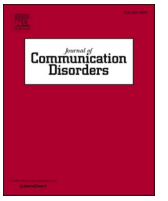
Journal of Communication Disorders 92 (2021) 106110

Contents lists available at ScienceDirect

Journal of Communication Disorders

journal homepage: www.elsevier.com/locate/jcomdis



A Longitudinal Study of Spanish Language Growth and Loss in Young Spanish-English Bilingual Children

Lindsey Hiebert [a,#,*], Raúl Rojas [b]

[a] Missouri State University
[b] University of Texas at Dallas

ARTICLE INFO

Keywords:
Bilingual
Longitudinal
Child
Language
Code Switching
Loss

ABSTRACT

This longitudinal study examined trajectories of Spanish language growth and loss in 34 Spanish-English bilingual children attending an English immersion school. Narrative retell language samples were collected in Spanish across 3 years using wordless, picture storybooks. Digital audio recordings were transcribed, coded, and analyzed for mean length of utterance in words, proportion of grammatical utterances, and moving-average type-token ratio. Code switching into English was also coded at the word level to determine its potential impact on moving-average type-token ratio. Growth curve models were used to estimate the change over time for each outcome measure. The findings indicated that the Spanish-English bilingual participants who attended an English immersion school demonstrated loss of Spanish grammatical and lexical production (as defined by encompassing maintenance and or significant deceleration) from preschool through kindergarten, and that the degree of loss in lexical production was impacted by whether code switching was included or excluded. The findings are discussed in the context of clinical decision-making when assessing the Spanish expressive language abilities of this specific population.

1. [1]Introduction

The Hispanic population is one of the fastest growing minority populations in the United States and comprises the largest percentage (67%) of the U.S. population over the age of five who speak a language other than English at home (Ryan, 2013). Although not all Hispanics speak Spanish, the majority (>60%) of the Hispanic population does (Ryan, 2013), and particularly so in major urban population centers. For instance, 89% of the Hispanic population in Dallas, Texas is reported to speak Spanish at home (Ennis, Ríos-Vargas, & Albert, 2011). The rapidly expanding Hispanic population is projected to make up 36% of all children enrolled in U.S. public schools by 2050, a percentage equal to that of white, non-Hispanic, school-age children in the same year (Forum on Child and Family Statistics, 2015). Nearly 25 percent of the U.S. Hispanic population is living below the poverty level, as compared to just over

* Corresponding author: Lindsey Hiebert, Missouri State University, Department of Communication Sciences and Disorders, 901 S. National Ave. Springfield, MO 65897, U.S.A.
E-mail address: lhiebert@missouristate.edu (L. Hiebert).
[#] Lindsey Hiebert, Department of Communication Sciences and Disorders, Missouri State University; Raúl Rojas, Department of Speech, Language, and Hearing, University of Texas at Dallas.

[1] LSA = language sample analysis; MLUw = mean length of utterance in words; PGU = percent of grammatical utterances; MATTR = moving-average type-token ratio; NTW = number of total words; GCM = growth curve model; UM = unconditional means

https://doi.org/10.1016/j.jcomdis.2021.106110
Received 14 November 2019; Received in revised form 23 March 2021; Accepted 10 April 2021
Available online 14 May 2021

10 percent of the white, non-Hispanic population (Macartney, Bishaw, & Fontenot, 2013). This growing student population is at an elevated risk for failing statewide tests and reading assessments, and is more likely to drop out of school due to difficulties with overall academic success (see Anderson, 2012 for a review; August & Shanahan, 2006; Forum on Child and Family Statistics, 2015; Francis et al., 2005; Han, 2012; Hansen, 2005; U.S. Census, 2010; Winsler et al., 2014). Additionally, Hispanic Spanish-English bilingual children (henceforth referred to as Spanish-English bilingual) are more likely to be misdiagnosed when assessed for learning disabilities or language impairment (Bedore & Peña; 2008; see Hammer et al., 2014, for a review).

Due to this growing population's apparent elevated risk of academic failure and clinical misdiagnosis, the need for additional research concerning their language development is pressing, and particularly so from a longitudinal, data-driven perspective (see Hammer et al., 2014, for a review; see Slavin & Cheung, 2005 for a review). One possible factor implicated in the academic challenges experienced by many Spanish-English bilingual children is a phenomenon that has been referred to as first language loss (Anderson 1999a, 1999b, 2001). First language will henceforth be referred to as the minority language (August & Shanahan, 2006; Gámez & González, 2017; see Iglesias & Rojas, 2012 for a review). The purpose of this longitudinal study was to estimate the dynamic changes of expressive Spanish language skills in Spanish-English bilingual children including trajectories of language growth and loss.

1.1. Longitudinal Studies of Bilingual Language Development

Although language loss is a phenomenon that has been shown to occur in bilingual children (see Anderson, 2012 for a review), the majority of longitudinal studies of Spanish-English bilingual language development predominantly show growth (Han 2012; Hoff & Ribot, 2017; Rojas & Iglesias, 2013; Simon-Cereijido, Gutiérrez-Clellen, & Sweet, 2013; Winsler et al., 2014). Studies that reported growth in the minority language have focused on expressive language skills. Studies that have examined measures such as picture vocabulary tests (Hoff & Ribot, 2017) and number of different words (Rojas & Iglesias, 2013; Simon-Cereijido et al., 2013) have shown that as a whole, Spanish-English bilingual children exhibit growth in lexical skills over time without controlling for the level of support in Spanish. In addition to measures of lexical diversity, gross complexity of expressive language skills, measured by mean length of utterance has also shown growth over time in larger numbers of participants (Rojas & Iglesias, 2013; Simon-Cereijido, Gutiérrez--Clellen, & Sweet, 2013). In contrast to the number of studies of minority language that show overall growth in expressive language skills of Spanish-English bilingual children, was a study that indicated there may be a deceleration of Spanish grammar skills in narrative language samples (Castilla-Earls, Francis, Iglesias, & Davidson, 2019).

Spanish-English bilingual speakers are inherently more heterogeneous than monolinguals (see Hammer et al., 2014 for a review; Han, 2012) and the possibility that Spanish loss may have had a negative impact on participants of these studies has not been explored. Another possible limiting factor in the lack of findings for Spanish loss in larger scope studies may be that school-age bilingual children are sometimes unwilling or unable to produce oral language in samples in their minority language (Mishina-Mori, 2011; Orellana, 1994), and therefore would be excluded from measures of their minority language altogether. Additionally, many prior studies of bilingual language development have neglected to report whether words code switched into English have been included or excluded from analyses (e.g., Kapantzoglou et al., 2019), or they have eliminated the code switched utterances altogether (e.g., Bedore, Peña, Gillam, & Ho, 2010; Castilla-Earles et al., 2019; Rojas et al., 2019). This lack of transparency or inclusion of code switched words and/or utterances may in part impact the limited findings on language loss, given that code switching from the minority language to English has been shown to be a primary indicator of language loss (Anderson, 1999a; Kravin, 1992; Guiberson et al., 2006).

Another reason for scarce findings on language loss may be the result of mixed experiences of the participants in studies on bilingual children, such as their academic program(s) of language instruction. There is a lack of research on bilingual language learner's oral language development that specifies and or stratifies the type of academic program of language instruction (e.g., structured English immersion, dual language, transitional bilingual) (see Hammer et al., 2014 for a review; see Slavin & Cheung, 2005 for a review). Structured English immersion programs, which focus nearly exclusively on English language and literacy skills, are the setting in which bilingual children receive the least amount of systematic academic instruction or support in their minority language. Spanish-English bilingual children educated in English only settings fall below monolingual English-speaking peers academically (Hoff, 2013), which is also true of bilingual children who experience language loss (see Anderson, 2012 for a review). A growing body of work indicates that the continued use and academic support of Spanish leads to children having greater success in learning English, and in overall academic achievement in both languages (e.g., Cardenás-Hagan et al., 2007; Dunn-Davidson, 2011; see Guiberson, 2013 for a review; Rojas, Hiebert, Gusewski, & Francis, 2019; Winsler et al., 2014). It may be inferred from such findings that Spanish-English bilingual children in English immersion environments may be at the most elevated risk for Spanish loss.

1.2. Existing Work on Language Loss

Spanish-English bilingual children who experience Spanish language loss are more at risk to fall behind academically, because they experience degradation of the linguistic foundation necessary to build higher-level language and literacy skills (see Anderson, 2012 for a systematic review; Dunn-Davidson, Hammer, & Lawrence, 2011; Menken & Kleyn, 2010). Language loss has been typically defined by the decline of previously acquired skills in the minority language during the process of second language (English) acquisition (Anderson, 1999b; Guiberson, Barrett, Jancosek, & Itano, 2006). Prior studies of language loss have described a decrease of both grammatical and lexical skills in the minority language and an increase of use of general terms or code switching into English over one to three years beginning in preschool (Anderson, 1999a, 1999b, 2001; Guiberson et al., 2006; Kravin, 1992). Guiberson et al. (2006) defined maintenance as the lack of discernible growth of Spanish. Schiff-Myers (1992) referred to arrested development of minority language skills in terms of maintenance or lack of positive growth as language attrition. For the purposes of this study the terms

arrested development, language attrition, and language loss are considered synonymous, because they describe a range of possible growth trajectories including negative growth as well as no significant growth over time.

Because the limited number of research studies on the phenomenon of language loss have predominantly measured expressive language skills over time, it is difficult to determine if the participants in the studies have truly 'lost' their language skills, or if rather these participants gradually lost the ability to effectively communicate expressively in the minority language. Thus, it may be more sensible to portray language loss as a description of expressive language skills that does not necessarily include receptive language skills that may be retained. Henceforth the phenomenon previously referred to as first language loss, attrition, or arrested development of the minority language will be referred to as Spanish (language) loss, which will embody the time course of loss or attrition via expressive language skills. Although the most commonly used term of 'loss' will be used, for the purposes of this study it will encompass a dynamic range and longitudinal progression of this phenomenon in Spanish-English bilinguals, which can span from non-significant growth (maintenance) to significant and negative growth (deceleration) of Spanish expressive language skills and can be longitudinally differentiated from significant and positive growth.

