CITATION DETAILS:

Montanari, S., **Mayr, R**. & Subrahmanyam, K. (2018). Bilingual speech sound development during the preschool years: The role of language proficiency and cross-linguistic relatedness. *Journal of Speech, Language and Hearing Research,* doi:10.1044/2018_JSLHR-S-17-0393, published online 18 September 2018.

See: https://jslhr.pubs.asha.org/

TITLE: Bilingual speech sound development during the preschool years: The role of language proficiency and cross-linguistic relatedness.

AUTHORS AND AFFILIATIONS:

Simona Montanari (California State University, Los Angeles, USA)

Robert Mayr (Cardiff Metropolitan University, United Kingdom)

Kaveri Subrahmanyam (California State University, Los Angeles, USA)

We acknowledge the support of grant #5SC3GM847583-3 to Kaveri Subrahmanyam, Marlene Zepeda, and Simona Montanari from NIGMS

KEY WORDS: bilingual preschoolers, speech sound development, Spanish, phonological assessment, error analysis, longitudinal study

ABSTRACT

Purpose: The purpose of this study was to investigate typical Spanish-English speech sound development longitudinally in a group of bilingual preschoolers enrolled in a Head Start Program and to examine the extent to which such development is linked to language proficiency. The study also aimed to identify whether speech development is related cross-linguistically, and to improve our understanding of error patterns in this population.

Method: Thirty-five bilingual preschool children produced single-word speech samples in Spanish and English both at the beginning of their first and their second year in a Head Start program. Conversational samples in both languages were also collected at these data points to calculate Mean Length of Utterance in words (MLUw) and thus assess the children's linguistic proficiency. The phonetically transcribed speech samples were compared over time in terms of segmental accuracy measures and error pattern frequencies. Correlation analyses were run to examine the relation between segmental accuracy measures across languages, and between speech sound production and MLUw.

Results: One-way within-subjects ANOVAs revealed significant improvements in accuracy over time in both languages, but not always for cross-linguistically unshared segments, nor for all consonant manner classes. Overall error rates decreased over time in both languages, although certain error types showed no change. Cross-linguistic interactions were low in both languages. The results also revealed significant cross-linguistic correlations in segmental accuracy between Spanish and English, as well as between MLUw and speech sound production in both languages on a range of measures, with language-specific differences in Year 2 of the Head Start Program, but not in Year 1.

Conclusions: This study is the first to document developmental changes in the speech patterns of Spanish-English bilingual preschool children over one year. Accuracy rates improved significantly in both languages, suggesting that enhanced exposure to the majority language at school may not impede phonological development in the home language. Bootstrapping effects were particularly pronounced on cross-linguistically shared sounds, which suggests that the same underlying skills are utilized in both languages, whereas language-specific singleton consonants and consonant clusters did not appear to benefit from exposure to the other language. The results also suggest an intricate link between phonological skills and morpho-syntactic performance at early stages of development, but a more complex pattern thereafter with differences that may be based on language-specific phonological properties.

Introduction

The Latino population is the largest and fastest growing minority in the United States and by 2015 it reached 57 million (U.S. Census Bureau, 2016). California, home to 15 million Hispanics in 2014, continues to have the nation's largest Latino population among all states, and the Los Angeles-Long Beach-Anaheim metropolitan areas host almost half of the state's Hispanic population (Krogstad, 2016). Many of the children in these families begin learning English when entering the school system, while they are still developing Spanish. Indeed, the number of English language learners in U.S. schools has doubled in recent years and Spanish-speaking children represented 77% of all English learners in 2015 (National Center for Education Statistics, 2017). Some of these children might be exposed to both Spanish and English from birth and, by preschool age, they will have varying degrees of proficiency in these languages; others will develop a first language (L1) from birth and will start learning English – their second language (L2) – around 3, 4 or 5 years of age, often when entering preschool or kindergarten (Baker, 2017).

The numbers above suggest that the majority of children on the caseloads of speech-language pathologists (SLPs) in California (as well as other states) speak Spanish, or Spanish and English, with Mexican Spanish as the dialect most commonly used (Fabiano-Smith, Oglivie, Maiefski, & Schertz, 2015). In contrast, only 4% of professionals certified by the American Speech, Language, and Hearing Association (ASHA) were Spanish-language service providers in 2016 (ASHA, 2017). This suggests that the majority of clinicians working with young children in California may lack the necessary information to provide accurate assessment or treatment, because they might rely only on English measures to evaluate bilingual children and might over- or under-identify them for special education services (Kraemer & Fabiano-Smith, 2017). Clearly, there is an urgent need to examine speech production in both Spanish and English during the preschool years and understand more about typical Spanish-English bilingual phonological development. These data can help SLPs determine whether a child's performance is within typical limits or not and, therefore, provide a valid basis for assessment and intervention in this growing population. The preschool years are important to study closely because it is often during this time that formal exposure to English begins. This means that while children are learning Spanish, they are suddenly immersed in English, in a process that may dramatically shift the amount of exposure that they have to each language. This is also the time when young bilinguals may be referred for speech and language services, because in the English classroom, many of these children might not be as intelligible as their monolingual peers or might produce speech patterns not typically seen among monolinguals (Gildersleeve-Neumann, Kester, Davis & Peña, 2008).

Bilingual Phonological Development

The *Unified Competition Model* (MacWhinney, 2005) provides a framework to examine phonological development in dual language environments. This theoretical perspective proposes that information that is frequent and reliable in the input has strong cue strength and is therefore acquired more easily and rapidly. In the case of bilingual phonological development, phonological properties that are common across languages will result in frequent and reliable speech cues, leading to rapid development – or *positive transfer* (Goldstein & Bunta, 2012) – of these properties in both languages. This means that bilingual children may use sounds and phonological patterns from one language for productions in the other, ultimately attaining phonological skills that are comparable to those of monolingual children (Fabiano-Smith & Goldstein, 2010b). At the same time, properties that are unique across languages – such as sounds that are specific to only one language – will result in less frequent and less strong cues, leading to *negative transfer* (Goldstein & Bunta, 2012) and, ultimately, to some speech production skills in bilinguals that are more limited or different from those of monolinguals (Fabiano-Smith & Goldstein, 2010b). The model also predicts that negative transfer should be particularly seen at early stages of bilingual development and when one language is dominant over the other, since children would particularly rely on phonological structures from their stronger language – based on more frequent and possibly more salient cues – for productions in their weaker language.

Empirical support for the Unified Competition Model (Scarpino, 2011) has shown that since Spanish and English have many similar consonant properties, consonant cue reliability is strong and children appear to successfully rely on many Spanish consonants in their acquisition of phonetically similar English consonants (*shared sounds*). For example, young bilinguals may perceive stops, most nasals (/m/ and /n/), and some fricatives (/f/ and /s/) as common between Spanish and English; they may categorize them into the same phonemic category despite their fine phonetic distinctions (e.g., cross-linguistic differences in Voice Onset Time on stops); and they may produce them with more frequency than sounds unique to only one language (*unshared sounds*). On the other hand, production errors and protracted development are expected with language-specific sounds that do not have phonetically similar equivalents in the other language (for example, the Spanish trill or English /v/, /θ/ or /z/) because, overall, these sounds are less frequent in the input and as such, they result in weaker and less reliable speech cues. Although it is true that there is some overlap between consonants shared between Spanish and English and early developing sounds (stops, nasals etc.), high accuracy with these segments has been found independent of the effect of developmental sequence (Goldstein, Fabiano & Iglesias, 2003), suggesting that sound similarity may be an important variable affecting bilingual phonological development as predicted by the Unified Competition Model.

Speech Sound Production in Spanish-English Bilingual Children: Phonetic Inventories, Singleton and Cluster Accuracy, and Phonological Processes

Table 1 summarizes recent research on the phonological skills of Spanish-English bilingual children in the U.S. between 3 and 7 years of age (only studies that look at *overall* production abilities are included; studies on the acquisition of specific aspects, such as VOT or dialectal features, are excluded). As can be seen, most studies have examined phonetic inventories, segmental accuracy, as measured by overall percentage of vowels correct (PVC), percentage of consonants correct-revised (PCC-R; Shriberg, Austin, Lewis, McSweeney, & Wilson, 1997) and accuracy of early, mid- and late-developing sounds, and phonological processes or error patterns. Most studies have also been cross-sectional, rather than longitudinal, and have been based on small samples. In fact, once monolingual participants are excluded, the number of bilingual participants is rarely over ten (Brice, Carson & O'Brien, 2009; Prezas, Hodson &

Schommer-Aikins, 2014; Bunta, Fabiano-Smith, Goldstein & Ingram, 2009; Fabiano-Smith & Goldstein, 2010a, 2010b; Fabiano-Smith & Barlow, 2010; Gildersleeve-Neumann et al., 2008; Gildersleeve-Neumann, Peña, Davis & Kester, 2009).

INSERT TABLE 1 ABOUT HERE

Overall, some investigations of Spanish-English bilingual preschoolers in the early stages of English language acquisition show that these children are less accurate on some measures than monolingual peers of a similar age and these differences may persist until age 5 (Bunta et al., 2009; Fabiano-Smith & Goldstein 2010a, 2010b; Gildersleeve-Neumann et al., 2008; Gildersleeve-et al., 2009; Montanari, Subrahmanyam, Zepeda, & Rodriguez, 2014). Gildersleeve-Neumann and colleagues, for instance, found that the bilingual preschoolers in their study had lower intelligibility ratings, made more segmental errors, especially with vowels which differ considerably between Spanish and English, and produced more uncommon error patterns than monolingual English speakers of the same age. Likewise, Bunta and colleagues found that bilingual 3-year-olds had significantly lower scores on consonant accuracy measures than age-matched monolinguals. Fabiano-Smith and Goldstein also found that Spanish-speaking monolinguals outperformed bilingual preschoolers on measures of consonant accuracy; however, no differences were found between bilingual children and their English monolingual peers. In line with the Unified Competition Model (MacWhinney, 2005), these results can be interpreted as evidence for competition in linguistic input between the two languages, less frequent and less strong speech cues, and, ultimately, more limited phonological skills in bilinguals as compared to monolinguals (Fabiano-Smith & Goldstein, 2010b).

At the same time, other studies have reported commensurate skills between bilingual and monolingual children when examining other phonological constructs or when assessing older and more proficient bilinguals. Fabiano-Smith and Barlow (2010), for example, found Spanish and English phonetic inventories comparable to those of monolinguals in eight Spanish-English bilingual preschoolers. Furthermore, Goldstein, Fabiano and Washington (2005) reported consonant accuracy rates equivalent to those of monolinguals (over 90%) for both English and Spanish in a group of 15 bilingual children aged 5;0–5;5 years. Spanish and English accuracy levels commensurate to those reported for monolinguals were also found in a larger group of bilingual children (*N*=50) aged between 4;7 and 7;1 (Goldstein, Bunta, Lange, Rodriguez & Burrows, 2010). These findings can be taken as evidence of interaction between the two phonological systems: despite reduced input in each language and/or later onset of English acquisition, one language might be aiding in the acquisition of the other, as predicted by the Unified Competition Model (MacWhinney, 2005), allowing for speech production skills in two languages that are within the typical range for the children's chronological age (Fabiano-Smith & Goldstein, 2010b). This interaction would be particularly felicitous in the case of languages that share a considerable number of phonological properties, such as Spanish and English, since children would be using structures in one language for productions in the other. Fabiano-Smith and Goldstein (2010b) found a significant difference between the accuracy of shared and unshared sounds among Spanish-English bilingual preschoolers; Gildersleeve-Neumann et al. (2008) also interpreted their subjects' fewer errors with English consonants than with vowels as a sign of children's reliance on phonetically similar Spanish consonants.

