critique of underlying views of culture. *Exceptional Children, 76*(3), 279–299. https://doi.org/10.1177/001440291007600303

ASHA Panel on Audiologic Assessment. (1997). *Guidelines for audiologic screening* (No. GL1997-00199). American Speech-Language-Hearing Association.

Bailey, A. L., & Huang, B. H. (2011). Do current English language development/proficiency standards reflect the English needed for success in school. *Language Testing, 28*(3), 343–365. https://doi.org/10.1177/0265532211404187

Barragan, B., Castilla-Earls, A., Martinez-Nieto, L., Restrepo, M. A., & Gray, S. (2018). Performance of low-income dual language learners attending English-only schools on the Clinical Evaluation of Language Fundamentals–Fourth Edition, Spanish. *Language, Speech, and Hearing Services in Schools, 49*(2), 292–305. https://doi.org/10.1044/2017_LSHSS-17-0013

Becker, K. R. (1997). Spanish/English bilingual codeswitching: A syncretic model on JSTOR. *Bilingual Review/La Revista Bilingüe, 22*(1), 3–30.

Bedore, L. M., & Leonard, L. (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., Peña, E. D., Joyner, D., & Macken, C. (2011). Parent and teacher rating of bilingual language proficiency and language development concerns. *International Journal of Bilingual Education and Bilingualism, 14*(5), 489–511. https://doi.org/10.1080/13670050.2010.529102

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

Bishop, D. V., & Leonard, L. B. (Eds.). (2000). *Speech and language impairments in children: Causes, characteristics, intervention and outcome*. Psychology Press.

Comeau, L., Genesee, F., & Lapaquette, L. (2003). The modeling hypothesis and child bilingual codemixing. *International Journal of Bilingualism, 7*(2), 113–126. https://doi.org/10.1177/13670069030070020101

Corona, K., Campos, B., & Chen, C. (2017). Familism is associated with psychological well-being and physical health: Main effects and stress-buffering effects. *Hispanic Journal of Behavioral Sciences, 39*(1), 46–65. https://doi.org/10.1177/0739986316671297

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language, 50*(4), 491–511. https://doi.org/10.1016/j.jml.2004.02.002

Delgado-Gaitan, C. (2001). *The power of community: Mobilizing for family and schooling*. Rowman & Littlefield.

Dollaghan, C. A. (1998). Spoken word recognition in children with and without specific language impairment. *Applied Psycholinguistics, 19*(2), 193–207. https://doi.org/10.1017/S0142716400010031

Dollaghan, C. A., & Horner, E. A. (2011). Bilingual language assessment: A meta-analysis of diagnostic accuracy. *Journal of Speech, Language, and Hearing Research, 54*(4), 1077–1088. https://doi.org/10.1044/1092-4388(2010/10-0093)

Garcia, E., Maez, L., & Gonzalez, G. (1983). Language switching in bilingual children: A national perspective. In E. Garcia (Ed.), *The Mexican-American child: Language, cognition, and social development* (pp. 56–73). Arizona State University.

Genesee, F., & Nicoladis, E. (2006). Bilingual first language acquisition. In E. Hoff & M. Shatz (Eds.), *Handbook of language development* (pp. 1–34). Blackwell. https://doi.org/10.21832/9781853598715-005

Gollan, T. H., & Ferreira, V. S. (2009). Should I stay or should I switch? A cost–benefit analysis of voluntary language switching in young and aging bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(3), 640–665. https://doi.org/10.1037/a0014981

Greene, K. J., Peña, E. D., & Bedore, L. M. (2012). Lexical choice and language selection in bilingual preschoolers. *Child Language Teaching and Therapy, 29*(1), 27–39. https://doi.org/10.1177/0265659012459743

Grosjean, F. (1997). Processing mixed language: Issues, findings and models. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 225–254). Erlbaum.

Grosjean, F. (2001). The bilingual's language modes. In J. L. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 1–22). Blackwell.

Gutiérrez-Clellen, V. F., & Kreiter, J. (2003). Understanding child bilingual acquisition using parent and teacher reports. *Applied Psycholinguistics, 24*(2), 267–288. https://doi.org/10.1017/S0142716403000158

Gutiérrez-Clellen, V. F., Simon-Cereijido, G., & Leone, A. E. (2009). Code-switching in bilingual children with specific language impairment. *International Journal of Bilingualism, 13*(1), 91–109. https://doi.org/10.1177/1367006909103530

Haubo, R., Christensen, B., & Hansen, M. K. (2014). *binomTools: Performing diagnostics on binomial regression models*. https://R-Forge.R-project.org/projects/binomtools/

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

Iluz-Cohen, P., & Walters, J. (2012). Telling stories in two languages: Narratives of bilingual preschool children with typical and impaired language. *Bilingualism: Language and Cognition, 15*(1), 58–74. https://doi.org/10.1017/S1366728911000538

Jackson-Maldonado, D., Maldonado, R., & Thal, D. J. (1998). Reflexive and middle markers in early child language acquisition: Evidence from Mexican Spanish. *First Language, 18*(54), 403–429. https://doi.org/10.1177/014272379801805407