There is a limited number of empirical research articles that have focused specifically on measuring the loss of the minority language in bilingual children in the U.S. This scant body of work has largely been qualitative and is predominantly made up of case studies that measure expressive language skills. There have also been studies that specifically targeted bilingual language development, loss, and shift (change in proficiency/dominance from the minority language to English over time; Wong Fillmore, 1991) but have measured this phenomenon via methods such as standardized assessment or parent survey/interview (e.g., Merino, 1983; Orellana, 1994; Wong Fillmore, 1991). One exception is Castilla-Earls et al. (2019), which measured shift via narrative language sample analysis.

Anderson (1999a, 1999b, 2001) conducted a series of studies on language loss, which together presented detailed qualitative examples of grammatical and lexical language components affected by Spanish loss. Her methods, measures of language loss, and results were similar to those of Kravin (1992). Both of these researchers focused on one (Anderson, 1999a; Kravin, 1992) or two children (Anderson 1999b, 2001). The children who participated in these studies spoke either Spanish (Anderson, 1999a, 1999b, 2001) or Swedish (Kravin, 1992) with their parents, but were exposed primarily to English in school. These studies used conversational language samples collected from one or two children over a 2-year period (mean age 5 years 7 months to 7 years 5 months; Anderson 1999a, 1999b, 2001), or over a 10-month period (age 5 years, 11 months to 6 years, 9 months). Collectively, the outcome measures from the conversational samples of these language loss studies include proportion and type of grammatical errors, proportions of errors in lexical morphology and verb use in the minority language, and proportion of code switching to English. The findings from Anderson's systematic body of work went further by also examining mean length of utterance (MLU; metric unspecified).

Guiberson et al. (2006) tracked Spanish loss in 10 participants (6 girls, 4 boys), but with fewer conversational language samples (1 per year over three years) than the studies by Anderson and Kravin. His study included Spanish-English bilingual preschool students whose ages at onset ranged from 33 to 37 months. Guiberson et al. tracked expressive language outcome measures used in previous studies of language loss (proportion of grammatical errors, and code switching to English), but added the measure of Spanish *D* in place of MLU. Spanish *D* is derived from type-token ratio, which was used as a lexical measure of Spanish loss. The participants from this study demonstrated either maintenance (8 participants) or loss (2 participants) of Spanish.

Castilla-Earls et al. (2019) longitudinally examined the proficiency shift of grammaticality in narrative retell language samples produced by Spanish-English bilingual school-age children from kindergarten to Grade 2 via the proportion of grammatical utterances (PGU). They found that all participants showed a decline in their Spanish grammatical skills over time regardless of the educational program they attended (e.g., English immersion, dual language). This study included the largest number of participants in comparison to the other studies focused on language loss with 1,080 participants. However, this study did not include a measure of lexical diversity.

The general findings from all of the previous studies of language loss indicate that children who experience language loss show increases in some language measures and decreases in others. The aforementioned studies found that loss of the minority language had occurred in their participants as demonstrated by (a) a decrease in minority language verb use (Anderson, 2001; Kravin, 1992), (b) increases in the proportion of grammatical (morphosyntactic) errors (Anderson, 1999a; Castilla-Earls et al., 2019; Kravin, 1992), (c) loss of lexical diversity (Guiberson et al., 2006), (d) use of general terms when speaking in the minority language (Anderson, 1999a, 2001; Guiberson et al., 2006; Kravin, 1992), and (d) increased code switching to English when speaking in the minority language (Anderson, 1999a; Guiberson et al., 2006; Kravin, 1992).

What can be gleaned from prior research on language loss are the methods, and patterns found that can inform and motivate further research on this dynamic phenomenon. In order to more precisely measure patterns of potential Spanish loss, larger participant samples and analytic approaches specifically designed to measure longitudinal change are necessary. Further research using a controlled and stable language sampling approach such as narrative language sample analysis (Heilmann et al., 2008), conducted over a longer period of time, would be necessary to find emerging patterns of Spanish loss that can inform and potentially augment these earlier findings.

1.3. Outcome Measures of Longitudinal Language Growth and Loss in Narrative Retells

Narrative retell language samples have been used extensively in prior work with bilingual children as they provide a functional measure of connected language production that allows for multiple language measures to be analyzed (e.g., Gusewski & Rojas, 2017; Heilmann et al., 2008; Miller et al., 2006). The present study used the oral narrative retell language samples collected to conduct language sample analysis (LSA) as a method of measuring overall language skills of the participants in Spanish over time. Narrative

LSA is recommended for measuring expressive language skills in Spanish-English bilingual children (Castilla-Earls et al., 2020; Heilmann, Miller, & Nockerts, 2010), and serves as an indicator of typical and functional expressive language use for monolingual and Spanish-English bilingual children (e.g., Gutiérrez-Clellen, Restrepo, Bedore, Peña, & Anderson, 2000; Heilmann et al., 2008; Kapantzoglou, Fergadiotis, & Restrepo, 2017). This approach has been successful in systematically mapping overall language growth in large-scale samples of Spanish-English bilingual children across a variety of educational settings (e.g., Castilla-Earls et al., 2019; Rojas et al., 2019). Similarly, it has also been used to track longitudinal language development with comparable sample sizes to the current study (e.g., Ferron et al., 2009; Gusewski & Rojas, 2017; McNeish & Matta, 2018), and can be leveraged to track Spanish loss in a similar population.

The measures used to track Spanish expressive language skills in the present study were chosen based on previous research that also used oral language sample measures (e.g., Anderson, 2001; Bedore et al., 2010; Castilla-Earls et al., 2019; Gámez & González, 2017; Kapantzoglou, Fergadiotis, & Auza Buenavides, 2019; Miller et al., 2006; Rojas & Iglesias, 2013), and included skills that have previously been identified as possible indicators of language loss (Anderson, 1999a, 1999b, 2001; Castilla-Earls et al., 2019; Guiberson et al., 2006; Kravin, 1992). Since there are both grammatical and lexical indicators of Spanish loss (see Anderson, 2012 for a review), grammatical measures of mean length of utterance in words (MLUw), and proportion of grammatical utterances (PGU), as well as the lexical measure moving-average type-token ratio (MATTR) were chosen.

MLUw measures the gross complexity of sentence structures by averaging the length of the children's utterances in words. MLUw has been shown to be a sensitive measure of longitudinal growth in the Spanish and English of Spanish-English bilingual children up to Grade 2 (Rojas & Iglesias, 2013). Further, MLU has been used as a measure of Spanish loss for the majority of existing Spanish language loss case studies (Anderson, 1999a, 1999b, 2001).

PGU indexes the ability to correctly use verb tense and sentence structure. Bedore (2001) found that a range of Spanish morphosyntactic skills (e.g., correct use of articles, gender agreement, object clitic pronouns, tense marking) typically begin to emerge in Spanish-English bilingual children in the U.S. by about the age of 3, an age comparable to that of the participants at the onset of the present longitudinal study. A gradual increase of grammatical errors over time may be indicative of potential Spanish loss, particularly when the errors include forms that should be used correctly (Anderson, 1999a, 1999b, 2001; Castilla-Earls et al., 2019; Guiberson et al., 2006; Shin, 2018).

Lexical measures such as number of different words (Rojas & Iglesias, 2013; Simon-Cereijido, Gutiérrez-Clellen, & Sweet, 2013) and *D* (Guiberson et al., 2006), which is derived from type-token ratio, have been integral in prior studies of longitudinal dual language growth. However, MATTR uses a moving window of a specified number of words (25 for the present study) to create a ratio of number different words to surface forms, which controls for the length of the narrative language sample. Fergadiotis, Wright, and Green (2015) demonstrated that MATTR was a stronger indicator of lexical diversity than *D* in adults (Malvern & Richards, 1997; McKee, Malvern, & Richards, 2000), which was the lexical measure used to track Spanish loss in Guiberson et al. (2006). More recently Kapantzoglou, Fergadiotis, and Auza Buenavides (2019) found that MATTR was the preferred measure of lexical diversity for child narrative language samples as well.

1.4. Research Questions

This longitudinal, 3-year investigation addressed two related research questions regarding expressive Spanish skills of Spanish-English bilingual preschool children, who were typically developing and enrolled in an English immersion program. First, what are the rates of Spanish growth or loss (maintenance or significant and negative growth) across two specific grammatical measures (mean length of utterance in words – MLUw; and proportion of grammatical utterances – PGU)? Second, what are the rates of Spanish growth or loss across the lexical measure of moving-average type-token ratio (MATTR) in Spanish when code switching into English is included and excluded?

Consistent with the existing literature on Spanish language growth and loss, it is hypothesized that bilingual children academically instructed in English only will demonstrate loss as defined by encompassing both maintenance (no significant growth) and significant negative growth of grammatical and lexical measures of language production (see Anderson, 2012 for a review; Guiberson et al., 2006). As code switching has been found to increase over time in children who experience the phenomenon of language loss (Anderson, 1999a, 1999b, 2001; Guiberson et al., 2006; Kravin, 1992), it is also hypothesized that MATTR including code switching in English will either show positive growth or a slower rate of decline when compared to MATTR excluding code switching. Code switching is also predicted to have an impact on grammatical measures that still include words in the non-target language (English). This impact will likely show an increase in the MLUw of some participants who demonstrate high levels of code switching into English. With some participants who will likely demonstrate restricted growth of MLUw, and others that show loss of Spanish expressive skills, the overall trajectory will likely demonstrate maintenance, which is included in the described definition of language loss.