Other investigators focusing on phonological processes have further found that bilingual children produce many of the same error patterns as their monolingual peers, but they also produce patterns that are not identical to those of monolingual children (Goldstein & Washington, 2001; Brice et al., 2009). Some of these patterns might be caused by negative transfer (Goldstein & Bunta, 2012), where children use consonants or vowels not shared by both languages in the other language (as when producing target /bʊk/, *book*, as [buk] or /klaβo/, "nail," as [klavo]) or they transfer fine-grained phonetic patterns from one language to the other (as when de-aspirating /t^host/, *toast*, as [tost]). Goldstein and Washington (2001), for instance, found that the most commonly occurring phonological processes produced by their 4-year-old subjects included stopping and final consonant deletion in English and liquid simplification and cluster reduction in Spanish, processes that were generally similar to those produced by typically developing monolingual English- and Spanish-speaking children. However, the percentage of occurrence of phonological processes differed for bilingual and monolingual children, and some processes were unique to the bilingual children. Gildersleeve-Neumann et al. (2008) also found that the English productions of their Spanish-English bilingual subjects contained a greater number of spirantization and final consonant deletions than those of monolingual peers. Nonetheless, Brice et al. (2009) and Prezas et al. (2014), who focused on cross-language comparisons, reported that percentage-of-occurrence means for all phonological processes were similar in the two languages of Spanish-English bilingual 4- and 5-year-olds, with the older children producing fewer phonological processes than the younger ones.

Overall, these findings suggest that bilingual children may perform similarly to monolingual peers and at comparable proficiency levels in both languages by age 5. However, as children are becoming bilingual during the preschool years, they may exhibit a fair amount of variability. It is therefore important to look closely at this time and examine how children's speech production skills relate to their overall linguistic ability in each of their languages.

Within-Language Relationships between Phonology and Other Domains and Cross-language Relationships between Phonological Skills

Differences in speech sound production between bilingual and monolingual children could be related to differences in language proficiency rather than bilingual status *per se* (in this study we will use the term *language proficiency* interchangeably with *language competence* or *language ability*, as done in other studies, e.g. Goldstein et al., 2010). For instance, as predicted by the Unified Competition Model (MacWhinney, 2005), negative transfer should be particularly seen at early stages of bilingual development and when one language is dominant over the other, since children would particularly rely on phonological structures from their stronger language – based on more frequent and possibly more salient cues – for productions in their non-dominant language (Munro, Ball, Müller, Duckworth & Lyddy, 2005; Mayr, Howells & Lewis, 2015). However, transfer from the dominant language should decrease as exposure to and practice with the non-dominant language increases, making cues specific to this language gain strength and exert more influence on the development of its phonological system. Goldstein and Bunta (2012), for instance, found that bilingual and monolingual 6-year-olds matched on language use and proficiency (as reported by parents) demonstrated comparable phonological skills. In addition, Goldstein et al.'s (2010) data revealed that a direct measure of language ability (mean length of utterance in words) was the best predictor of segmental accuracy between ages 4;7–7;1. These findings suggest that speech sound production is not independent of performance in other language domains.

Evidence that different language components affect one another is common in the bilingual literature. For example, Marchman, Martínez-Sussmann and Dale (2004) and Conboy and Thal (2006) documented strong, positive within-language correlations between lexical and grammatical abilities in bilingual toddlers. Evidence for within-language associations between phonology and other language components has also recently been found (Kehoe, 2011, 2015; Scarpino, 2011). Scarpino, for instance, found that a large portion of the variance in the phonological skills of 199 Spanish-English bilinguals between the ages of 37 and 77 months was predicted by vocabulary scores. These results are expected given that the development of a phonological system requires a lexicon in which to store the words phonologically and this lexicon develops in tandem with the phonological system (Scarpino, 2011). Furthermore, Cooperson, Bedore and Peña (2013), who sought to determine whether there were associations between English and Spanish phonological skills and performance in other language domains in kindergarten, found that the language measures that were most strongly correlated with phonology at 5 years of age were morpho-syntactic measures such as mean length of utterance (MLU) and percent grammaticality. In fact, the correlations of phonology with percent grammaticality, MLU, and the morpho-syntax scores obtained from the Bilingual English Spanish Assessment (BESA, Peña, Gutiérrez-Clellen, Iglesias, Goldstein, & Bedore, 2014) were stronger than the correlations of phonology with semantic (BESA semantics scores) and lexical measures (i.e. number of different words). These findings suggest that, by 5 years of age, speech sound production might be as related to morpho-syntactic performance as it is to vocabulary and thus this relationship should be taken into account when examining speech patterns in bilingual children.

The only two studies in the current literature that have looked at the link between morpho-syntactic and phonological skills have used mean length of utterance in words (MLUw) as an indicator of linguistic ability. MLU is the most widely used measure of morpho-syntactic skills and the best predictor of language ability in both monolinguals (Brown, 1973) and bilinguals (Bedore, Peña, Gillan & Ho, 2010). Because of the high correlations attested cross-linguistically between MLU measured in morphemes and MLU measured in words (Parker & Brorson, 2005), MLUw has been typically used to make comparisons between languages with different degrees of morphological complexity (Anderson, 1999). Aguilar-Mediavilla and Serra-Raventos (2006), who compared the speech sound production of five Spanish-Catalan bilingual 4-year-olds with language impairment to two typically-developing control groups, one matched on age and the other on proficiency (based on MLUw), found that children with greater MLUw had more advanced segmental abilities in Spanish than those with poorer language skills. Furthermore, Goldstein et al. (2010), the only study that specifically investigated the relationship between MLUw and phonological skills in Spanish-English bilingual children, found that this measure was the best predictor of segmental accuracy in English and Spanish between the ages of 4;3 and 7;1. Given these findings, it is important to examine the relationship between MLUw and phonological skills in the preschool years, when most of the variability in speech performance has been observed.

In addition to documenting a link between phonological and morpho-syntactic performance *within each* language, Cooperson et al. (2013) found that phonological skills were also moderately correlated *across languages.* Scarpino (2011) also documented positive, significant relationships between English and Spanish speech production abilities in a large sample of Spanish-English bilingual 4-6-year-olds. In fact, her results showed that speech production proficiency in one language was the best predictor of segmental accuracy in the other language. Other studies of bilingual children learning other language pairs have documented correlations and interaction between phonological systems (Mayr et al., 2015; Kehoe, 2015). Taken together, these findings suggest, in line with the predictions of the Unified Competition Model (MacWhinney, 2005), that speech production abilities in one language may be highly predictive of speech production abilities in the other language, at least when the languages share similar sounds and phonological structures. That is, Spanish production skills might be aiding in the acquisition of English, allowing for phonological skills in two languages that are commensurate to those of monolingual peers (as shown by Fabiano-Smith & Barlow, 2010; Goldstein et al., 2005; and Goldstein et al., 2010) and also similar across languages (as found by Cooperson et al., 2013, and Scarpino, 2011).

Therefore, it appears that there are both *within-language* (i.e. language-specific) relations between phonology and skills in other language domains, and *between-language* (i.e. language-general) relations between phonological skills in different languages. Links between phonology and grammatical and lexical performance within a language suggest that children's phonological abilities develop together with skills in other language components, in particular, vocabulary and syntax. At the same time, cross-linguistic relationships in speech sound production may suggest that sound learning is also based on the same speech motor abilities that provide the foundation for the acquisition of two sound systems. Indeed, physiological factors have a profound effect on children's early sounds irrespective of the language they are learning (Velleman & Vihman, 2007) and physiology has an impact in later stages as well (Kent, 2000). In turn, the articulatory ease and perceptual discriminability of sounds predict universal markedness patterns, similarly shaping phonological development cross-linguistically (Locke, 1983). Thus, it is possible that speech sound production in two languages is partially dependent on the same articulatory constraints not only at the onset of speech, but also at later stages of development.

Purpose of the Study

It is the goal of the present study to provide longitudinal data on Spanish-English phonological skills during the preschool years and examine the links between speech sound and morpho-syntactic performance as well as between phonological skills in the two languages. In particular, we assess English and Spanish singleton consonants and consonant clusters between 3;7 and 4;7 in a group of 35 American English-Mexican Spanish bilingual children with different proficiency levels in each language and at each age as measured by mean length of utterance in words (MLUw). We focus on changes in a variety of phonological measures including (a) overall accuracy of consonant (PCC-R) and cluster (PClC) production, (b) accuracy of production of different sound classes, (c) accuracy of production of sounds and clusters shared and unshared between English and Spanish, and (d) phonological processes/error types. We also examine the extent to which speech production skills are correlated both with MLUw and across languages both in the first and second year of preschool.

On the basis of the extant literature, we hypothesized that (a) phonological skills would increase over time in both languages; (b) accuracy with shared sounds and clusters would be higher than with unshared sounds and clusters; (c) phonological processes would systematize and decrease over time as the children got older and became increasingly proficient in both Spanish and English; (d) singleton consonant and cluster accuracy would be correlated with language proficiency, as measured by MLUw, and across languages, suggesting shared speech motor abilities.

Methods

Participants

Data from 35 typically-developing children (16 boys, 19 girls) from a larger pool of 48 children who participated in a longitudinal study of dual language development in bilingual Spanish-English speakers were selected for this study. Children's data were included in this analysis if audio recordings of the Bilingual English Spanish Assessment (BESA, Peña et al., 2014) phonology subtest and conversational speech samples were available in both languages and both in the fall of their first preschool year (at approximately 3 and a half, Time 1) and in the fall of their second preschool year (at 4 and a half, Time 2). Data from a total of 35 children were available for analysis.

The children's mean age was 3;7 (age range: 3;0-4;2) at Time 1 and 4;7 (age range: 4;0-5;2) at Time 2. All children came from Mexican-origin families who lived in Los Angeles County and were enrolled in a Head Start program during the time of the study. All families came from low SES backgrounds as determined by the children's eligibility to participate in the program, which is specifically designed to promote school readiness in children between birth and age 5 from low-income families (https://www.acf.hhs.gov/ohs). Indeed, all mothers had limited education, with the majority of them having completed high school (65.7%) or only primary education (28.6%), as shown in Table 2. All children were typically developing and had no hearing, speech, language, cognitive, or neurological deficits based on parental reports and program screening tests. In order to determine the children's length and amount of exposure to, use of, and proficiency in each language, a detailed parent questionnaire was administered in Spanish. All parents reported a Mexican background/nationality and the Mexican Spanish variety as the language spoken in the home. All parents indicated that Spanish was the language that the child learned first and that was most used at home. As shown in Table 2, 100% of the parents indicated that they spoke only or primarily Spanish with their children from birth. At the same time, all parents reported that their children had also had some exposure to English before entering the preschool program through siblings, media, and the larger community.