Kapantzoglou, M., Fergadiotis, G., & Restrepo, M. A. (2017). Language sample analysis and elicitation technique effects in bilingual children with and without language impairment. *Journal of Speech, Language, and Hearing Research, 60*(10), 2852–2864. https://doi.org/10.1044/2017_JSLHR-L-16-0335

Kapantzoglou, M., Restrepo, M. A., Gray, S., & Thompson, M. S. (2015). Language ability groups in bilingual children: A latent profile analysis. *Journal of Speech, Language, and Hearing Research, 58*(5), 1549–1562. https://doi.org/10.1044/2015_JSLHR-L-14-0290

Kapantzoglou, M., Thompson, M. S., Gray, S., & Restrepo, M. A. (2016). Assessing measurement invariance for Spanish sentence repetition and morphology elicitation tasks. *Journal of Speech, Language, and Hearing Research, 59*(2), 254–266. https://doi.org/10.1044/2015_JSLHR-L-14-0319

Kline, R. B. (2010). *Principles and practice of structural equation modeling* (3rd ed.). Guilford Press.

Kohnert, K. (2010). Bilingual children with primary language impairment: Issues, evidence and implications for clinical actions. *Journal of Communication Disorders, 43*(6), 456–473. https://doi.org/10.1016/j.jcomdis.2010.02.002

Lanvers, U. (2001). Language alternation in infant bilinguals: A developmental approach to codeswitching. *International Journal of Bilingualism, 5*(4), 437–464. https://doi.org/10.1177/13670069010050040301

Leonard, L. B. (2014). *Children with specific language impairment*. MIT Press.

Little, R. J., & Rubin, D. B. (1989). The analysis of social science data with missing values. *Sociological Methods & Research, 18*(2–3), 292–326. https://doi.org/10.1177/0049124189018002004

MacSwan, J. (2000). The threshold hypothesis, semilingualism, and other contributions to a deficit view of linguistic minorities. *Hispanic Journal of Behavioral Sciences, 22*(1), 3–45. https://doi.org/10.1177/0739986300221001

MacSwan, J., & Rolstad, K. (2006). How language proficiency tests mislead us about ability: Implications for English language learner placement in special education. *Teachers College Record, 108*(11), 2304–2328.

Marchman, V. A., & Martínez-Sussmann, C. (2002). Concurrent validity of caregiver/parent report measures of language for children who are learning both English and Spanish. *Journal of Speech, Language, and Hearing Research, 45*(5), 983–997. https://doi.org/10.1044/1092-4388(2002/080)

Marsh, H. W., Hau, K.-T., & Grayson, D. (2005). Goodness of fit in structural equation models. In A. Maydeu-Olivares & J. J. McArdle (Eds.), *Multivariate applications book series. Contemporary psychometrics: A festschrift for Roderick P. McDonald* (pp. 275–340). Erlbaum.

Marton, K., Abramoff, B., & Rosenzweig, S. (2005). Social cognition and language in children with specific language impairment (SLI). *Journal of Communication Disorders, 38*(2), 143–162. https://doi.org/10.1016/j.jcomdis.2004.06.003

Mayer, M. (1967). *A boy, a dog, and a frog*. Dial Press.

Mayer, M. (1973). *Frog on his own*. Dial Press.

Meuter, R. F., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language, 40*(1), 25–40. https://doi.org/10.1006/jmla.1998.2602

Miccio, A. W., Hammer, C. S., & Rodríguez, B. L. (2009). Code-switching and language disorders in bilingual children. In B. E. Bullock & A. J. Toribio (Eds.), *The Cambridge handbook of linguistic code-switching* (pp. 241–252). Cambridge University Press. https://doi.org/10.1017/CBO9780511576331.015

Miller, J. F., & Iglesias, A. (2012). *Systematic Analysis of Language Transcripts (SALT), Research Version 2012*. SALT Software, LLC.

Montanari, S., Ochoa, W., & Subrahmanyam, K. (2019). A longitudinal investigation of language mixing in Spanish–English dual language learners: The role of language proficiency, variability, and sociolinguistic factors. *Journal of Child Language, 46*(5), 913–937. https://doi.org/10.1017/S0305000919000278

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

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAD in instructed SLA: The case of complexity. *Applied Linguistics, 30*(4), 555–578. https://doi.org/10.1093/applin/amp044

Paradis, J., & Nicoladis, E. (2007). The influence of dominance and sociolinguistic context on bilingual preschoolers' language choice. *International Journal of Bilingual Education and Bilingualism, 10*(3), 277–297. https://doi.org/10.2167/beb444.0

Peña, E. D., Bedore, L. M., Gutierrez-Clellen, V. F., Iglesias, A., & Goldstein, B. A. (2010). *Bilingual English–Spanish Oral Screener (BESOS): Pre-K, 1st, and 3rd grade (Unpublished test)*.

Peña, E. D., Gutiérrez-Clellen, V., Iglesias, A., Goldstein, B., & Bedore, L. M. (2014). *BESA: Bilingual English–Spanish Assessment manual*. AR-Clinical.