The findings from this study will provide information that may be pertinent to instructional paradigms of Spanish-English bilinguals in English immersion settings, as well as, clinically relevant information that may help to inform language assessment in Spanish-English bilingual children. Furthermore, the findings will provide timely and important information regarding the effects of Spanish loss on one of the fastest growing populations of school-age children in the U.S. (Forum on Child and Family Statistics, 2015; U.S. Census, 2010). This study also introduces a number of improvements that aim to expand the limited Spanish loss knowledge base including longitudinally tracking a larger number of participants, using the specific academic environment of English immersion, using an analytic approach (growth curve modeling) specialized for longitudinal data analysis (Singer & Willet, 2003), and investigating a wider range of expressive language measures collected via the narrative retell elicitation method, which has been shown to be clinically and academically relevant for Spanish-English bilingual children (e.g., Gutiérrez-Clellen, Restrepo, Bedore, Peña, & Anderson,

2000; Heilmann et al., 2008; Kapantzoglou et al., 2017).

2. Method

2.1. Participants and Language Samples

This study used archival data from a larger-scope longitudinal project focused on the developmental growth of dual language and self-regulation skills in preschool- and early school-age Spanish-English bilingual children. The present study received approval from the Office of Research Compliance at The University of Texas at Dallas. The larger project, which is ongoing, uses a multi-year longitudinal design in which preschool children ages 3 to 4 are being followed across several academic years in a variety of programs of language instruction (i.e., English immersion; dual language). Although the larger, ongoing project continues to recruit new cohorts of preschool students, the present study is based on the first three years of narrative data from the initial cohort, which consisted of narrative retell samples in Spanish that were transcribed, cleaned, and checked for reliability.

Participants who met the following criteria were included in the study: (a) their parents reported them to be Spanish-English speaking children; (b) they were at least 3 years of age upon entering preschool; (c) they were enrolled in an English immersion program; (d) they passed the Bilingual English Spanish Oral Screening (BESOS; Peña, Bedore, Iglesias, Gutiérrez-Clellen, & Goldstein, 2008) at the age of 4 indicating low risk for language impairment; (e) their parents did not report any previous or current history of special education services upon enrollment in the study; (f) they were not retained in any academic grade for the duration of the study; and (g) had normal hearing as indicated by passing annual hearing screenings. A total of 37 participants who produced 168 narrative samples in Spanish over a 3-year period, met the aforementioned inclusionary criteria. It should be noted that 18 of the participants were not asked to provide narrative language samples in this study until the fall of preschool-4.

Inclusionary criteria were also applied to individual narrative language samples stipulating that: (a) the language sample contained at least 10 complete and intelligible utterances, which meets or exceeds narrative sample length criteria established in prior studies with Spanish-English bilingual children (e.g., Gusewski & Rojas, 2017; Miller et al., 2006; Rojas & Iglesias, 2013); and (b) 20% or more of the total words in the narrative language sample were produced in Spanish in order to capture potential changes in Spanish production over time that could range from producing narratives primarily in Spanish to primarily in the non-target language (English), which indicates a minimal level of Spanish proficiency (Pearson, Fernandez, Lewedeg, & Oller, 1997). Code switching was coded at the word level, and the proportion was determined by dividing the total number of English (non-target language) words produced by the total number of words in the narrative including both Spanish and English words. The proportion of code switching in words and utterances is reported in Table 1. Eleven narrative samples were not successfully collected. An additional 26 narrative samples did not meet the inclusionary criteria due to containing too few utterances or greater than 80% of all words were in English. The exclusion of these 37 narrative samples reduced the total number of participants from 37 to 34, who produced a total of 131 narrative language samples in Spanish over 6 semesters (waves of data collection).

The final sample of participants in the present study included 34 (15 male, 19 female) Spanish-English bilingual children who attended an English immersion school in Dallas, Texas and were followed over a 3-year period from preschool to kindergarten. Sixteen participants were followed from the fall of preschool-3, and 18 participants were followed from the fall of preschool-4. The mean age of the participants at the onset of data collection was 51.2 months of age. This study focused on Spanish-English bilingual participants who were enrolled in this English immersion school for a duration of six consecutive academic semesters to capture dynamic changes in Spanish development experienced within this specific type of academic setting. The participants attended a school where the majority of the students were Hispanic (94%). All but one of the participants qualified for free or reduced lunch, indicative of 97% of the participants residing in households that fell below the federal poverty income eligibility guidelines established by the U.S. Department of Agriculture.

Participant recruitment and enrollment included a questionnaire for parents, which they completed in the preferred language (see Supplemental Material A.1-A.2) with one-on-one personal assistance provided by Spanish-English bilingual examiners. The

Table 1
Participant and Spanish Language Sample Analysis (LSA) Descriptive Statistics

	PreK-3 Fall	PreK-3 Spring	PreK-4 Fall	PreK-4 Spring	Kinder Fall	Kinder Spring
Language Samples	9	11	27	24	29	31
NTW	80.0 (55.9)	96.0 (50.2)	146.2 (63.0)	154.57 (61.3)	186.7 (80.4)	167.1 (68.4)
% CS English words	21.1 (24.9)	16.7 (30.4)	26.5 (38.0)	26.7 (34.2)	25.2 (31.7)	22.3 (27.7)
% Utts with CS utterances	32.9 (31.9)	21.0 (34.4)	30.8 (39.1)	30.0 (32.5)	32.8 (37.3)	36.2 (36.6)
Spanish LSA measures						
MLUw	4.22 (2.17)	4.48 (2.36)	5.57 (4.36)	6.25 (3.83)	6.20 (4.65)	6.05 (4.85)
PGU	0.71 (0.20)	0.74 (0.17)	0.73 (0.12)	0.70 (0.17)	0.68 (0.18)	0.70 (0.21)
MATTR (including CS)	0.67 (0.11)	0.75 (0.03)	0.72 (0.12)	0.71 (0.08)	0.69 (0.13)	0.66 (0.13)
MATTR (excluding CS)	0.69 (0.17)	0.76 (0.06)	0.70 (0.15)	0.67 (0.12)	0.64 (0.17)	0.64 (0.15)

Note. PreK-3 = preschool-3; PreK-4 = preschool-4; Kinder = kindergarten; NTW = number of total words; % CS English words = proportion of code switched English words; % Utts with CS utterances = proportion of utterances with code switched English words; MLUw = mean length of utterance in words; PGU = proportion of grammatical utterances; MATTR (including CS) = moving-average type-token ratio (code switching included); MATTR (excluding CS) = moving-average type-token ratio (code switching excluded).

questionnaire gathered information regarding relative language use (Spanish and English) at home with different speakers (mother, father, older siblings, peers), as well as the highest level of maternal education completed (adapted from Francis et al., 2005). As reported in the parent questionnaire (see Supplemental Material A.1-A.2), the 34 participants came from households with a mean maternal level of education that included vocational training, but no university attendance. The mean home language use indicated that the families, on average, reported using both languages relatively equally in their homes. Specifically, parents reported that interactions with their children were on average primarily in Spanish, while the interactions of the children with their older siblings and peers were in both languages relatively equally.

The children's narratives were produced primarily in Spanish, with more than 70% mean total words produced in Spanish over the 3-year duration of the study (see Table 1). Across the 6 waves of data collection, most of the participants (85%) produced three or more narrative samples. It is important to note that a comparison with the participants who produced three or more narrative language samples across the 6 waves yielded parallel findings (see Supplemental Material B). As specified by Luke (2004), maximum likelihood estimation was utilized to effectively handle missing data.

Descriptive statistics, including the number of narratives included at each wave of data collection are outlined in Table 1. Prior work with bilingual children has noted that participants are sometimes reluctant to supply language samples in their minority language (Mishina-Mori, 2011; Orellana, 1994). The relatively low number of samples produced by participants (aside from those not collected) in the first wave of data collection may have been related to the young mean age or early academic experiences of the participants during their first semester in preschool-3. However, the number of narrative retell samples produced by participants increased across subsequent waves of data collection. The distribution of the 34 participants who produced their *first* narrative language sample (as previously noted, the majority produced two or more samples over time) at each wave is as follows: Nine participants did so in wave 1, four in wave 2, 17 others in wave 3, no participants provided their first narrative sample in wave 4, two additional participants in wave 5, and two participants produced their first narrative language sample in wave 6. It is important to note that this distribution indicates that not every participant provided a language sample at each wave.

2.2. Longitudinal Language Sampling Protocol

Narrative retell language samples were elicited during the fall and spring semesters of each academic year from preschool through kindergarten, yielding a total of 6 waves of observation collected over 3 years. Each semester, the participants were asked to provide a narrative retell sample in Spanish. Sampling occurred no earlier than three weeks into the academic semester to allow students time to adjust to classroom procedures prior to being sampled. Four wordless picture storybooks, counterbalanced across participants and time, were used to elicit the narrative retell language samples: *Frog, Where Are You?,* (Mayer, 1969); *Frog Goes To Dinner,* (Mayer, 1974); *Frog On His Own,* (Mayer, 1975a); and *One Frog Too Many* (Mayer, 1975b). Spanish-English bilingual examiners first told the participants one of the stories using scripts in Spanish for each wordless picture storybook (Miller & Iglesias, 2018), thereby providing an example of the narrative task. Participants were then asked to retell the same story in Spanish using the pictures from the book as a prompt. The narratives were digitally recorded, and tracked through use of alphanumeric codes, to ensure participant anonymity.