Classroom observations provided additional information about the language environment in the school. The Head Start program followed an English immersion model; for instance, for most children, formal and consistent exposure to English had begun in preschool. However, some teachers and classroom assistants were bilingual and Spanish was not discouraged in the classroom. Thus, the participants in this study received input in, and spoke, both Spanish and English, but Spanish tended to be the home language and English the school language.

INSERT TABLE 2 ABOUT HERE

Data Collection

In order to examine phonological skills as related to language abilities, MLUw values were calculated from conversational speech samples collected at the preschool both at Time 1 and Time 2 by separate research assistants in Spanish and English. A pre-determined set of toys and books that included stuffed animals, dolls, cars and books depicting common and culturally-appropriate themes (going to school, attending a picnic, a birthday party, etc.) was used to ensure consistency across the samples. The samples were then transcribed orthographically and analyzed by two additional research assistants using Systematic Analysis of Language Transcripts (Miller & Iglesias, 2008) to calculate MLUw values. Inter-judge reliability of transcription and coding, calculated on 30% of the data, reached 96.7% for English and 98.2% for Spanish. The obtained MLUw values confirmed that most children (24 out of 35) were dominant in Spanish at Time 1 (see Table 2). Specifically, a paired-samples *t*-test comparing the participants' MLUw across the two languages suggested a significantly higher MLUw overall in Spanish (*mean:* 2.33, *SD* = .58) than English (*mean:* 2.03, *SD* = .56): $t(34)= -2.987$, $p=.005$. However, by Time 2, English MLUw values (*mean:* 3.13, *SD* = .89) were no longer significantly different from Spanish ones (*mean:* 3.06, *SD* = .88): $t(33)= 0.527$, $p=.601$, suggesting that most children had similar proficiency levels across their two languages.

As in several previous studies of Spanish-English bilingual phonological development (see Table 1), the phonology subtest of the Bilingual English Spanish Assessment (BESA; Peña et al., 2014) was used to elicit single word samples and assess singleton consonant and cluster production abilities both at Time 1, when children had been in the program for no longer than three months, and a year later (Time 2). As shown in the Appendix, the assessment contains 31 separate target items for English and 28 separate target items for Spanish, varying in length and lexical stress pattern. The assessment targets all singleton consonants in Spanish and English (except /ʒ/ in English) between one and seven times, either in syllable-initial, syllable-final, or both positions, as well as commonly occurring consonant clusters (e.g., /pl/, /kw/, /tr/, etc.). In particular, the English assessment targets a total of 47 consonants (15 plosives, 9 nasals, 12 fricatives, 3 affricates, and 8 approximants) as well as 16 tautosyllabic clusters. The Spanish assessment targets a total of 54 consonants (14 plosives, 7 nasals, 17 fricatives, 1 affricate, 3 flaps, 6 trills, and 6 approximants) as well as 8 tautosyllabic clusters.

Procedure

Single words were elicited at the preschool using high-quality pictures that depicted the target items. Children were asked to name the object depicted in each picture. In the case of no response, children were given prompts or were allowed to repeat the answer provided by the examiner, as in previous studies, on the basis that spontaneous and imitated responses tend to be very similar (Goldstein, Fabiano, & Iglesias, 2004). Each session was administered in one language at a time (Spanish or English) by the same research assistants who had collected the conversational speech samples. All research assistants were native speakers of English and Spanish and only interacted with the child in the language of testing. Half the children completed the BESA Spanish phonology subtest first, while the other half completed the English version first.

Samples were recorded using an Edirol R-09HR High-Resolution WAVE/MP3 recorder and a desktop microphone in close proximity to the child. All digital files were transferred from the digital recorder and stored on a desktop computer. Each sample was then independently transcribed by the first author and by two graduate students in narrow phonetic transcription, using the conventions of the International Phonetic Alphabet (IPA, 1999). One student transcribed all Spanish data, while the other worked on the English data. All transcribers were phonetically trained and bilingual in Spanish and English. The two independent transcriptions in each language were then compared for reliability purposes. Inter-rater reliability, calculated for 100% of the target consonants and clusters, was 96% for Spanish and 94% for English. In order to perform intra-judge reliability on their own transcriptions, the three transcribers re-transcribed 100% of the samples; intra-rater reliability was 98% for the first author and 96% for the graduate assistant for Spanish, and 96% for the first author and 95% for the graduate assistant for English. Disagreements on sounds were discussed by listening to the recordings several more times until consensus was reached. The consensus transcriptions were used for the analysis.

Analyses

Overall consonant accuracy was calculated in terms of Percent of Consonants Correct-Revised (PCC-R) (Shriberg et al., 1997) as in most studies of Spanish-English bilingual children. PCC-R indicates the percentage of consonants that the child has produced correctly out of the total number of consonants targeted, although speech sound distortions – which are common in typical speech development in monolinguals as well – are not coded as errors. Likewise, productions that differed on fine phonetic detail (such as substituting less aspirated syllable-initial voiceless stops as in *toast* [toust]) were not coded as errors as long as they were target-like in terms of place and manner of articulation as well as voicing. PCC-R was also calculated for different sound classes (stops, fricatives, nasals, etc.) and for sounds shared between Spanish and English, that is, sounds that are “phonetically similar” (Flege, 1981) and present in both languages (/p, b, t, d, k, g, m, n, f, s, ð, tʃ, l, w, j/) and for unshared sounds, that is, sounds that exist only in English (/ŋ, v, z, ʃ, θ, h, dʒ, ɹ/) or in Spanish (/ɲ, β, ɣ, x, r, ɾ[1]/), as in Fabiano-Smith and Goldstein (2010b). Likewise, we calculated the percentage of consonant clusters that the children had produced correctly out of the total number of clusters targeted (PClC) as well as the accuracy of production of different classes of consonant clusters and of shared (/kl, bl, pl, fl, kw) and unshared (/bɹ, pɹ, fɹ, tɹ, gɹ, dɹ, nt(s), nd, st, skɹ, tɾ, bɾ, kɾ, fɾ/) clusters. Only tautosyllabic clusters were used in the analyses. As in previous studies, we took the varieties spoken by the children into account and compared the children’s productions to adult targets in American English and Mexican Spanish (as outlined in Goldstein, 2001).

Statistical comparisons of sound and cluster accuracy were completed using the Statistical Package for the Social Sciences (SPSS Statistics Version 22). A series of one-way ANOVAs with *time* as within-subjects factor (2 levels: Time 1; Time 2) was conducted in each language to examine the effect of time on singleton consonant and cluster accuracy measures. Moreover, additional one-way ANOVAs with *cross-linguistic overlap* as within-subjects factor (2 levels: shared; unshared) were carried out to investigate whether the participants’ phonological skills differed on cross-linguistically shared and unshared items. Correlation analyses were run on the accuracy measures in Spanish and English to assess whether speech production abilities were related cross-linguistically. Moreover, we correlated MLUw values and segmental and cluster accuracy scores at each time and in each language to examine the relation between phonological skills and language proficiency, and the extent to which it changed over time. The alpha level was adjusted for multiple comparisons throughout using the Holm-Bonferroni method (Holm, 1979).

The next analysis involved examining the types of phonological processes produced by the children and their percentage of occurrence at each time, focusing on the changes in error patterns over

[1]Although the flap exists in English as an allophone of /t/ and /d/, we coded it as a Spanish-only sound because it is phonemic only in Spanish (as in Fabiano-Smith & Goldstein, 2010b).

time. Errors on singletons were coded in terms of the following categories: affrication, deaffrication, deletion, devoicing, flap/trill deviation, flapping, fricative simplification, fronting, gliding, glottal substitution, spirantization, and stopping, and errors on clusters were classified according to the following processes: reduction, simplification, deletion, consonant insertion, and metathesis. These categories were chosen as they have been attested in previous studies on phonological development in monolingual and bilingual children learning English and Spanish as well as other language pairs (e.g. Gildersleeve-Neumann et al., 2008, 2009; Goldstein & Bunta, 2012; Mayr, Jones & Mennen, 2014; Mayr et al., 2015; Prezas et al., 2014; Shriberg & Kwiatkowski, 1980).

Results

Accuracy Measures

To examine whether the participants' speech production skills increased over time in the two languages, a series of one-way within-subjects ANOVAs was conducted. The alpha level was adjusted for multiple comparisons in each language using the Holm-Bonferroni method (Holm, 1979). The results, depicted in Tables 3 and 4, revealed significantly better performance in both languages at Time 2 than Time 1 for overall accuracy on singleton consonants (PCC-R English: $F(1,34)= 19.218$, $p< .001$, $\eta^2= .361$; PCC-R Spanish: $F(1,34)= 13.011$, $p= .001$, $\eta^2= .277$) and consonant clusters (PClC English: $F(1,34)= 61.515$, $p< .001$, $\eta^2= .644$; PClC Spanish: $F(1,34)= 27.291$, $p< .001$, $\eta^2=.445$), as well as for shared and unshared clusters (PClC-shared English: $F(1,34)= 23.272$, $p< .001$, $\eta^2= .406$; PClC-unshared English: $F(1,34)= 62.2$, $p< .001$, $\eta^2= .647$; PClC-shared Spanish: $F(1,34)= 18.066$, $p< .001$, $\eta^2=.347$; PClC-unshared Spanish: $F(1,34)= 17.50$, $p<.001$, $\eta^2=.340$). Shared and unshared singletons also showed significant improvements over time, however only in English (PCC-R shared: $F(1,34)= 20.256$, $p< .001$, $\eta^2= .373$; PCC-R unshared: $F(1,34)= 11.654$, $p= .002$, $\eta^2= .255$). The same held true for specific manner classes, with only English stops, fricatives, nasals and approximants showing a significant effect of *time* (PCC-R stops: $F(1,34)= 5.867$, $p= .021$, $\eta^2= .147$; PCC-R fricatives: $F(1,34)= 14.735$, $p= .001$, $\eta^2= .302$; PCC-R nasals: $F(1,34)= 12.989$, $p= .001$, $\eta^2= .276$; approximants: $F(1,34)= 27.344$, $p< .001$, $\eta^2= .446$).

Taken together, these results are consistent with our hypothesis that the participants' speech production skills would increase over time in both languages.