Peña, E. D., Gutiérrez-Clellen, V. F., Iglesias, A., Goldstein, B. A., & Bedore, L. M. (2018). *Bilingual English Spanish Assessment (BESA)*. Brookes.

Poulisse, N., & Bongaerts, T. (1994). First language use in second language production. *Applied Linguistics, 15*(1), 36–57. https://doi.org/10.1093/applin/15.1.36

Pratt, A. S., Adams, A. M., Peña, E. D., & Bedore, L. M. (2020). Exploring the use of parent and teacher questionnaires to screen for language disorders in bilingual children. *Topics in Early Childhood Special Education*. Advance online publication. https://doi.org/10.1177/0271121420942308

Pray, L. (2005). How well do commonly used language instruments measure English oral-language proficiency? *Bilingual Research Journal, 29*(2), 387–409. https://doi.org/10.1080/15235882.2005.10162841

Raichlin, R., Walters, J., & Altman, C. (2019). Some wheres and whys in bilingual codeswitching: Directionality, motivation and locus of codeswitching in Russian–Hebrew bilingual children. *International Journal of Bilingualism, 23*(2), 629–650. https://doi.org/10.1177/1367006918763135

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

Reyes, I. (2004). Functions of code switching in schoolchildren's conversations. *Bilingual Research Journal, 28*(1), 77–98. https://doi.org/10.1080/15235882.2004.10162613

Ribot, K. M., & Hoff, E. (2014). "¿Cómo estás?" "I'm good." Conversational code-switching is related to profiles of expressive and receptive proficiency in Spanish–English bilingual toddlers. *International Journal of Behavioral Development, 38*(4), 333–341. https://doi.org/10.1177/0165025414533225

Rice, M. L., Wexler, K., & Cleave, P. L. (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech and Hearing Research, 38*(4), 850–863. https://doi.org/10.1044/jshr.3804.850

Rojas, R., Iglesias, A., Bunta, F., Goldstein, B., Goldenberg, C., & Reese, L. (2016). Interlocutor differential effects on the expressive language skills of Spanish-speaking English learners. *International Journal of Speech-Language Pathology, 18*(2), 166–177. https://doi.org/10.3109/17549507.2015.1081290

Rubin, D. B. (1976). Inference and missing data. *Biometrika, 63*(3), 581–592. https://doi.org/10.1093/biomet/63.3.581

Sankoff, D., & Poplack, S. (1981). A formal grammar for code-switching. *Paper in Linguistics, 14*(1), 3–45. https://doi.org/10.1080/08351818109370523

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

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics, 30*(4), 510–532. https://doi.org/10.1093/applin/amp047

Smyk, E., Restrepo, M. A., Gorin, J. S., & Gray, S. (2013). Development and validation of the Spanish–English Language Proficiency Scale (SELPS). *Language, Speech, and Hearing Services in Schools, 44*(3), 252–265. https://doi.org/10.1044/0161-1461(2013/12-0074)

Spaulding, T. J. (2010). Investigating mechanisms of suppression in preschool children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 53*(3), 725–738. https://doi.org/10.1044/1092-4388(2009/09-0041)

Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Erlbaum. https://doi.org/10.4324/9781410604491

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (6th ed.). Allyn & Bacon.

Treffers-Daller, J. (2019). What defines language dominance in bilinguals? *Annual Review of Linguistics, 5,* 375–393. https://doi.org/10.1146/annurev-linguistics-011817-045554

Walters, J. (2005). Bilingualism: The sociopragmatic–psycholinguistic interface. In E. Nicoladis & S. Montanari (Eds.), *Bilingualism across the lifespan: Factors moderating language proficiency*. Walter de Gruyter.

Wiig, E. H., Secord, W. A., & Semel, E. (2013). *Clinical Evaluation of Language Fundamentals–Fifth Edition: CELF-5*. Pearson.

Wiig, E. H., Semel, E., & Secord, W. (2006). *Clinical Evaluation of Language Fundamentals-Spanish (CELF-4 Spanish)*. The Psychological Corporation.

Woolford, E. (1983). Bilingual code-switching and syntactic theory. *Linguistic Inquiry, 14*(3), 520–536. https://www.jstor.org/stable/4178342

Yip, V., & Matthews, S. (2006). Assessing language dominance in bilingual acquisition: A case for mean length utterance differentials. *Language Assessment Quarterly, 3*(2), 97–116. https://doi.org/10.1207/s15434311laq0302_2

Yow, W. Q., Patrycia, F., & Flynn, S. (2016). Code-switching in childhood. In E. Nicoladis & S. Montanari (Eds.), *Lifespan perspectives on bilingualism* (pp. 81–100). American Psychological Association. https://doi.org/10.1037/14939-006

Zentella, A. C. (1990). Integrating qualitative and quantitative methods in the study of bilingual code switching. *Annals of the New York Academy of Sciences, 583*(1), 75–92. https://doi.org/10.1111/j.1749-6632.1990.tb12186.x