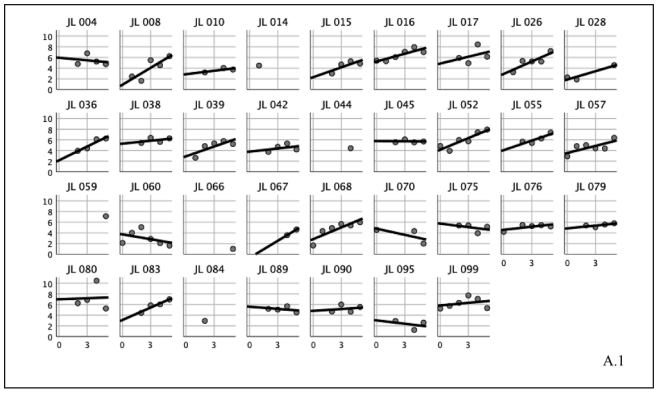
2.3. Coding of Outcome Measures

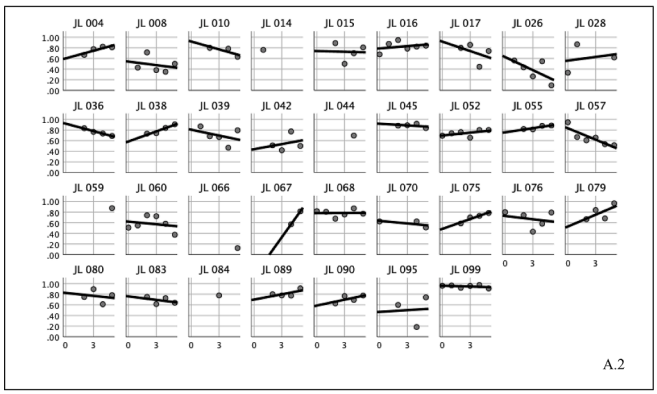
Narrative language sample analyses (LSA) was conducted using the Systematic Analysis of Language Transcripts (SALT) software (Miller & Iglesias, 2018) in order to transcribe and calculate mean length of utterances in words (MLUw), percent of grammatical utterances (PGU), and moving-average type-token ratio (MATTR). MLUw and MATTR (25 word, moving window) were automatically calculated by SALT. Each utterance was coded at the utterance level, as either grammatical or not. Utterances without grammatical errors were coded as grammatical. Utterances were segmented using modified communication units to account for the PRO-drop nature of Spanish as recommended for language sample analysis with Spanish-speaking children (Gutiérrez-Clellen et al., 2000).

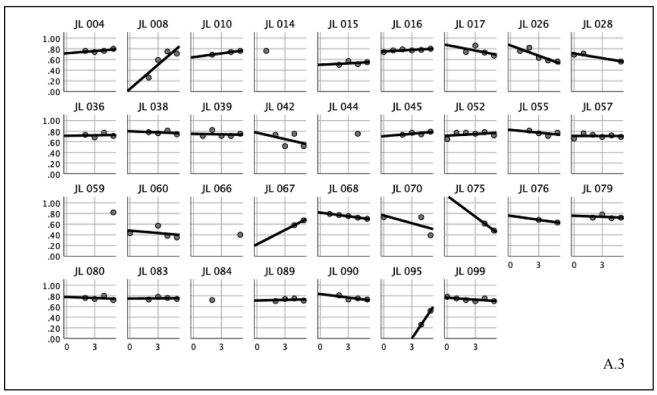
It should be noted that MLUw in the present study included all complete and intelligible words produced in Spanish and English, and excluded words omitted by the child or words included in mazes. Including all complete and intelligible words produced in Spanish and English in the calculation of MLUw allowed for the full complexity of every utterance to be represented.

PGU was calculated by dividing the number of utterances that did not include morphosyntactic errors or omissions (coded as grammatical at the utterance level) by the total number of complete and intelligible utterances in each narrative. Utterances that included one or more morphosyntactic errors were considered ungrammatical. Morphosyntactic errors included omissions of obligatory words or morphemes, errors of gender or number agreement, and errors of verb tense. Consistent with MLUw, PGU included all complete and intelligible words produced in Spanish and English, and excluded words included in mazes. Therefore, morphosyntactic errors in either language rendered an utterance as ungrammatical. Instances when the child code switched from Spanish to English, where English influenced either the word order or morpheme would not be counted as ungrammatical if it was considered grammatical in the other language. To illustrate, "*the rana's babies"* would be coded as a grammatical utterance, whereas "*los dog se duerme*" would be coded as an ungrammatical utterance.

The present study also considered the added complexity in the lexical measure (MATTR) that may occur due to code switching, which is the alternating shifting between two languages often used by bilingual speakers (Gutiérrez-Clellen, 2009). Code switching is considered to be a normal phenomenon among fluent bilingual speakers when used within the grammatical structures of the languages being used (see Paradis, 2012 for a systematic review), but as it has been noted as increasing over time in prior studies of language loss (Anderson, 1999a, 1999b, 2001; Guiberson et al., 2006; Kravin, 1992). Thus, the present study accounts for words that were code switched into the non-target language (English) in order to yield the most accurate LSA measure of lexical diversity (MATTR) at the







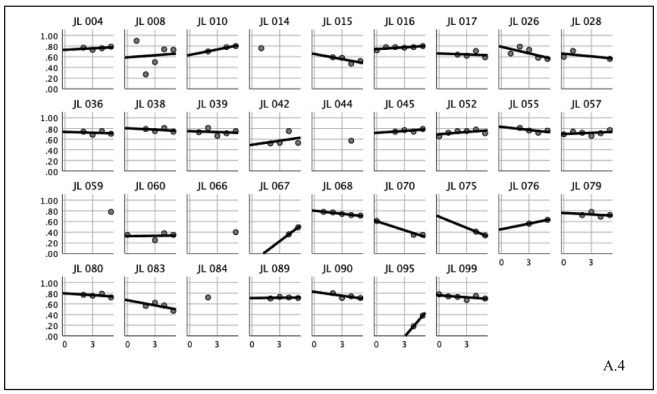


Figure A. Individual linear growth trajectories for mean length of utterance in words (*A.1*); proportion of grammatical utterances (*A.2*), and moving-average type-token ratio including *(A.3)* and excluding (*A.4*) code switching in Spanish.

word level.

Trained research assistants with native to near-native oral and literate proficiency in Spanish orthographically transcribed the digital audio recordings of narrative language samples using SALT software. Transcribers with at least one year of intensive transcription experience (10 hours or more per week) were trained to clean transcripts as expert research assistants. Cleaning consisted of inspecting completed transcripts for any word, utterance segmentation, or code errors and correcting those errors. Reliability was then conducted in a third round of editing by expert research assistants by randomly selecting 25% of the transcripts to determine inter-rater reliability of transcription accuracy, segmenting, and coding agreement. Word-for-word transcription accuracy ($M = 94\%$; $SD = 7\%$), utterance segmentation accuracy ($M = 97\%$; $SD = 6\%$), and coding agreement ($M = 91\%$; $SD = 10\%$) were indexed by dividing the total agreements by the product of total agreements plus disagreements. Coding agreement included English code switched words, coded at the word level, as well as grammatical utterances, coded at the utterance level.

2.4. Analytic Approach

The longitudinal design of the study modeled the change over time in MLUw, PGU, and MATTR in Spanish. It allowed for dynamic and simultaneous comparisons of change over time within and across participants (Singer & Willett, 2003). Time, measured in academic semesters (waves) of data collection, served as a time-varying covariate. Specifically, this study examined expressive Spanish language skills over 6 consecutive academic semesters: fall semester of preschool-3 = wave 1; spring semester of preschool-3 = wave 2; fall semester of preschool-4 = wave 3; spring semester of preschool-4 = wave 4; fall semester of kindergarten = wave 5; and spring semester of kindergarten = wave 6.

Growth curve modeling (GCM) was used to analyze the longitudinal data collected using IBM SPSS Statistics 26.0 software for Mac (IBM corp., 2018). GCM allows for simultaneous analyses of the same measure obtained at several data points over time for the same individual (Hox & Stoel, 2005). Thus, multiple narrative language samples can be used to observe changes over time in Spanish within and across the participants. GCM also had the added benefit of handling missing data at single or multiple time points throughout the data collection period, as it can change and infer patterns based on the available data at the intra- and inter-individual levels by using maximum likelihood estimation (Luke, 2004). Maximum likelihood helped to balance the data and allowed for maximization of the number of participants and language samples used in the analysis, without requiring exclusion of participants with missing data from one or more transcribed language samples.

A series of models were estimated for each of the outcome measures. Unconditional means (UM) models as well as linear and curvilinear GCMs (with academic semester serving as a time-varying covariate) were estimated for each of the expressive language measures (Singer & Willett, 2003). Visual inspection of empirical growth plots of the individual trajectories for each participant in the sample indicated a range of linear and curvilinear trajectories across the outcome measures (see Figure A).

The UM models were conducted first to measure the variance of the outcomes without including time as a predictor (Singer & Willett, 2003). The UM models also served as the baseline models for the proportional reduction of within-person residual variance (pseudo-R^2_ε) by subsequent comparison to GCMs that specified the effect of time on each outcome variable. Linear and quadratic unconditional growth curve models were then estimated with academic semester serving as a time-varying covariate to determine if the prototypical growth trajectories were linear or curvilinear. All possible model specifications of fixed and randomly varying intercepts and slopes were estimated for the unconditional GCMs to determine if the baseline (intercept) alone varied across participants, or the baseline and the rate of growth varied. The order in which the models were specified was from fixed to random and linear to quadratic resulting in a total of eight unconditional GCMs for each outcome measure.

The best fitting model for each measure was found by using the deviance statistic or -2 log-likelihood (-2LL) as the primary goodness-of-fit index, where lower -2LL values were better (Field, 2013). To determine whether models had a statistically significant better fit, the model fit was confirmed with a χ^2 distribution test of degrees of freedom between nested and adjacent models to assess -2LL differences (Field, 2013). The estimates from the best fitting models were used to generate the prototypical growth trajectories for each of the outcome measures (see Figures B-D).