INSERT TABLES 3 AND 4 ABOUT HERE

Inspection of Tables 3 and 4 shows consistently higher mean values on shared than unshared items at both time points. To examine whether the participants' accuracy differed on shared and unshared singletons and clusters, we ran a series of one-way ANOVAs with *cross-linguistic overlap* (2 levels: shared and unshared) as within-subjects factor. The alpha level was again adjusted using the Holm-Bonferroni method (Holm, 1979). The results revealed significant effects of *cross-linguistic overlap* for both languages and on both singletons and clusters (English time 1 PCC-R: $F(1,34)= 137.324$, $p<.001$, $\eta^2= .802$, adj. $\alpha= .0167$; English time 2 PCC-R: $F(1,34)= 286.577$, $p<.001$, $\eta^2= .894$, adj. $\alpha= .0125$; English time 1 PClC: $F(1,34)= 61.617$, $p<.001$, $\eta^2= .644$, adj. $\alpha= .025$; English time 2 PClC: $F(1,34)= 57.478$, $p<.001$, $\eta^2= .628$, adj. $\alpha= .05$; Spanish time 1 PCC-R: $F(1,34)= 161.653$, $p<.001$, $\eta^2= .826$, adj. $\alpha= .0125$; Spanish time 2 PCC-R: $F(1,34)= 185.987$, $p<.001$, $\eta^2= .845$, adj. $\alpha= .0167$; Spanish time 1 PClC: $F(1,34)= 32.872$, $p<.001$, $\eta^2= .492$, adj. $\alpha= .025$; Spanish time 2 PClC: $F(1,34)= 24.638$, $p<.001$, $\eta^2= .42$, adj. $\alpha= .05$). This confirms our hypothesis that the participants would be more accurate on shared than unshared items in both languages.

Error Patterns

A total of 1185 English tokens and 1002 Spanish tokens were identified as errors. Of these, 1534 occurred on singleton targets and 653 in clusters. Tables 5 and 6 depict a breakdown of error patterns in English and Spanish, respectively, with all error types accounting for at least 1% of tokens displayed.

INSERT TABLES 5 AND 6 ABOUT HERE

Inspection of the tables shows that the most common processes on singletons were *deletion* and *stopping*, which occurred in both languages, as well as *devoicing* in English, and *flap/ trill deviation* in Spanish. Interestingly, the number of deletions decreased over time, while there was little change for the other processes. Stopping (e.g. /ˈfɛðɚ/, *feather,* realized as [ˈfɛdɚ]) and flap/trill deviations (e.g. /aˈros/, "rice," realized as [aˈjos] and /ˈseɲoɾ/, "man," realized as [ˈseɲol]) may have persisted as these processes involved late acquired sounds. Indeed, 67% of English stopping errors at Time 2 were substitutions for /θ/ and /ð/. In Spanish, in contrast, the persistence of stopping errors is likely a result of incomplete acquisition of the stop-spirant alternation rule (Barlow, 2002; Fabiano-Smith et al., 2015), which requires underlying stops to be realized as stops when they occur utterance-initially, after homorganic nasals, and in the case of /d/ after laterals, while spirants occur in all other contexts (Branstine, 1991). For instance, 93% of stopping errors at Time 2 occurred on /β/, /ð/ and /ɣ/ when children produced items such as /ˈklaβo/, "nail," /ˈraðio/, "radio," and /aˈɣuxa/, "needle," as [ˈklabo], [ˈradio] and [aˈguxa], respectively. Conversely, all spirantization errors at Times 1 and 2 involved contexts where /b/, /d/ and /g/ would have been required. As to devoicing errors, they all involved word-final obstruents in English, as when children produced /ˈfɹɑg/, *frog,* as [ˈfɹɑk]. The persistence of these patterns is not surprising since the existence of a voicing contrast in final position is marked in the languages of the world (Eckman, 1977) and known to be developmentally protracted (Goldstein & Bunta, 2012; Stoel-Gammon & Dunn, 1985). At the same time, since Spanish does not allow a voicing contrast in final position, it is also possible that devoicing errors were the result of cross-language transfer.

In addition to singletons, the children also made errors on their cluster productions. As shown in Tables 5 and 6, these predominantly involved reduction (e.g. /ˈpleɪt/, *plate,* → [ˈpeɪt]) and simplification (e.g. /ˈpɹɛzənt/, *present,* → [ˈpwɛzənt]) patterns, while cluster deletions (e.g. /ˈglæs/, *glass,* → [ˈæs]) were rare. While the error rate on clusters decreased substantially in both languages, it was implemented differently in Spanish and English. Thus, on English clusters there was a steady decrease over time that affected reduction and simplification processes equally, while in Spanish only cluster reductions decreased but simplifications increased. Note that the decrease in reductions was more substantial in Spanish than English and likely resulted in a trade-off between reductions and simplifications, with clusters reduced at Time 1 in Spanish (e.g. /ˈkɾus/, "cross," → [ˈkus]) exhibiting simplification patterns at Time 2 (e.g. /ˈkɾus/ → [ˈklus]) (cf. Greenlee, 1974; Kirk & Demuth, 2005).

Finally, the error patterns were examined for cross-linguistic interactions. Due to the absence of monolingual control groups, we were unable to determine whether our participants' patterns were due to acceleration or deceleration effects (Goldstein et al., 2005, 2010; Paradis & Genesee, 1996). Instead, in the present study, interactions were defined narrowly in terms of negative transfer or the use language-specific segments or clusters in the other language, as in Goldstein and Bunta (2012). According to this definition, the incidence of interactions was low, in line with previous studies (e.g., Fabiano-Smith & Goldstein, 2010b; Goldstein & Bunta, 2012). Overall, only 27 English tokens at Time 1 (1.2%) and 32 at Time 2 (1.4%) could be attributed to transfer from Spanish. They include spirantization errors, e.g. /ˈwægən/, *wagon,* realized as [ˈwæɣən], /h/ produced as [x] and the use of [ɾ] in English clusters. On the Spanish productions, 23 tokens suggested transfer from English at Time 1 (1.1%) and 16 at Time 2 (0.8%). These mostly included Spanish /r/ and /ɾ/ being realized as [ɹ], in both singleton contexts and in clusters. Taken together, the results confirm our hypothesis that error patterns would systematize and decrease over time as children became more proficient in Spanish and English.

Correlations between Phonological Skills and MLUw and Cross-linguistic Correlations between Phonological Skills

In order to examine whether phonological development is related to MLUw, we ran a series of correlation analyses, conducted separately for Time 1 and Time 2, and for each language. The results, depicted in Table 7, show moderate positive correlations on PCC-R and PClC at both data points in English (PPC-R time 1: r=.403, p=.016, adj. α= .05; PCC-R time 2: r=.460, p=.006, adj. α= .025; PClC time 1: r= .538, p=.001, adj. α= .025; PClC time 2: r= .433, p= .011, adj. α= .05), and at Time 1 for Spanish (PCC-R: r= .415, p=.013, adj. α= .05; PClC: r= .462, p= .005, adj. α= .025). The magnitude of association was based on Evans' (1996) account: 00-.19: "very weak"; .20-.39: "weak"; .40-.59: "moderate"; .60-.79: "strong"; .80-1.0: "very strong". These results suggest that as children increase the linguistic complexity of their utterances, their speech production becomes more accurate, in line with our hypothesis. However, this pattern did not hold across the board, since the Spanish phonological measures at Time 2 were not significantly correlated with MLUw (PCC-R: r=.157, p=.368, adj. α= .05; PClC: r= .265, p=.124, adj. α= .025).

INSERT TABLE 7 ABOUT HERE

In addition to correlations with MLUw, we wanted to determine whether phonological development is related across Spanish and English. To this end, we carried out a series of cross-linguistic correlation analyses. The results, depicted in Table 8, indicate moderate to strong positive correlations on singleton consonants overall (PCC-R time 1: r= .732, $p<.001$, adj. α= .0125; PCC-R time 2: r=.431, p= .01, adj. α= .0167) and consonant clusters overall (PClC time 1: r= .770, $p<.001$, adj. α= .01; PClC time 2: r= .479, p=.004, adj. α= .0125) as well as on all shared items (PCC-R shared time 1: r= .797, $p<.001$, adj. α= .0083; PCC-R shared time 2: r= .606, $p<.001$, adj. α= .0083; PClC shared time 1: r= .642, $p<.001$, adj. α= .0167; PClC shared time 2: r= .573, $p<.001$, adj. α= .01). These results confirm our hypothesis that phonological abilities are related across languages, in line with Scarpino's (2011) study. However, this was not the case for unshared items, with the exception of clusters at Time 1 (PCC-R unshared time 1: r= -.061, p= .727, adj. α= .05; PCC-R unshared time 2: r= .025, p= .888, adj. α= .05; PClC unshared time 1: r= .598, $p<.001$, adj. α= .025; PClC unshared time 2: r= .277, p=.107, adj. α= .025), suggesting that they develop independently in the two languages.

INSERT TABLE 8 ABOUT HERE

Discussion

The purpose of this study was to examine Spanish-English phonological development as related to language proficiency. Specifically, we examined English and Spanish speech production abilities between 3;7 and 4;7 in a group of 35 American English-Mexican Spanish bilingual children with different proficiency levels in each language and at each age as measured by MLUw. We focused on changes in a variety of phonological measures including (a) overall accuracy of consonant (PCC-R) and cluster (PClC) production, (b) accuracy of production of different sound classes, (c) accuracy of production of sounds and clusters shared and unshared between English and Spanish, and (d) phonological processes/error types. We also examined the extent to which phonological skills were correlated both with MLUw and across languages during the preschool years. In the following discussion we examine first the growth of phonological skills and error patterns during the preschool years. We then focus on the within-language links between morpho-syntactic performance and speech sound production and the between-language relationships between phonological skills.

Growth of Phonological Skills during the Preschool Years

The first main finding of this study is that accuracy on singleton consonants and clusters increased significantly in both languages over the course of one year, although growth appeared more pronounced in English than in Spanish. Indeed, while children produced most manner classes more accurately at Time 2 than at Time 1 in English, statistically significant differences between Time 1 and Time 2 were only found for overall PCC-R in Spanish. Mean PCC-R was 71.11% in English and 76.19% in Spanish at Time 1, and 82.18% in English and 80.64% in Spanish at Time 2. This means that within a year, children's English and Spanish consonant accuracy improved between 4 and 11 percentage points, clearly showing a developmental pattern. Development was even more pronounced for clusters, whose accuracy improved from around 46-49% at Time 1 to approximately 75% at Time 2 in both languages. These findings are important because the only previous longitudinal studies of Spanish-English bilingual preschoolers' phonological skills (Gildersleeve-Neumann et al., 2008, 2009) failed to document growth over time. It is possible that in those studies, the interval between the two data collection time points was too short since the participants were tested eight months apart in English and six months apart in Spanish. In the current study, on the other hand, preschoolers' phonological skills grew in both Spanish and English over a one-year period, providing evidence that speech sound production can improve in both languages even when children are only instructed in English. At the same time, growth was more evident in English and, unlike in Spanish, it affected most sound classes, confirming previous findings that the acquisition of the majority language occurs relatively fast in this population and may soon outpace home language development (Fabiano-Smith & Goldstein, 2010b; Ruiz-Felter, Cooperson, Bedore, & Peña, 2016). Children's language environment – Spanish at home and English at school – may also help explain this asymmetry: it is possible that increased input in, and emphasis on, English *at school* specifically increased accuracy in English relative to Spanish (Ruiz-Felter et al., 2016).