Table 2
Growth Curve Models MLUw, PGU, MATTR (Including CS), and MATTR (Excluding CS) in Spanish

	Parameter	*MLUw*	*PGU*	*MATTR (Including CS)*	*MATTR (Excluding CS)*
Fixed effects: γ (SE)					
Intercept	γ00	3.27*** (0.37)	0.72*** (0.03)	0.71*** (0.02)	0.73*** (0.03)
Linear Slope	γ10	0.90*** (0.24)	-0.01 (0.01)	-0.01 (0.005)	-0.02** (0.007)
Quadratic Slope	γ20	-0.10* (0.04)			
Goodness-of-fit					
-2LL		433.1*	-118.1	-220.9	-151.1**

Note. MLUw = mean length of utterance in words; PGU = proportion of grammatical utterances, MATTR (including CS) = moving-average type-token ratio (code switching included); MATTR (excluding CS) = moving-average type-token ratio (code switching excluded); SE = standard error; -2LL = -2 log-likelihood deviance statistic.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

3. Results

The best fitting unconditional growth curve models (GCMs) for MLUw, PGU, and MATTR in Spanish had randomly varying intercepts and fixed slopes. Table 2 specifies the best fitting models for each of the expressive language measures.

3.1. Mean Length of Utterances in Words (MLUw)

The best-fitting model for MLUw (see Table 2) was the fixed quadratic model. This model had the lowest goodness-of-fit deviance statistic (-2LL = 433.1, $p < .05$, for a χ^2 distribution on 1 *df*). This final model also had the highest overall proportional reduction (25%) of within person residual variance (pseudo-R^2_ε) relative to the UM model by estimating the effect of time.

The fixed effects of the final model, estimated that the average initial status (onset of growth at the fall of preschool-3) of MLUw was $\gamma 00 = 3.27$, $p < .001$, with a positive and significant linear rate of change ($\gamma 10 = 0.90$, $p < .001$), and significant curvilinear deceleration ($\gamma 20 = -0.10$, $p < .05$) over time, specifically from semester to semester. The variance components estimated significant within person variance (σ^2_ε) over time and between person variance (σ^2_0) at initial status. Figure B illustrates the prototypical growth trajectory of MLUw in Spanish.

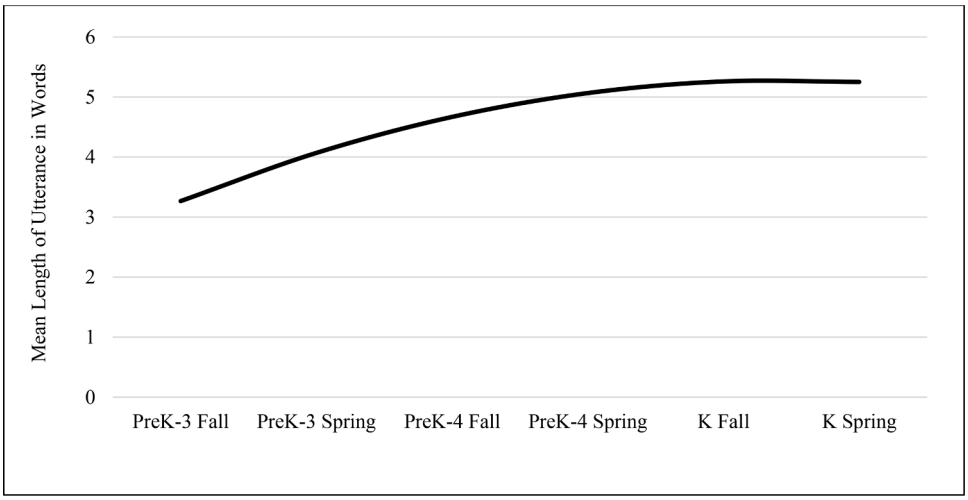


Figure B. Prototypical growth trajectories for grammatical measure of mean length of utterance in words in Spanish.

3.2. Percent of Grammatical Utterances (PGU)

The best-fitting model for PGU (see Table 2) was the fixed linear model. This model had a lower goodness-of-fit deviance statistic than the UM model (-2LL = -118.1, $p = .35$, for a χ^2 distribution on 1 *df*). Although this model did not have statistically significant better fit, it did have an overall higher proportional reduction of 1% in within person residual variance (relative to the UM model); it was selected as the final model as it converged, and demonstrated a lower -2LL deviance statistic with a 1 degree of freedom increase ($df = 4$) by including the effect of time. Technically, the best-fitting model for PGU indicated that the addition of time was non-informative, estimating negligible to no growth over time.

The fixed effects of the final model, estimated that the average initial status of PGU was $\gamma 00 = 0.72$, $p < .001$, with a negative and nonsignificant linear rate of change ($\gamma 10 = -0.01$, $p = .35$), indicating that there was no significant growth or decline of expressive grammaticality over time. The variance components estimated significant within person variance (σ^2_ε) and between person variance (σ^2_0) at initial status. Figure C illustrates the prototypical growth trajectory of PGU in Spanish.

3.3. Moving-Average Type-Token Ratio (MATTR)

The GCM testing process was estimated for MATTR in Spanish including *and* excluding code switching into English. The best-fitting models for both iterations of MATTR were fixed linear GCMs (see Table 2). The final model for MATTR including code switching had a lower goodness-of-fit deviance statistic than the UM model (-2LL = -220.9, $p = .08$ for a χ^2 distribution on 1 *df*). Although the fit of this model did not have statistically significant better fit, it did have an overall higher proportional reduction of 2% in within person residual variance (relative to the UM model); it was selected as the final model as it converged, and demonstrated a lower -2LL deviance statistic with a 1 degree of freedom increase ($df = 4$) by including the effect of time. Technically, the best-fitting model for MATTR including code switching indicated that the addition of time was non-informative, estimating negligible to no growth over time in MATTR including code switching.

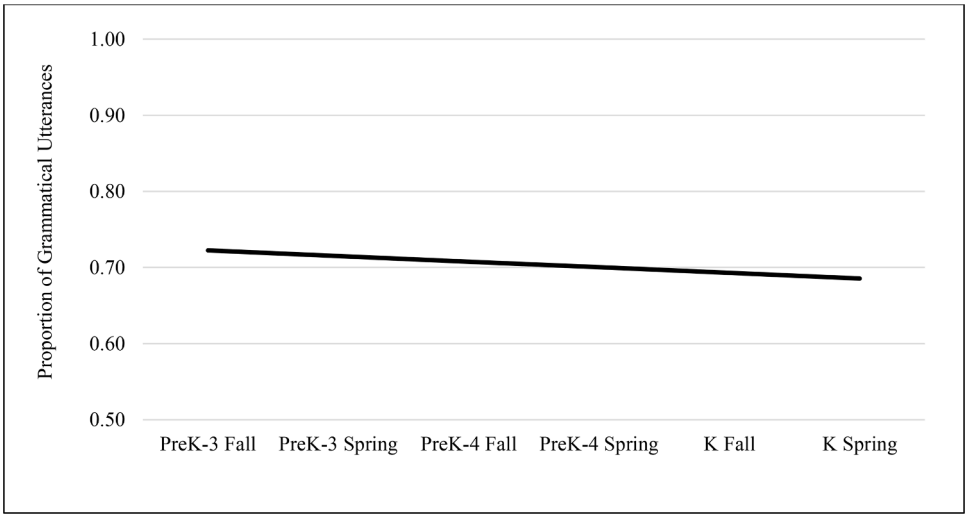


Figure C. Prototypical growth trajectories for grammatical measure of proportion of grammatical utterances in Spanish.

The fixed effects of the final model for MATTR including code switching estimated an average initial status of $\gamma 00 = 0.71, p < .001$, with a negative and nonsignificant linear rate of change ($\gamma 10 = -0.01, p = .08$). The variance components estimated significant within person variance (σ^2_ε), and between person variance (σ^2_0) at initial status. Figure D illustrates the prototypical growth trajectory for MATTR including code switching in Spanish.

The final model for MATTR excluding code switching had the lowest goodness-of-fit deviance statistic (-2LL = -151.1, $p < .01$, for a χ^2 distribution on 1 *df*). This final model had the highest overall proportional reduction (6%) of within person residual variance (pseudo-R^2_ε) relative to the UM model.

The fixed-effects of the final model for MATTR excluding code switching estimated an average initial status of $\gamma 00 = 0.73, p < .001$, with a significant and negative linear rate of change ($\gamma 10 = -0.02, p < .01$). The variance components for MATTR excluding code switching estimated significant within person variance (σ^2_ε) over time, and between person variance (σ^2_0) at initial status. Figure D illustrates the prototypical growth trajectory for MATTR excluding code switching in Spanish.

Differences were found between MATTR including code switching and MATTR excluding code switching. The mean initial status (onset of growth) of MATTR excluding code switching was estimated to be 2% higher relative to MATTR including code switching. However, the mean rate of negative deceleration of MATTR excluding code switching was 1% faster per semester relative to MATTR including code switching.

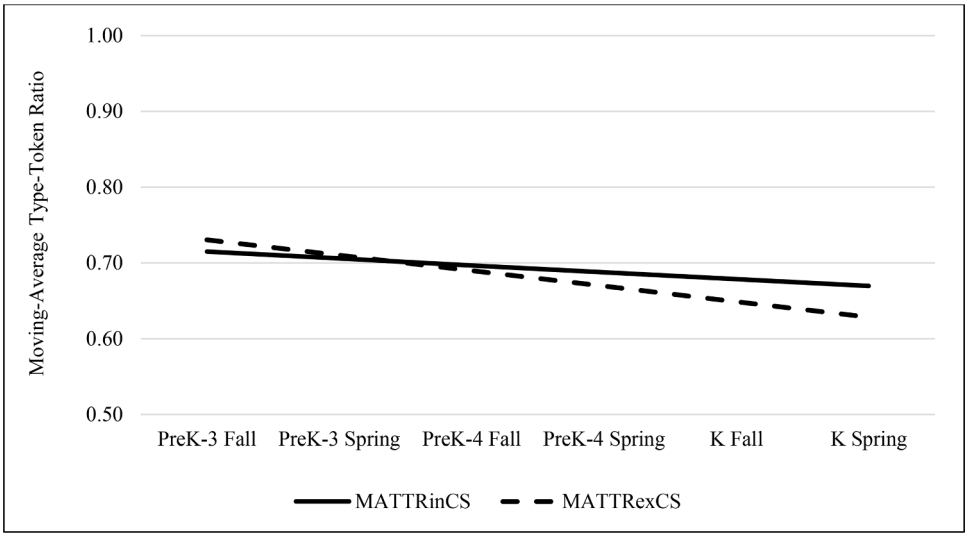


Figure D. Prototypical growth trajectories for the lexical measure of moving-average type-token ratio in Spanish including (MATTRinCS) and excluding code switching (MATTRexCS).

4. Discussion

This 3-year longitudinal study used growth curve modeling (GCM) to map individual and group changes in Spanish expressive language measures of grammaticality (MLUw and PGU), and lexical diversity (MATTR including and excluding code switching) over 6 consecutive academic semesters in typically developing Spanish-English bilingual children who attended an English immersion school. The study's research questions examined whether bilingual children demonstrated Spanish growth or loss across grammatical and lexical skills, and at what rates.

Ultimately the best fitting models for the outcome measures were fixed GCMs, for which the intercepts varied randomly but the linear and curvilinear rates of growth were fixed. A common finding across the outcome measures, with the partial exception of the initial significant linear growth rate for MLUw (followed by significant deceleration), was that participants showed loss of Spanish expressive language skills over time. Although the rate of negative growth was not significant for all of the outcome measures, language loss as defined in this study included maintenance or non-significant growth over time.

4.1. Spanish Growth or Loss of Grammatical Skills: Mean Length of Utterance in Words (MLUw)

The findings mostly supported the hypothesis that Spanish-English bilingual children academically instructed in English would demonstrate loss of grammatical skills, indexed by MLUw, from preschool through kindergarten. The participants showed significant linear acceleration along with significant deceleration of MLUw in Spanish (see Table 1; Figure B).

Prior smaller scale longitudinal studies that measured MLU (Anderson, 1999a, 1999b, 2001) demonstrated loss in this measure of gross morphosyntax. In this study, MLUw demonstrated growth for 2.5 years, albeit followed by a decline toward the end of kindergarten. The findings provided minimal support for the hypothesis that code switching to English may have driven the period of overall growth for MLUw from the fall of preschool-3 to the fall of kindergarten. Table 1 shows a relative increase in the mean proportion of total words produced in English (code switched words) in waves 3, 4, and 5. Although there was some increased use in code switching, it may not have been significant, and was not maintained across all waves of the study. The measure of MLUw with code switching excluded may have yielded different results, however calculating MLUw excluding code switching at the word or utterance level would arguably yield a metric of MLUw not representative of children's gross morphosyntactic skill. Given that this study focused on the growth and loss of Spanish language skills over time, it was important to capture the most representative measure of gross morphosyntactic skill, which was MLUw including all grammatical components produced in each utterance irrespective of target language. Given these results, MLUw may not be among the best measures to represent growth trajectories of Spanish language development.

4.2. Spanish Growth or Loss of Grammatical Skills: Percent of Grammatical Utterances (PGU)

The findings relevant to the grammatical measures supported the hypothesis that Spanish-English bilingual children academically instructed in English would demonstrate loss of grammatical skills as indexed by PGU, from preschool through kindergarten. The participants showed loss as defined in this study (maintenance; no significant growth or decline) of PGU in Spanish (see Table 1; Figure C).

In addition to MLUw, this study examined another measure of grammatical skill in Spanish, PGU, which measured grammaticality at the utterance level. PGU provides a distinct dimension of grammatical skill, as a child who produces an elevated MLUw will not necessarily also demonstrate a high degree of grammaticality. The findings for PGU were indicative of precisely this distinction, as PGU demonstrated essentially no change over time, while MLUw demonstrated overall growth during the same period of time. The children demonstrated an average initial status of 71% grammaticality in Spanish during the fall of preschool-3, which approximated, but did not reach, mastery level, when defined as $\geq$ 80%. Yet over the next 2.5 years they did not demonstrate any measurable growth in their grammaticality although the length of their utterances on average increased when asked to speak in Spanish. This is important as the children in this study never reached mastery level in their Spanish grammaticality, indicating that PGU may be a more sensitive measure of Spanish loss than MLUw during the preschool and kindergarten years. The lack of growth of Spanish PGU may be associated with the children in this study being academically instructed in English from an early age, and therefore were required to focus on English grammar. As a consequence of the lack of support in Spanish, some children may gradually lose Spanish grammar skills, particularly unshared aspects with English grammar, such as gender agreement (Anderson & Márquez, 2009; Shin, 2018). It is important to note that a recent study by Castilla-Earls et al. (2019) tracked longitudinal development of PGU across different programs of bilingual education from kindergarten through Grade 2, which also found a decline in Spanish PGU irrespective of program type. Therefore, the findings from this study serve as a downward extension to Castilla-Earls et al. (2019), demonstrating that loss of Spanish grammaticality can be detected as early as during the preschool years.

4.3. Spanish Growth or Loss of Lexical Diversity: Moving-Average Type-Token Ratio (MATTR)

The findings supported the hypothesis that English instructed bilingual children would demonstrate loss of lexical diversity in Spanish, as specifically indexed by MATTR with code switching excluded, from preschool through kindergarten. This hypothesis also considered the likelihood that code switching to English would increase over time, with the expectation that MATTR including code switching could demonstrate positive growth or a slower decline. The participants showed negative and significant loss of MATTR excluding code switching and non-significant decline of MATTR including code switching, indicating that most participants

demonstrated Spanish loss of verbal productivity and lexical diversity. Although participants started with a slightly higher level of Spanish lexical diversity in preschool when English vocabulary was not accounted for, this lexical diversity in Spanish suffered from a significant decline over time (see Table 1; Figure D). Put another way, the participants were using more Spanish at the beginning (fall semester) of preschool-3 than in later academic semesters. This was not surprising, since previous findings on Spanish loss of lexical diversity reported an increase in the use of general terms, which would likely drive down the ratio of diverse word use (e.g., Anderson, 1999a; Guiberson et al., 2006). Given the results of this study, and the findings of increased use of code switching and general terms from previous studies of Spanish loss, MATTR would be expected to continue decelerating at a faster rate when calculated using types and tokens in the target language only.

The MATTR findings indicated that on average as the narratives increased in length over time (in Table 1, see number of total words; NTW), the ratio of types to tokens decreased for most children, meaning that the children were not using as many different words as their number of total words increased. The finding of Spanish MATTR loss is consistent with previous studies of language loss, which have shown that lexical diversity also declines over time (see Anderson, 2012 for a review). Fergadiotis et al. (2015) and Kapantzoglou et al. (2019) demonstrated that MATTR is the most accurate measure of lexical diversity when accounting for narrative length, and therefore the loss of MATTR in Spanish may be more indicative of the dynamic changes in Spanish lexical diversity over time for bilingual Spanish-English children, and perhaps particularly for those who are English instructed.

4.4. Qualitative Differences of Individual Participants Showing Overall Growth or Loss

Individual participants demonstrated a wide range of trajectories of linear and curvilinear growth and loss of Spanish grammatical and lexical skills, as illustrated on Figure A. The majority of participants varied in their rates of growth or loss across measures. There were, however, three total participants that experienced loss across all Spanish outcome measures in this study. Participants JL 060, JL 070, and JL 076 all had negative trajectories of MLU, PGU, and MATTR across six academic semesters (see Figure A). Upon closer inspection of their demographic information and individual narrative language samples, these three participants shared a number of similarities. All three were female, as was the case for most participants in prior studies of Spanish loss (Anderson, 1999a, 1999b, 2001; Guiberson et al., 2006). However, there clearly aren't enough participants who have been followed longitudinally to suggest that females are more at risk for experiencing language loss. All three participants had mothers with education at or above the mean level of maternal education for the sample (vocational training, or some community college). Although on average, the study participants were reported to live in relatively balanced bilingual homes, the three participants who demonstrated loss across all outcome measures predominantly used English when communicating specifically with peers and siblings.

A distinguishing feature of the narrative language samples from these three participants was they demonstrated a considerably higher overall proportion of code switching (50.7%) than the sample mean (23.1%), when averaged across all six waves. Prior studies have reported an increase in use of code switching for children experiencing loss (Anderson, 1999a, 1999b, 2001; Guiberson et al., 2006; Kravin, 1992), which seems to hold true for these participants who showed negative growth trajectories across all outcome measures. Inspection of their narrative retells revealed utterances consisting of mostly labeling characters in the story, with little expansion on ideas or use of Spanish verbs. Finally, these children produced frequent morphologic Spanish errors in their narratives, with the most common being gender disagreement. These participants produced numerous gender-disagreement errors when marking storybook characters, mostly using the feminine, singular definite article '*la*' to substitute for masculine articles. This specific error pattern was also documented in Anderson and Márquez (2009), which investigated Spanish loss in grammar skills of typical and language impaired bilingual children.

In contrast to the three participants who demonstrated loss across all outcome measures there were two participants who demonstrated overall positive growth trajectories (JL 016 and JL 052; see Figure A). Both of these participants were male, and had mothers with education levels that corresponded to the group mean. These children were reported to use more Spanish than English in the home. Their narrative language samples included higher levels of subordination, and they produced fewer morphologic errors, relative to the children who demonstrated overall Spanish language loss. The narratives in fact increased in complexity over time, indicating they were continuing to learn more Spanish over the 3-year duration of this study. Another distinction of these participants was that they essentially did not code switch into English (0.01%) across all waves of observation.

The qualitative features of the language samples and demographic information from these individual participants, who appeared to experience overall Spanish loss or growth across multiple domains of language, do not represent a cause of language loss or growth, but rather provide a deeper look at their expressive language skills. No statistical analyses were conducted to connect code switching, specific morphologic errors, or subordination index as predictors of growth or loss of the measures MLUw, PGU, or MATTR over time. Additional research would be required to make connections between individual differences and patterns of growth trajectories. Although prior studies have made such connections, for example, to birth order (Wong Fillmore, 1991), or socioemotional skills (Simon-Cereijido et al., 2013), there is still relatively little known about what drives patterns of language growth in bilingual children. Furthermore, although the individual characteristics of the children who were described qualitatively (e.g., gender, home language, maternal level of education) were similar, this was not supported by the larger scale statistical analyses that used these covariates as predictors of growth patterns. In other words, the individual differences described did not hold true for the larger group of participants in this study.