Overall, accuracy rates were very similar to those in other studies. For example, English PCC was 72.31% for the children (mean age: 3;6) in Fabiano-Smith and Goldstein (2010a), 70.18% for the English-dominant bilinguals (mean age: 3;6) in Gildersleeve-Neumann et al. (2008), and 72.64% for the participants (mean age: 3;10) in Montanari et al. (2014). In Spanish, our participants had Time 1 accuracy rates that were similar to those reported for the bilinguals in Gildersleeve-Neumann et al. (2009) (77.7%) but also comparable to those of the monolinguals in Fabiano-Smith and Goldstein (2010a) (74.08%). A year later (at 4 and a half), around 80% of consonant and 75% of cluster productions were accurate in both languages. These values suggest that by kindergarten, the children will likely reach consonant and cluster accuracy levels around 90% as reported for Spanish-English bilingual 5-6-year-olds from Texas and Pennsylvania (Goldstein et al., 2010; Goldstein et al., 2005; Ruiz-Felter et al., 2016). Taken together, our findings, based on children in California, contribute to and validate previous work on the speech sound production abilities of young Spanish-English bilinguals in the USA. A clinical implication is that professionals working with this population should expect growth in phonological skills in both languages during the preschool years, despite English-only schooling, and should make use of the typical performance levels reported in the extant research to assess bilingual children and identify atypical development.

Although we did not systematically compare accuracy rates for early-, middle-, and late-acquired sound classes, the results showed patterns that were similar to those found in previous studies (e.g. Fabiano-Smith & Goldstein 2010a; Ruiz-Felter et al., 2016), revealing, in both Spanish and English, higher mean accuracy rates for early-developing sounds (i.e. stops and nasals) than late-developing sounds (i.e. trills and flaps in Spanish and fricatives in English). Conversely, as shown in Tables 3 and 4, variation across children for the different manner classes tended to increase from early- to late-developing segments, especially in Spanish. This means that accuracy was more uniform for sounds that are typically acquired early while children differed in how good they were at producing late-acquired segments. These findings suggest that variability in performance, especially with late-developing sounds, may be typical of bilingual children at this age and thus professionals should take this variability into account when providing assessment and intervention.

Error Patterns

The error analysis revealed patterns that were similar to, but also different from, those reported in previous research on bilingual Spanish-English children (Gildersleeve-Neumann et al., 2008, 2009; Goldstein & Washington, 2001; Prezas et al., 2014). Indeed, at 3;7, the most common errors affecting English singletons were consonant deletion, stopping and devoicing, typical errors that have been documented in the literature for English-learning children (Stoel-Gammon & Dunn, 1985). However, the percentage of occurrence of these processes was quite low, with deletion occurring on 9.6% of tokens at Time 1, stopping on 5.4%, and devoicing on 3.3%. These rates are much lower than those reported by Gildersleeve-Neumann et al. (2008) for their 3-year-old bilingual and monolingual subjects, suggesting that the bilinguals in our study had quite advanced English speech production abilities – and possibly overall higher proficiency – compared to Gildersleeve-Neumann et al.'s participants. In Spanish, the most common errors affecting singletons were consonant deletion, flap/ trill deviations, and stopping, which also occurred infrequently at 9.2%, 7.6%, and 2.9% of the time, respectively. Gildersleeve-Neumann et al. (2009) also reported a low percentage of occurrence of stopping errors (6%), but consonant deletions and flap and trill deviations were much higher. By Time 2, when children were 4;7, consonant deletion had decreased in both languages, and the percentages of occurrence of this process (5.4% in English and 3.2% in Spanish) were similar to those reported for the 4-year-old bilinguals in Prezas et al. (2014) and Goldstein and Washington (2001). These results are in line with the finding that errors with consonant deletion largely decrease between 3 and 5 resulting in a significant increase in speech intelligibility (Stoel-Gammon & Dunn, 1985).

However, there were errors with singletons that showed less change over time in both languages. In English, children continued to devoice and stop consonants at a rate of 2.8% and 4.5%, respectively, possibly due to the late-developing nature of the sounds affected. Final consonant devoicing was also the most frequent error for the children in Gildersleeve-Neumann et al. (2008), although error rates were much higher than those in the present study. On the other hand, Goldstein and Washington (2001) reported a similar percentage of occurrence of stopping for 4-year-old bilinguals. In Spanish, stopping and flap/trill errors persisted at Time 2, occurring in 4.8% and 7.3% of instances, respectively, as documented in other studies (Goldstein & Washington, 2001). As pointed out above, the persistence of stopping errors is likely a result of incomplete acquisition of the stop-spirant alternation rule (Barlow, 2002; Fabiano-Smith et al., 2015). As to flap/trill errors, the results are in line with previous studies that document high rates of liquid simplifications at this age (Gildersleeve-Neumann et al., 2009; Goldstein & Washington, 2001; Prezas et al., 2014), although rates were higher in those studies, possibly due to differences in the analysis procedures and/or to the larger number of participants in the current study. Irrespective of their specific percentage of occurrence, the extant research confirms that flap and trill deviations are the most common production errors among Spanish-English bilingual preschoolers, as also documented for same-age Spanish-speaking monolinguals (Goldstein & Iglesias, 1996).

When it comes to cluster errors, these predominantly involved reduction and simplification patterns. In English, cluster reduction errors decreased from 39.1% to 18.9% over one year, suggesting a level of improvement that was parallel to the one observed among the English-dominant bilinguals in Gildersleeve-Neumann et al. (2008). Cluster simplification errors similarly decreased from 11.7% to 4.9% between Time 1 and Time 2, suggesting that by the second year of preschool, errors with consonant clusters decline, significantly improving speech intelligibility (Stoel-Gammon & Dunn, 1985). In Spanish, errors with consonant clusters included primarily reductions at Time 1 (39.8%); however, by Time 2, reductions had decreased (9.6%), while simplifications had increased (15.4%). These results might suggest a trade-off between reductions and simplifications, with clusters reduced at Time 1 exhibiting simplification patterns at Time 2 (cf. Greenlee, 1974; Kirk & Demuth, 2005). Overall, these findings are in line with previous studies that document high rates of cluster reduction errors in both Spanish and English among Spanish-English bilingual preschoolers (Gildersleeve-Neumann et al. 2008; 2009; Prezas et al., 2014).

Consonant and Cluster Accuracy as Related to Language Ability

An important finding of this study was that singleton consonant and cluster accuracy were moderately correlated with a direct measure of language ability, i.e. MLUw, in the same language, suggesting that speech production abilities develop together with skills in other language domains. This was especially true in English at both times and in Spanish at Time 1, when PCC-R and PClC were significantly and positively linked to MLUw. These results are in line with Cooperson et al.'s (2013) findings which also showed a close relation between phonological and morpho-syntactic development. A possible interpretation is that overall language ability drives speech sound performance. As predicted by the Unified Competition Model (MacWhinney, 2005), limited exposure to a non-dominant language may indeed result in infrequent and unnoticeable speech and language cues in the input, leading to protracted development of overall language abilities and, in turn, to restricted practice with, and ability to, reproduce its sounds. On the other hand, extensive exposure to a language – and to its frequent speech cues – may result in higher language skills, increasing experience with its sounds, words, and sentences and improving speech sound proficiency.

Alternatively, it could be speech sound production that drives overall language ability, rather than the other way around. For instance, in the context of language impairment, Chiat (2001) argued that it is morpho-syntactic performance that is dependent on a child's phonological capabilities and not the other way around. Impaired phonological processing can indeed disrupt the mapping processes through which word and sentence structure is acquired, affecting lexical and syntactic development in children with language impairment. Applied to typical phonological development, this theory predicts that specific speech motor abilities might be required in order for children to combine words and produce utterances of increasing length. Thus, children with advanced speech production skills would also demonstrate increased language ability as measured by MLUw, while limited speech production abilities would go hand in hand with lower utterance length. While our findings demonstrate a link between segmental and cluster accuracy and morpho-syntactic skills, we cannot ascertain whether it is speech production that affects utterance length or vice versa as correlation does not imply causation and, therefore, both interpretations are plausible.

Interestingly, by Time 2, singleton and cluster production were correlated with morpho-syntactic performance only in English, but not in Spanish. These results are surprising considering the children displayed similar MLUw values – around three words per utterance – in their two languages at 4;7. It is possible that MLUw is a rough measure that no longer captures morpho-syntactic ability in Spanish at later stages of development. It is also possible that the link between speech sound and morpho-syntactic performance is particularly strong in the early stages of multi-word production when speech motor abilities have been shown to be a prerequisite for increasing utterance length (Dromi, 1987). This was indeed the case at Time 1, when children's average utterance length was around two words in both languages. However, by Time 2, when children produced an average of three words per utterance, speech motor abilities might have been a prerequisite for utterance length in English but not in Spanish, a language with relatively simple syllabic structures and a higher statistical frequency of words with three or more syllables (Astruc, Payne, Post, Vanrell & Prieto, 2012). Indeed, Spanish-learning children can produce multisyllabic words much earlier than children learning English or German (Lleó, 2006) as they do not have to deal with complex syllabic structures and they can concentrate on increasing the number of syllables. In other words, the speech production abilities needed to produce 3-word utterances in Spanish may possibly be different from those needed to produce 3-word utterances in English, and hence the relationship between segmental production and utterance length may be mediated by the phonological characteristics of the language involved. There is, however, a need for further investigation of the relationship between phonological and morpho-syntactic measures at later stages of multi-word production.

In practice, the findings presented in this section suggest that the variability in phonological performance that has been observed in the preschool years might be explained by the heterogeneity of bilingual children's proficiency levels. For instance, children who are highly proficient in their two languages might be as accurate on both English and Spanish phonological measures as monolinguals, or even more so. While our study does not allow us to reach this conclusion given the absence of monolingual control groups, this was the case for the 6-year-olds in Goldstein and Bunta (2012) whose phonological skills were compared to age-matched monolinguals with the same levels of language use and proficiency. However, at early stages of bilingual development, children with limited linguistic abilities in one language may particularly rely on phonological structures from their stronger language – based on more frequent cues (MacWhinney, 2005) – for productions in their non-dominant language, showing delay in speech sound production abilities as compared to monolinguals (as reported in Gildersleeve-Neumann et al., 2008, 2009; and Fabiano-Smith and Goldstein, 2010a, 2010b). Therefore, speech language pathologists may want to take language abilities into account when assessing and providing intervention to young bilinguals.

Interdependence between Phonological Systems

It is important to note that errors predominantly affected language-specific sounds, that is, sounds that are present only in English or Spanish, in particular, English /v, z, ʃ, θ/ and Spanish /β, ɣ, x, r, ɾ/. Accuracy rates with unshared sounds and clusters were indeed significantly lower than those with shared ones, confirming the hypothesis of positive and negative transfer in bilingual phonological development (Goldstein & Bunta, 2012). This means that, as predicted by the Unified Competition Model (MacWhinney, 2005), interaction between the two phonological systems might have allowed children to transfer phonetically similar sounds (for example, stops, nasals, and obstruent + /l/ clusters) from Spanish to English. However, this interaction might have also resulted in negative transfer in the case of language-specific sounds, therefore resulting in production errors. Although it is true that shared sounds tend to be early developing and unshared ones tend to be acquired late – and thus it is impossible to ascertain whether the earlier acquisition of shared sounds is indeed a result of enhanced cue strength or rather lack of markedness – the strong positive cross-linguistic relations that were found on shared items point to shared speech production abilities with common sounds. Accuracy with unshared items, on the other hand, was largely not related between Spanish and English, indicating that these structures probably develop independently in the two languages and thus may take longer to develop.

Therefore, unlike what has been found for vocabulary and grammar (Marchman et al. 2004; Conboy & Thal, 2006), the results of this study document a strong interdependence of bilingual skills in the area of speech sound production. These findings suggest that bilingual phonological learning may rely more on general or shared abilities than lexical and morpho-syntactic development. Since physiological factors have a profound effect on the sounds children produce both at early (Velleman & Vihman, 2002) and later stages of development (Kent, 2000) irrespective of the language they are learning, it is possible that the same speech motor production abilities support – and the same articulatory constraints limit – speech sound performance in two languages. These common speech motor abilities would provide the foundation for the acquisition of two sound systems and allow children to reach similar accuracy levels in their two languages by preschool age.

The interdependence of the children's phonological systems was also reflected in interaction patterns. Thus, we found evidence of Spanish-only sounds being used in English productions and vice versa. In English, spirantization errors appeared to have this origin, as when /g/ was produced as [ɣ] (e.g. [ˈwæɣən] for *wagon*), /h/ was realized as [x] (e.g. [ˈxænd] for *hand*) and [ɾ] was used in English clusters (e.g. [ˈtɾeɪn] for *train*). Similarly, Spanish trills and flaps were at times realized as [ɹ], both in singleton contexts (e.g., [ˈpeɹo] for /ˈpero/, "dog" and [ˈfloɹ] for /ˈfloɾ/, "flower") and in clusters (e.g., [ˈtɹen] for /ˈtɾen/, "train"). The incidence of language-specific consonants or consonant clusters in the other language was low, in line with previous studies that have documented limited transfer between phonological systems (e.g., Fabiano-Smith & Goldstein, 2010b; Goldstein & Bunta, 2012). However, other errors that did not involve language-specific segments could have originated from the interaction between the two languages, and hence, they could have been cross-linguistic in nature. For example, devoicing errors in English, which all involved word-final obstruents and did not decrease over time (as in Gildersleeve-Neumann et al., 2008), could have originated from the absence of a voicing contrast in final position in Spanish, which would turn final /z/, /ɡ/, and /dʒ/ into [s], [k] and [tʃ], respectively (e.g. [ˈnoʊs] for *nose*, [ˈfɹɑk] for *frog*, and [ˈbɹɪtʃ] for *bridge*). Similarly, the persistence of stopping in Spanish with the segments /β/, /ð/ and /ɣ/ may not only have been developmental in nature but could have also been the result of transfer from English /b/, /d/ and /ɡ/. The fact that stopping errors increased over time following a year of consistent exposure to English in preschool lends support to this interpretation.

Conclusion and Clinical Implications

The current investigation is the first longitudinal study that has examined the growth of Spanish-English phonological skills during the preschool years and the links between speech sound production and morpho-syntactic performance. The results revealed that consonant and cluster accuracy improved significantly in both languages over the course of one year, despite children's enrollment in an English-only program. This suggests that production abilities in the home language continue to develop even when children begin formal schooling in English. Growth was more pronounced in English, however, and, unlike in Spanish, it affected all sound classes, confirming previous findings that the acquisition of the majority language occurs relatively fast in this population and may soon outpace development in the home language (Ruiz-Felter et al., 2016).

Another important finding of this study was that a direct measure of language ability (i.e. MLUw) was significantly correlated with accuracy on singletons and clusters, suggesting that children's phonological abilities develop together with skills in other language domains. Thus, children with higher linguistic competence exhibited more advanced phonological skills than those with more limited language abilities. These results confirm previous findings (Cooperson et al., 2013; Goldestein et al., 2010) and suggest either that speech motor abilities are a prerequisite for increasing utterance length (Chiat, 2001; Dromi, 1987), or that morpho-syntactic proficiency drives phonological ability with increased linguistic abilities providing stronger cues to develop and practice articulatory gestures (MacWhinney, 2005).

Interestingly, phonological and morpho-syntactic skills remained related over time in English, but not in Spanish, suggesting that the relationship between phonological and multi-word production may be more complicated than predicted and it may possibly be mediated by the phonological characteristics of the language involved. Clearly, more research is needed to elucidate this issue further.

Speech sound production was not only related to morpho-syntactic performance in the same language but also to speech sound production in the other language. That is, children with advanced phonological skills in Spanish also displayed advanced phonological skills in English. These findings were interpreted as evidence of shared speech motor production abilities (Scarpino, 2011) and interaction between the two phonological systems (Paradis, 2001). In line with the predictions of MacWhinney's (2005) Unified Competition Model, the results showed that, since Spanish and English have many similar consonant properties, consonant cue reliability is strong and phonological skills in English may be bolstered by production abilities in Spanish.

Applied clinically, the findings reported here suggest that young bilinguals' phonological skills should grow in both languages during the preschool years, although more progress in English may typically occur as children begin formal schooling in the majority language. Professionals working with developing Spanish-English bilinguals should make use of the typical performance levels reported in extant research to provide assessment and identify atypical development. In terms of phonological processes, Spanish-English bilingual preschoolers should be expected to display errors that are similar to those produced by Spanish- and English-speaking monolinguals. At the same time, low levels of non-target productions in the form of transfer from the other language and protracted errors with sounds unique to each language may be typical while children master the unshared aspects of their two phonological systems. The findings of this study also suggest that MLUw, a direct measure of language proficiency, is significantly correlated with segmental accuracy, especially at early stages of multi-word production. Thus, it may be prudent to obtain MLUw values for comparison with phonological measures, and different levels of speech sound proficiency should be expected for children with varying language abilities. A related implication is that intervention aimed at bolstering children's overall linguistic proficiency could potentially have a facilitative effect on their phonological development as well. Finally, speech production measures in the other language add insight into the child's overall phonological proficiency. If a child is performing well in Spanish, it should be expected that he/she will perform well in English as well, especially with shared sounds. In turn, intervention in one language that involves shared sounds and structures might aid the acquisition of the same sounds and structures in the other language as well. These clinical implications notwithstanding, some caution needs to be exercised in translating them into firm recommendations for clinical decision making, since we did not carry out a formal clinical intervention study.

Limitations and Future Directions

This work has several limitations that should be considered when interpreting the results. First, the sample in this study was of moderate size, and phonological and language proficiency levels were only assessed at two points in time, separated by one year. Clearly, we were only able to document developmental changes that occurred within this timeframe. Future studies should include larger, homogeneous samples of children from different Spanish-English bilingual communities in the U.S. as well as children speaking other language pairs, and, given the heterogeneous nature of bilingual populations, children with different proficiency levels and onsets of English acquisition. Future studies should examine bilinguals' developing production abilities over a longer period of time, possibly documenting changes from age 2, when home language should still be dominant, to age 3, when exposure to English in preschool begins, to the outset of formal schooling in kindergarten. It is particularly important that future studies be longitudinal because only this methodology can show changes in sound production that occur over time for the same subjects, truly revealing developmental patterns.

Speech production accuracy was found to be correlated with language proficiency in both Spanish and English at Time 1. However, at Time 2, speech and morpho-syntactic production remained related in English but not in Spanish. Although we proposed that this relationship may be mediated by the phonological characteristics of Spanish, further investigations of this issue are needed, especially at later stages of multi-word production. Indeed, it is still unclear whether the language-speech correlation is steady throughout early childhood and the preschool years, or whether it waxes and wanes over time, and at what point speech and language development become decoupled, perhaps due to phonological attainment exhibiting ceiling effects. Moreover, we only examined the link between one measure of language proficiency, MLUw, and phonological skills. Future studies should further explore the intricate relationship between overall linguistic competence and speech sound production, examining how different measures of language ability – including grammatical but also lexical measures – interact and predict speech production accuracy at different points in development. Comparisons between measures of speech production proficiency and those obtained from comprehensive language assessments at different ages are particularly recommended. Related to this issue is the need to explore whether intervention aimed at bolstering children's linguistic proficiency has any facilitative effect on phonological development as well.

Furthermore, in the current study, the strong positive cross-linguistic relations that were found on shared items were interpreted as evidence of shared speech motor production abilities. However, it is important to note that cross-linguistically shared sounds are usually unmarked and constitute developmentally early acquired sounds, while language-specific sounds tend to occur more rarely in the languages of the world and tend to be acquired late (Eckman, 1977; Maddieson, 1984). This makes it difficult to ascertain whether higher accuracy with shared sounds was indeed a result of enhanced cue strength, or whether it arose from markedness considerations, or a combination of the two. In order to assess the real impact of sound similarity on bilingual phonological development, future studies should examine accuracy with shared and unshared sounds taking into consideration the effects of developmental sequence (Goldstein et al., 2003).

Finally, the correlation analyses in the current study found that morpho-syntactic performance and speech production in the other language were related to phonological skills. However, the relative contribution of these and other factors to speech production proficiency still remains unknown. Future studies that use regression analysis are hence needed to investigate which predictor variables account for the greatest amount of variance in the phonological proficiency of young Spanish-English bilinguals. Studies should not only include multiple measures of language ability but also other variables related to language experience and environment, for example, the effects of language use by parents, siblings and teachers. Investigating these issues will lead to a better understanding of the complex factors that affect the development of sound systems in bilingual children.

References

Aguilar-Mediavilla, E. & Serra-Raventos, M. (2006). Phonological profile of Spanish-Catalan children with specific language impairment at age 4: Are there any changes over time? *Folia Phoniatrica et Logopaedica, 58,* 400-414.

American Speech, Language and Hearing Association (2017, March). *Demographic profile of ASHA members providing bilingual services.* Retrieved from http://www.asha.org/uploadedFiles/Demographic-Profile-Bilingual-Spanish-Service-Members.pdf

Anderson, R. (1999). Loss of gender agreement in L1 attrition: Preliminary results. *Bilingual Research Journal, 23,* 319–338.

Astruc, L., Payne, E., Post, B., Vanrell, M, & Prieto, P. (2012). Tonal targets in early child English, Spanish, and Catalan. *Language and Speech, 56*(2) 229 –253.

Baker, C. (2017). *Foundations of bilingual education and bilingualism (6th ed.)* (selected pages). Clevedon: Multilingual Matters.

Barlow, J. (2002). Error patterns in Spanish-English bilingual phonological development. In Scarabela, B. Fish, S. & Do, A. H.-J. (eds.), *BUCLD 26: Proceedings of the Boston University Conference on Language Development* (pp. 60-71). Boston, MA: Cascadilla Press.

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

Branstine, Z. (1991). Stop/spirant alternations in Spanish: On the representation of contrast. *Studies in the Linguistics Sciences, 21*, 1-22.

Brice, A., Carson, C. & O'Brien, J. D. (2008). *Comprehensive Assessment of Spanish Articulation-Phonology (CASA-P).* Unpublished assessment tool. University of South Florida, St. Petersburg: St. Petersburg, FL.