4.5. Conclusion and Clinical Implications

Measures that track proportion of change over time (PGU; MATTR), may be better suited to investigate Spanish loss than measures

that track raw change (MLUw) over time. MLUw may also be less sensitive to detecting language loss due to the potential increase of code switching over time, which in turn may inflate gross complexity of utterances as bilingual children become more proficient in English. A different measure such as subordination index, which was considered in reviewing the qualitative differences among participants, may be more sensitive for tracking changes in grammatical complexity in older children. Finally, this study along with others found minority language loss in bilingual children who were academically instructed in English (Anderson, 1999a, 1999b, 2001; Anderson & Márquez, 2009; Kravin 1992). Therefore, language of academic instruction is another important factor to consider, which can range from English only to dual language instruction.

The results from this study also highlight the importance of considering code switching in lexical measures, and the need for transparency in future studies that address lexical diversity in bilingual children. The analyses of MATTR either including or excluding code switching demonstrated the differences in both the intercepts and growth rates that could be found within the same measure and same group of participants. Because code switching is a common occurrence in bilingual language production (see Paradis, 2012 for a systematic review), as well as an indicator of language loss (Anderson, 1999a; Guiberson et al., 2006; Kravin, 1992), future studies need to indicate whether their analytic approach includes code switched words and utterances. This becomes particularly relevant in clinical diagnostics, when studies of typical bilingual language development such as this one should provide an accurate comparison of language production. Studies that exclude any utterances that contain any code switched words provide an incomplete depiction of typical bilingual language development. The decision to include code switching would be beneficial when researchers or clinicians want a holistic view of the language use in bilingual children, while excluding code switched words is appropriate when a single language is being targeted for a specified purpose.

The findings from this study may have potential clinical implications, particularly within the scope of Spanish-English bilingual assessment. Spanish-English bilingual children instructed in academic environments that offer negligible to no academic support in Spanish may have an increased risk of experiencing Spanish loss. It is possible that Spanish loss may contribute to the misdiagnosis of language impairment in the Spanish-English bilingual preschool and school-age population (Bedore & Peña, 2008). Language loss is particularly important to consider in clinical assessment, because it has been shown to be more evident and occur more rapidly in children who present with language impairments (Anderson & Márquez, 2009; Castilla-Earls et al., 2016; Ebert, Pham, & Kohnert, 2014; Restrepo & Kruth, 2000). Additionally, it may be advantageous to conduct assessment at more than one point in time in a clinical setting to determine the growth trajectory in the minority language in order to determine if language loss should be considered in making an informed clinical diagnosis. Bilingual children who experience language loss may indeed be typically developing like those in the present study, however, those who experience loss at a faster rate may be at risk.

The longitudinal findings from this study can be cautionary for speech-language pathologists working with Spanish-English bilingual children in schools with English immersion academic programs. This study provides empirical, longitudinal evidence that Spanish-English bilingual children who are typically developing can simultaneously demonstrate Spanish growth *and* loss over time across a range of grammatical and lexical expressive language skills (see Figure A). The results also offer foundational knowledge of typically developing Spanish language skills within a particular context of academic instruction, by more precisely outlining dynamic patterns of Spanish growth and loss in bilingual children (see Figures B-D). This study can provide increased awareness as well as an initial basis of knowledge for clinicians working with young Spanish-English bilingual children in the schools, and can potentially inform an approach to Spanish-English bilingual clinical assessment. Future work should explore possible diagnostic markers that can differentiate loss from impairment.

4.6. Future Directions

Although the results of this longitudinal study may have a range of clinical implications for Spanish-English bilingual children and add to the limited knowledgebase on Spanish loss, the findings should be interpreted with caution. The sample size of 34 participants, although larger than many previous studies of language loss, was still relatively small. The participants of this study represented Hispanic, Spanish-English bilingual children from homes experiencing low socioeconomic status, with primarily non-college educated mothers, living in the Dallas, Texas, metropolitan area. It is possible that students in other areas of the U.S. coming from diverse backgrounds will not show the same degree or patterns of Spanish growth and loss. The participants from this study, however, were more representative of Hispanic children in the U.S. than those from the majority of prior work on Spanish loss (Anderson, 1999a, 1999b, 2001). The present study included Spanish-English bilingual preschool children in English immersion classrooms, and therefore the results cannot be generalized to bilingual children who speak other languages or attend other types of academic programs (e.g. dual language or transitional bilingual). However, this study represents an important first step for research on Spanish loss with the benefit of having studied this phenomenon within the context of one particular academic setting. Strategically comparing and contrasting the individual errors from children with typical development who are undergoing Spanish loss relative to those of children with language impairment may help to differentiate a true disorder from language loss.

Although this study examined growth and loss in Spanish expressive language measures, examining receptive language skills longitudinally would provide an important contribution to the limited body of work on the process of minority language growth and loss. In addition, future work on Spanish loss should also consider the dynamic co-development of English using both expressive and receptive language measures. More in-depth longitudinal analysis of code switching may also be indicated, as it has been reported to increase in prior studies of language loss as well as in this study. Growth curve modeling of code switching may be a useful measure of potential Spanish growth and loss in future work. However, such work would need to longitudinally profile trajectories of code switching in bilingual children when speaking in the minority language (e.g., Spanish), as well as when speaking in English. Finally, a more in-depth qualitative look into participants who demonstrate loss may provide a better understanding of specific error patterns or

indicators of language loss in typically developing participants.

This study is the first in a series of studies that will address the dynamic co-development and impact of Spanish loss on the English oral language and reading skills of Spanish-English bilingual children. Future planned studies will extend the waves of data collection to determine how the patterns of Spanish growth and loss evidenced in this study may persist or change after kindergarten, and how these patterns may impact measures of academic achievement. In addition to the extended waves, future planned work will include more participants, some from other contexts of academic instruction such as dual language, in order to compare mainstream academic settings and their relative impact on the dynamic growth and loss of Spanish and English language skills over time.

Acknowledgements

This research has been supported in part by a UT Dallas Dissertation Research Award awarded to Lindsey Hiebert; by Grant GA-2013-016–Jerry M. Lewis, M.D. Mental Health Research Foundation, and Grant 13180 – Anonymous Donor awarded to Raúl Rojas. The authors wish to thank the dedication from all present and prior members of the Bilingual Language Laboratory as well as the participants and their families.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jcomdis.2021.106110.

References

Anderson, R. T. (1999a). Impact of language loss on grammar in a bilingual child. *Communication Disorders Quarterly, 21*, 4–16.
Anderson, R. T. (1999b). Loss of gender agreement in first language attrition: Preliminary results. *Bilingual Research Journal, 23*, 319–338.
Anderson, R. T. (2001). Lexical morphology and verb use in child first language loss: A preliminary case study investigation. *International Journal of Bilingualism, 5*(4), 377–401.
Anderson, R. (2012). First language loss in Spanish-speaking children. In B. A. Goldstein (Ed.), *Bilingual language development & disorders in Spanish-English speakers* (pp. 193–212). Baltimore, MD: Paul H. Brookes Publishing Co.
Anderson, R., & Márquez, A. (2009). The article paradigm in Spanish-speaking children with SLI in language contact situations. In J. Grinstead (Ed.), *Hispanic child languages: Typical and impaired development* (pp. 29–55). Philadelphia: John Benjamins.
August, D., & Shanahan, T. (2006). Developing literacy in second-language learners. *Report of the national literacy panel on language minority children and youth*. Mahwah, NJ: Lawrence Erlbaum Associates.
Bedore, L. M. (2001). Assessing morphosyntax in Spanish-speaking children. *Seminars in Speech and Language, 2*(1), 65–77.
Bedore, L. M., & Peña, E. D. (2008). Assessment of bilingual children for identification of language impairment: Current findings and implications for practice. *International Journal of Bilingual Education and Bilingualism, 11*(1), 1–29.
Bedore, L. M., Pena, E. D., Gillam, R. B., & Ho, T. H. (2010). Language sample measures and language ability in Spanish-English bilingual kindergarteners. *Journal of Communication Disorders, 43*(6), 498–510.
Cárdenas-Hagan, E., Carlson, C. D., & Pollard-Durodola, S. D. (2007). The cross-linguistic transfer of early literacy skills: The role of initial L1 and L2 skills and language of instruction. *Language, Speech, and Hearing Services in Schools, 38*(3), 249–259.
Castilla-Earles, A., Bedore, L., Rojas, R., Fabiano-Smith, L., Pruitt-Lord, S., Restrepo, M. A., & Peña, L. (2020). Beyond scores: Using converging evidence to determine speech and language services eligibility for dual language learners. *American Journal of Speech-Language Pathology, 29*, 1116–1132.
Castilla-Earls, A., Francis, D., Iglesias, A., & Davidson, K. (2019). The Impact of the Spanish-to-English Proficiency Shift on the Grammaticality of English Learners. *Journal of Speech, Language, and Hearing Research, 62*, 1739–1754.
Castilla-Earls, A. P., Restrepo, M. A., Perez-Leroux, A. T., Gray, S., Holmes, P., Gail, D., & Chen, Z. (2016). Interactions between bilingual effects and language impairment: Exploring grammatical markers in Spanish-speaking bilingual children. *Applied Psycholinguistics, 37*, 1147–1173.
Dunn-Davidson, M., Hammer, C., & Lawrence, F. R. (2011). Associations between preschool language and first grade reading outcomes in bilingual children. *Journal of Communication Disorders, 44*(4), 444–458.
Ebert, K. D., Pham, G., & Kohnert, K. (2014). Lexical profiles of bilingual children with primary language impairment. *Bilingualism: Language and Cognition, 17*, 766–783.
Ennis, S. R., Ríos-Vargas, M., & Albert, N. G. (2011). *The Hispanic population: 2010. 2010 Census Briefs* (pp. 2–16). U.S: Census Bureau.
Federal Interagency Forum on Child and Family Statistics. (2015). Retrieved from http://www.childstats.gov.
Fergadiotis, G., Wright, H. H., & Green, S. B. (2015). Psychometric evaluation of lexical diversity indices: Assessing length effects. *Journal of Speech, Language, and Hearing Research, 58*, 840–852.
Ferron, J. M., Bell, B. A., Hess, M. R., Rendina-Gobioff, G., & Hibbard, S. T. (2009). Making treatment effect inferences from multiple-baseline data: The utility of multilevel modeling approaches. *Behavior Research Methods, 41*, 372–384.
Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th edition). Thousand Oaks, CA: SAGE Publications Inc.
Francis, D., Carlson, C., Fletcher, J., Foorman, B., Goldenberg, C., Vaughn, S., … Papanicolaou, A. (2005). Oracy/literacy development of Spanish-speaking children: A multi-level program of research on language minority children and the instruction, school and community contexts, and interventions that influence their academic outcomes. *Perspectives (The International Dyslexia Association, 31*, 8–12.
Gámez, P. B., & González, D. (2017). A comparison of narrative skill in Spanish-English bilinguals and their functionally monolingual Spanish-speaking and English-only peers. *International Journal of Bilingualism, 56*, 1–20.
Guiberson, M., Barrett, K. C., Jancosek, E. G., & Yoshinaga-Itano, C. (2006). Language Maintenance and Loss in Preschool-Age Children of Mexican Immigrants Longitudinal Study. *Communication Disorders Quarterly, 28*(1), 4–17.
Guiberson, M. (2013). Bilingual Myth-Busters Series: Language Confusion in Bilingual Children. *Perspectives on Communication Disorders and Sciences in Culturally and Linguistically Diverse Populations, 20*(1), 5–14.
Gusewski, S., & Rojas, R. (2017). Tense marking in the English narrative retells of dual language preschoolers. *Language, Speech, and Hearing Services in Schools, 48*, 183–196.
Gutiérrez-Clellen, V. F., Restrepo, M. A., Bedore, L., Peña, E., & Anderson, R. (2000). Language sample analysis in Spanish-speaking children. *Language, Speech, and Hearing Services in Schools, 31*(1), 88–98.
Gutiérrez-Clellen, V. F., Simon-Cereijido, G., & Erickson Leone, A. (2009). Code-switching in bilingual children with specific language impairment. *International Journal of Bilingualism, 13*(1), 91–109.