Brice, A. E., Carson, C. K. & O'Brien, J. D. (2009). Spanish-English articulation and phonology of 4- and 5-year-old preschool children. *Communication Disorders Quarterly, 31*, 3–14.

Brown, R. A. (1973). *A first language: The early stages.* Cambridge: Harvard University Press.

Bunta, F., Fabiano-Smith, L., Goldstein, B.A. & Ingram, D. (2009). Phonological whole-word measures in 3-year-old bilingual children and their age-matched monolingual peers. *Clinical Linguistics and Phonetics, 23,* 156–175.

Conboy, B. T. & Thal, D. J. (2006). Ties between the lexicon and grammar: Cross-sectional and longitudinal studies of bilingual toddlers. *Child Development, 77*(3), 712-735.

Cooperson, S.J., Bedore, L.M. & Peña, E.D. (2013). The relationship of phonological skills to language skills in Spanish-English-speaking bilingual children. *Clinical Linguistics and Phonetics, 27,* 371–389.

Core, C.C. & Scarpelli, C.S. (2015). Phonological development in young bilinguals: Clinical implications. *Seminars in Speech and Language, 36,* 100–108.

Dromi, E. (1987). *Early lexical development.* Cambridge: Cambridge University Press.

Eckman, F. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning, 27,* 315–330.

Evans, J. D. (1996). *Straightforward statistics for the behavioral sciences*. Pacific Grove, CA: Brooks/Cole Publishing.

Fabiano, L. & Goldstein, B.A. (2005). Phonological cross-linguistic influence in sequential Spanish–English bilingual children. *Journal of Multilingual Communication Disorders, 3,* 56–63.

Fabiano-Smith, L. & Barlow, J. (2010). Interaction in bilingual phonological acquisition: Evidence from phonetic inventories. *International Journal of Bilingual Education and Bilingualism*, *13,* 1–17.

Fabiano-Smith, L. & Goldstein, B.A. (2010a). Early, middle, and late developing sounds in monolingual and bilingual children: An exploratory study. *American Journal of Speech-Language Pathology, 19*, 66–77.

Fabiano-Smith, L. & Goldstein, B.A. (2010b). Phonological acquisition in bilingual Spanish-English speaking children. *Journal of Speech, Language, and Hearing Research, 53,* 160–178.

Fabiano-Smith, L., Oglivie, T., Maiefsku, O. & Schertz, J. (2015). Acquisition of the stop-spirant alternation in bilingual Mexican Spanish–English speaking children: Theoretical and clinical implications. *Clinical Linguistics and Phonetics, 29*, 1–26.

Flege, J.E. (1981). The phonological basis of foreign accent: A hypothesis. *TESOL Quarterly, 15,* 443–455.

Gildersleeve-Neumann, C.E., Kester, E.S., Davis, B.L & Peña, E.D. (2008). English speech sound development in preschool-aged children from bilingual English–Spanish environments. *Language, Speech and Hearing Services in Schools*, *39*, 314–328.

Gildersleeve-Neumann, C.E., Peña, E.D., Davis, B.L. & Kester, E.S. (2009). Effects on L1 during early acquisition of L2: Speech changes in Spanish at first English contact. *Bilingualism: Language and Cognition, 12,* 259–272.

Goldman, R. & Fristoe, M. (2000). *Goldman–Fristoe Test of Articulation 2*. Circle Pines, MN: American Guidance Service.

Goldstein, B.A. (2001). Transcription of Spanish and Spanish-influenced English. *Communication Disorders Quarterly, 23,* 54–60.

Goldstein, B.A. & Bunta, F. (2012). Positive and negative transfer in the phonological systems of bilingual speakers. *International Journal of Bilingualism, 16*, 388–401.

Goldstein, B.A., Bunta, F., Lange, J., Rodriguez, J. & Burrows, L. (2010). The effects of measures of language experience and language ability on segmental accuracy in bilingual children. *American Journal of Speech-Language Pathology, 19,* 238–247.

Goldstein, B., Fabiano, L., & Iglesias, A. (2003, April). Phonological representation in bilingual Spanish-English speaking children. Poster presented at *The 4th International Symposium on Bilingualism,* Tempe, AZ.

Goldstein, B.A., Fabiano, L. & Iglesias, A. (2004). Spontaneous and imitated productions in Spanish-speaking children. *Language, Speech, and Hearing Services in Schools, 35,* 5–15.

Goldstein, B.A., Fabiano, L. & Washington, P.S. (2005). Phonological skills in predominantly English-speaking, predominantly Spanish-speaking, and bilingual Spanish-English speaking children. *Language, Speech, and Hearing Services in Schools, 36,* 201–218.

Goldstein, B.A. & Iglesias, A. (1996). Phonological patterns in Puerto Rican Spanish-speaking children with phonological disorders. *Journal of Communication Disorders, 29*(5), 367-386.

Goldstein, B.A. & Washington, P.S. (2001). An initial investigation of phonological patterns in typically developing 4-year-old Spanish-English bilingual children. *Language, Speech and Hearing Services in Schools, 32*, 153–164.

Greenlee, M. (1974). Interacting processes in the child's acquisition of stop-liquid clusters. *Papers and Reports on Child Language Development, 7,* 85–100.

Hodson, B. W. (1985). *Assessment of Phonological Processes—Spanish*. San Diego, CA: Los Amigos.

Hodson, B. W. (2004). *Hodson Assessment of Phonological Patterns (3rd ed.*). Austin, TX: Pro-Ed.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(2), 65-70.

Ingram, D. (2002). The measurement of whole-word productions. *Journal of Child Language, 29,* 713–733.

International Phonetic Association (1999). *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.

Kehoe, M. (2011). Relationships between lexical and phonological development: A look at bilingual children – a commentary on Stoel-Gammon's 'Relationships between lexical and phonological development in young children.' *Journal of Child Language, 38,* 75–81.

Kehoe, M. (2015). Lexical-phonological interactions in bilingual children. *First Language, 35*(2) 93-125.

Kent, R. D. (2000). Research on speech motor control and its disorders: A review and prospective. *Journal of Communication Disorders, 33,* 391–427.

Khan, L., & Lewis, N. (2002). *The Khan–Lewis Phonological Analysis (2nd ed.).* Circle Pines, MN: American Guidance Service.

Kirk, C. & Demuth, K. (2005). Asymmetries in the acquisition of word-initial and word-final consonant clusters. *Journal of Child Language, 32,* 709–734.

Kraemer, R. & Fabiano-Smith, L. (2017). Language assessment of Latino English-learning children: A records abstraction study. *Journal of Latinos and Education,* http://dx.doi.org/10.1080/15348431.2016.1257429

Krogstad, J. M. (2016). *Key facts about how the U.S. Hispanic population is changing.* Retrieved from http://www.pewresearch.org/fact-tank/2016/09/08/key-facts-about-how-the-u-s-hispanic-population-is-changing/

Lleó, C. (2006). The acquisition of prosodic word structures in Spanish by monolingual and Spanish-German bilingual children. *Language and Speech*, *49,* 207–231.

Locke, J. L. (1983). *Phonological acquisition and change.* New York: Academic Press.

MacWhinney, B. (2005). A unified model of language acquisition. In J. F. Kroll & A. M. B. De Groot (eds.), *Handbook of bilingualism: Psycholinguistic approaches,* pp. 49-67. Oxford: Oxford University Press.

Maddieson, I. (1984). *Patterns of sounds.* Cambridge: Cambridge University Press.

Marchman, V. A., Martínez-Sussmann, C. & Dale, P. S. (2004). The language-specific nature of grammatical development: Evidence from bilingual language learners. *Developmental Science, 7*(2), 212-24.

Mayr, R., Howells, G. & Lewis, R. (2015). Asymmetries in phonological development: The case of word-final cluster acquisition in Welsh-English bilingual children. *Journal of Child Language, 42,* 146–179.

Mayr, R., Jones, D. & Mennen, I. (2014). Speech learning in bilinguals: Consonant cluster acquisition, In Thomas, E.M. & Mennen, I. (eds.), *Advances in the study of bilingualism* (pp. 3–24). Bristol: Multilingual Matters.

Miccio, A. W. & Hammer, C. S. (2006). *Bilingual Phonological Assessment (BiPA).* Unpublished assessment instrument.

Miller J., & Iglesias A. (2008). *Systematic Analysis of Language Transcripts* (Research Version 2008) [Computer software]. Madison, WI: SALT Software.

Montanari, S., Subrahmanyam, K., Zepeda, M., & Rodriguez, C. (2014). The English phonological skills of Latino Spanish-English dual-language preschoolers living in Los Angeles: Implications for practice. In A.E. Arias, M. Gutiérrez, A. Landa, & F. Ocampo (eds.), *Perspectives in the study of Spanish language variation. Papers in honor of Carmen Silva-Corvalán* (pp. 465-498). Santiago de Compostela, Spain: Universidade de Santiago de Compostela Press.

Munro, S., Ball, M. J., Müller, N., Duckworth, M. & Lyddy, F. (2005). Phonological acquisition in Welsh–English bilingual children. *Journal of Multilingual Communication Disorders*, *3,* 24–49.

National Center for Education Statistics (2017). *Fast facts.* Retrieved from https://nces.ed.gov/fastfacts/display.asp?id=96

Paradis, J. (2001). Do bilingual two-year-olds have separate phonological systems? *International Journal of Bilingualism, 5*, 19-38.

Paradis, J. & Genesee, F. (1996). Syntactic acquisition in bilingual children: Autonomous or interdependent? *Studies in Second Language Acquisition, 18,* 1-25.

Parker, M. D. & Brorson, K. (2005). A comparative study between mean length of utterance in morphemes (MLUm) and mean length of utterance in words (MLUw). *First Language*, *25*(3), 365-376.

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

Prezas, R.F., Hodson, B.W. & Schommer-Aikins, M. (2014). Phonological assessment and analysis of bilingual preschoolers' Spanish and English word productions. *American Journal of Speech-Language Pathology*, *23,* 176–185.

Ruiz-Felter, R., Cooperson, S.J., Bedore, L.M. & Peña, E.D. (2015). Influence of current input–output and age of first exposure on phonological acquisition in early bilingual Spanish–English-speaking kindergarteners. *International Journal of Language and Communication Disorders, 51,* 368–383.

Scarpino, S. E. (2011). *The effects of language environment and oral language ability on phonological production proficiency in bilingual Spanish-English speaking children.* Unpublished doctoral dissertation, Pennsylvania State University.

Shriberg, L. D. & Kwiatkowski, J. (1980). *Natural Process Analysis: A procedure for phonological analysis of continuous speech samples.* New York: Macmillan.

Shriberg, L.D., Austin, D., Lewis, B.A., McSweeny, J.L. & Wilson, D.L. (1997). The Percentage of Consonants Correct (PCC) metric: Extensions and reliability data. *Journal of Speech, Language, and Hearing Research*, *40*, 708–722.\

Stoel-Gammon, C. & Dunn, C. (1985). *Normal and disordered phonology in children*. Baltimore: University Park Press.