Hammer, C. S., Hoff, E., Uchikoshi, Y., Gillanders, C., Castro, D. C., & Sandilos, L. E. (2014). The language and literacy development of young dual language learners: A critical review. *Early Childhood Research Quarterly, 29*(4), 715–733.
Han, W. J. (2012). Bilingualism and academic achievement. *Child Development, 83*(1), 300–321.
Hansen, A. L. (2005). Research brief: Hispanic student achievement. *Education Partners, Inc.*Retrieved from http://www.educationpartnerships.org/.
Heilmann, J., Miller, J. F., Iglesias, A., Fabiano-Smith, L., Nockerts, A., & Andriacchi, K. D. (April-June 2008). Narrative transcription accuracy and reliability in two languages. *Topics in Language Disorders, 28*(2), 178–188.
Heilmann, J. J., Miller, J. F., & Nockerts, A. (2010). Using language sample databases. *Language, Speech, and Hearing Services in Schools, 41*, 84–95.
Hoff, E. (2013). Interpreting early language trajectories of children from low-SES and language minority homes: Implications for closing the achievement gap. *Developmental Psychology, 49*(1), 4–14.
Hoff, E., & Ribot, K. M. (2017). Language growth in English monolingual and Spanish-English bilingual children from 2.5 to 5 years. *The Journal of Pediatrics, 190*, 241–245.
Hox, J., & Stoel, R. D. (2005). Multilevel and SEM approaches to growth curve modeling. *Encyclopedia of Statistics in Behavioral Sciences, 3*, 1296–1305.
Iglesias, A., & Rojas, R. (2012). In B. A. Goldstein (Ed.), *Bilingual language development & disorders in Spanish-English speakers* (pp. 193-212). Baltimore, MD: Paul H. Brookes Publishing Co.
Kapantzoglou, M., Fergadiotis, G, & Auza Buenavides, A. (2019). Psychometric evaluation of lexical diversity indices in Spanish narrative samples from children with and without developmental language disorder. *Journal of Speech, Language, and Hearing Research, 62*, 70–83.
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*, 2852–2864.
Kravin, H. (1992). Erosion of a language in bilingual development. *Journal of Multilingual and Multicultural Development, 13*(4), 307–325.
Luke, D. A. (2004). *Multilevel modeling*. New Delhi, India: Sage Publications, Inc.
Macartney, S., Bishaw, A., & Fontenot, K. (2013). Poverty rates for selected detailed race and Hispanic groups by state and place: 2007-2011. *American Community Survey Briefs* (pp. 11–17). U.S: Census Bureau.
Malvern, D. D., & Richards, B. J. (1997). A new measure of lexical diversity. In A. Ryan, & A. Wray (Eds.), *Evolving models of language* (pp. 58–71). Clevedon, United Kingdom: Multilingual Matters.
Mayer, M. (1969). *Frog, where are you?* New York, NY: Penguin Books.
Mayer, M. (1974). *Frog goes to dinner*. New York, NY: Penguin books.
Mayer, M. (1975a). *Frog on his own*. New York, NY: Dial.
Mayer, M. (1975b). *One frog too many*. New York, NY: Dial.
McKee, G., Malvern, D., & Richards, B. (2000). Measuring Vocabulary Diversity Using Dedicated Software. *Literary and Linguistic Computing, 15*(3), 323–337.
McNeish, D., & Matta, T. (2018). Differentiating between mixed-effects and latent-curve approaches to growth modeling. *Behavior Research Methods, 50*, 1398–1414.
Menken, K., & Kleyn, T. (2010). The long-term impact of subtractive schooling in the educational experiences of secondary English language learners. *International Journal of Bilingual Education and Bilingualism, 13*(4), 399–417.
Merino, B. J. (1983). Language Loss in Bilingual Chicano Children. *Journal of Applied Developmental Psychology, 4*, 277–294.
Miller, J., & Iglesias, A. (2018). *Systematic Analysis of Language Transcripts (SALT). Research Version 16 [Computer Software]. Middleton*. WI: SALT Software, LLC.
Miller, J. F., Heilmann, J., Nockerts, A., Iglesias, A., Fabiano, L., & Francis, D. J. (2006). Oral language and reading in bilingual children. *Learning Disabilities Research and Practice, 21*, 30–43.
Mishina-Mori, S. (2011). A longitudinal analysis of language choice in bilingual children: The role of parental input and interaction. *Journal of Pragmatics, 43*(13), 3122–3138.
Orellana, M. F. (1994). *Superhuman forces: Young children's English language acquisition and Spanish language loss. (Doctoral dissertation)*. Retrieved from Educational Recourses Information Center (ERIC).
Paradis, J. (2012). Cross-linguistic influence and code-switching. In B. A. Goldstein (Ed.), *Bilingual language development & disorders in Spanish-English speakers* (pp. 73–91). Baltimore, MD: Paul H. Brookes Publishing Co.
Pearson, B. Z., Fernández, S. C., Lewedeg, V., & Oller, K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics, 18*(1), 41–58.
Restrepo, M. A., & Kruth, K. (2000). Grammatical characteristics of a Spanish-English bilingual child with specific language impairment. *Communication Disorders Quarterly, 21*, 66–76.
Rojas, R., Hiebert, L., Gusewski, S., & Francis, D. J (2019). Moving forward by looking back: Understanding why some Spanish-speaking English learners fall behindD. J. Francis (Ed.), Moving forward by looking back: Understanding why some Spanish-speaking English learners fall behind. *Identification, Classification, and Treatment of Reading and Language Disabilities in Spanish-speaking EL students. New Directions for Child and Adolescent Development, 166*, 1–35.
Rojas, R., & Iglesias, A. (2013). The language growth of Spanish-speaking English language learners. *Child Development, 84*(2), 630–646.
Ryan, C. (2013). Language use in the United States: 2011. *American Community Survey Reports*.
Schiff-Myers, N. B. (1992). Considering arrested language development and language loss in the assessment of second language learners. *Language, Speech, and Hearing Services in Schools, 23*, 28–33.
Shin, N. (2018). Child heritage speakers' morphosyntax. In Kim Potowski (Ed.), *Handbook of Spanish as a Heritage Language* (pp. 235–253). New York, NY: Routledge.
Simon-Cereijido, G., Gutiérrez-Clellen, V. F., & Sweet, M. (2013). Predictors of growth or attrition of the first language in Latino children with specific language impairment. *Applied Psycholinguistics, 34*, 1219–1243.
Singer, J. D., & Willet, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press, Inc.
Slavin, R. E., & Cheung, A. (2005). A synthesis of research on language of reading instruction for English language learners. *Review of Educational Research, 75*(2), 247–284.
U.S. Census Bureau (2010). *State and County QuickFacts.*Retrieved October 5, 2015 from http://www.quickfacts.census.gov.
Winsler, A., Burchinal, M. R., Hsiao-Chuan, T., Peisner-Feinberg, E., Espinosa, L., Castro, D. C., LaForett, D. R., Kim, Y. K., & De Feyter, J. (2014). Early development among dual language learners: The roles of language use at home, maternal immigration, country of origin, and socio-demographic variables. *Early Childhood Research Quarterly*, 1–15.
Wong Fillmore, L. (1991). When learning a second language means losing the first. *Early Childhood Research Quarterly, 6*, 323–346.