United States Census Bureau (2016). *Annual estimates of the resident population by sex, age, race, and Hispanic origin for the United States and States: April 1, 2010 to July 1, 2015*. Retrieved from https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk

Velleman, S. L., & Vihman, M. M. (2007). Phonology in infancy and early childhood: Implications for theories of language learning. In M. C. Pennington (Ed.), *Phonology in context* (pp. 25-50) (Palgrave Advances in Linguistics). Luton: Palgrave Macmillan.

Woodcock, R.W., Muñoz-Sandoval, A., Ruef, M., & Alvarado, C. (2005). *Woodcock-Muñoz Language Survey-Revised*. Itasca, IL: Riverside Publishing

Table 1. Comparison of studies on the phonological skills of Spanish-English bilingual children in the U.S. between 3 and 7 years of age.

Study	*N*	Languages assessed	Participant groups	Ages	Design	Analyses	Findings
Brice *et al.* (2009)	16	S & E	S-E BILS	4 years (*N*=8); 5 years (*N*=8)	Cross-sectional	Articulation tests in Spanish (CASA-P) and English (GFTA-2; KLPA-2): PCC and phonological processes	Significantly greater accuracy on stops and liquids/ glides in E than S; cross-linguistic differences in stopping and fronting.
Bunta *et al.* (2009)	24	S & E	S-E BILS (*N*=8); S MONS (*N*=8); E MONS (*N*=8)	3-4 years	Cross-sectional	Single-word samples (BESA): Phonological whole-word measures (pMLU and proximity) and PCC	MONS > BILS on E measures, but only on PCC for S; cross-linguistic differences in PCC and pMLU.
Cooperson *et al.* (2013)	186	S & E	S-E BILS	5-6 years	Cross-sectional	BESA: phonology, semantics and morpho-syntax subtests; narrative samples: percent grammatical utterances, MLUw, NDW	Correlation between phonology and morphosyntax within both languages, as well as correlation in phonology across languages
Fabiano & Goldstein (2005)	3	S & E	S-E BILS	5;0; 6;2; 7;0	Multiple case study	Single word, conversation and narrative samples	Phonemic and syllabic cross-linguistic effects
Fabiano-Smith & Barlow (2010)	24	S & E	S-E BILS (*N*=8); S MONs (*N*=8); E MONS (*N*=8)	3-4 years	Cross-sectional	Single-word samples (BESA): analysis of complexity of phonetic inventories	No difference in complexity of phonetic inventories between MONS and BILS
Fabiano-Smith & Goldstein (2010a)	24	S & E	S-E BILS (*N*=8); E MONS (*N*=8); S MONS (*N*=8)	3-4 years	Cross-sectional	Single-word samples (BESA): PCC; analysis of early-, middle- and late-developing sounds	MONS > BILS on PCC in S, but not E. Confirmation of EML sounds in E; EML norms for S established
Fabiano-Smith & Goldstein (2010b)	24	S & E	S-E BILS (*N*=8); E MONS (*N*=8); S MONS (*N*=8)	3-4 years	Cross-sectional	Single-word samples (BESA): PCC (including shared & unshared	MONS > BILS on PCC in S, but not E, but limited to specific manner classes;

						sounds); phonetic inventory analysis	BILS and S MONS: shared sounds > unshared sounds
Gildersleeve-Neumann *et al*. (2008)	33	E	S-E BILS: (balanced, *N*=3); S-E BILS (E-dominant, *N*=20); E MONS (*N*=10)	Study onset: 3-4 years	Longitudinal: 8-month period	Single-word samples: analysis of phonetic inventories; PCC; PVC; phonological processes	No significant difference in PCC and PVC over time; no significant between-group difference in accuracy; PVC > PCC
Gildersleeve-Neumann *et al.* (2009)	6	S	S-E BILS	Study onset: 3;5	Longitudinal: 6-month period	Single-word samples: analysis of phonetic inventory; PCC; PVC; phonological processes	No significant change in PCC over time; accuracy on PVC lower at Time 2 than Time 1
Goldstein & Bunta (2012)	30	S & E	S-E BILS (*N*=10); E MONS (*N*=10); S MONS (*N*=10);	6;0	Cross-sectional	Single-word samples (BESA): Phonological whole-word measures; PCC-R; PVC; phonological processes	BILS = MONS in S, except on stops; BILS > MONS in E on Proximity, PVC, PCC-R, and PCC-R for nasals.
Goldstein & Washington (2001)	12	S & E	S-E BILS	4;0-4;11	Cross-sectional	Single-word samples: PCC; PVC; phonological processes	No significant differences across S and E on any measure.
Goldstein *et al.* (2005)	15	S & E	S-E BILS: balanced (*N*=5); S-dominant (*N*=5); E-dominant (*N*=5)	5-6 years	Cross-sectional	Single word samples: PCC-R (including early-, mid- & late-developing sounds); substitution patterns; cross-linguistic effects; syllable structure analysis	No between-group differences on any of the measures.
Goldstein *et al.* (2010)	50	S & E	S-E BILS: 5 groups, based on language use, proficiency and output (each *N*=10)	4;3-7;1	Cross-sectional	Single-word samples (BESA): PCC-R and PVC; narrative sample to calculate MLUw	Segmental accuracy predicted by MLUw and reported language use

Prezas *et al.* (2014)	56	S & E	S-E BILS	4-5 years	Cross-sectional	Single-word productions (HAPP-3 for English; APP-S for Spanish): accuracy (TOMPD score)	Accuracy and deviation patterns affected by age, but not gender. No cross-linguistic differences.
Ruiz-Felter *et al.* (2016)	91	S & E	S-E BILS: balanced (*N*=20); S-dominant (*N*=43); E-dominant (*N*=28)	5-6 years	Cross-sectional	Single-word samples (BESA): PCC-R; PVC; early-, mid- and late-developing sounds	Current input-output accounts for more of the variance than age of first exposure to E.
Scarpino (2011)	199	S & E	S-E BILS	3;0-6;5	Cross-sectional	BiPA: PCC-R; pMLU; PWP; Picture Vocabulary and Story Recall Test (Woodcock *et al*., 2005)	Cross-linguistic correlation in phonological accuracy as well as correlation between lexical and phonological skills

Key: S= Spanish; E= English; BILS = bilinguals; MONS= monolinguals; PCC-R= percentage consonants correct revised; PVC-R= percentage vowels correct revised; APP-S= Assessment of Phonological Processes—Spanish (Hodson, 1985); BESA= Bilingual English-Spanish Assessment (Peña *et al.*, 2014); BiPA = Bilingual phonological assessment (Miccio & Hammer, 2006); CASA-P= Comprehensive Assessment of Spanish Articulation–Phonology (Brice *et al.*, 2008); GFTA-2= Goldman–Fristoe Test of Articulation 2 (Goldman & Fristoe, 2000); HAPP-3= Hodson Assessment of Phonological Patterns—Third Edition (Hodson, 2004); KLPA-2= Khan–Lewis Phonological Analysis–Second Edition (Khan & Lewis, 2002); MLUw = mean length of utterance in words; NDW= number of different words; pMLU = phonological mean length of utterance; PWP= percent whole-word proximity (Ingram, 2002); TOMPD= total occurrences of major phonological deviations.

Table 2. Participants' gender, age at T1, mother's education, home language use and proficiency as reported by parents, MLUw in English and Spanish at Time 1 and Time 2.

Participants	Gender	Age at T1	Mother's Education	Home language (parent-child)	Home language (sibling-child)	Child's Strongest Language	MLUw Eng. T1	MLUw Span. T1	MLUw Eng. T2	MLUw Span. T2
001	M	3;8	Primary	Primarily Spanish	Primarily English	Speaks both equally	2.2	1.87	5.02	2.07
002	F	3;10	Secondary	Spanish only	Spanish only	English	2.09	2.34	3.23	4.38
003	M	3;7	Secondary	Primarily Spanish	English only	English	2.17	2.11	3.28	1.87
004	M	3;4	Secondary	Spanish only	English only	Speaks both equally	1.69	2.05	3.54	2.26
005	F	3;9	Secondary	Primarily Spanish	N/A	Speaks both equally	3.21	2.97	3.84	2.34
006	M	3;9	Primary	Spanish only	Primarily Spanish	Spanish	1.69	1.7	2.37	2.66
007	M	3;6	Primary	Spanish only	Spanish only	Speaks both equally	1.91	2.49	2.07	3.26
008	F	3;4	Secondary	Spanish only	English only	Spanish	1.75	2.59	3.83	2.78
009	F	3;10	Secondary	Spanish only	Primarily Spanish	Spanish	1.55	2.47	2.26	2.42
010	M	3;10	Secondary	Primarily Spanish	English only	English	2.54	2.54	4.35	2.94
011	M	3;7	Primary	Spanish only	Spanish only	Spanish	2.21	2.23	3.65	2.81
012	F	3;2	Primary	Spanish only	Spanish only	Speaks both equally	1.84	2.34	2.67	3.09
013	M	4;0	Primary	Spanish only	N/A	English	2.43	2.43	2.98	3.03
014	F	3;9	Secondary	Spanish only	Primarily Spanish	Spanish	1	3.11	1.65	3.19
015	M	3;11	Secondary	Spanish only	Spanish only	English	1.57	2.33	Not Avail	4.2
016	F	3;7	Secondary	Primarily Spanish	N/A	Spanish	2.87	2.66	4.06	3.78
017	F	3;5	Secondary	Spanish only	Spanish only	Spanish	1.67	2.01	4.12	5.73
018	F	3;5	Secondary	Primarily Spanish	N/A	English	1.41	1.46	1.8	2.14
019	F	3;0	Secondary	Spanish only	N/A	Spanish	1.52	1.61	2.07	3.23
020	F	3;6	Primary	Spanish only	Primarily Spanish	English	2.32	2.19	2.53	2.85
021	M	4;0	Secondary	Spanish only	Primarily Spanish	Spanish	3.32	3.14	2.38	3.81
022	F	3;2	Secondary	Spanish only	N/A	Spanish	2.87	3.51	2.88	3.38
023	M	3;5	Secondary	Spanish only	English only	English	1.38	1.3	3.45	2.32
024	M	3;9	Primary	Spanish only	Primarily Spanish	Spanish	1.9	1.86	2.2	2.76
025	M	3;3	Secondary	Primarily Spanish	N/A	Spanish	2	2.02	2.55	2.6
026	M	3;4	No Schooling	Spanish only	Primarily Spanish	Speaks both equally	1.37	1.39	1.97	2.3
027	F	3;2	Secondary	Spanish only	Primarily Spanish	Spanish	1.2	2.21	2.87	2.66
028	F	3;6	Secondary	Primarily Spanish	Primarily English	Speaks both equally	1.82	1.72	3.13	1.82
029	F	3;8	Some College	Primarily Spanish	English only	English	2.53	2.63	4.02	2.88
030	F	4;2	Secondary	Primarily Spanish	N/A	Spanish	1.98	2.69	2.73	3.21
031	F	3;10	Secondary	Spanish only	Spanish only	Spanish	1.91	3.48	2.82	3.36
032	M	3;8	Secondary	Primarily Spanish	Spanish only	Speaks both equally	2.09	1.55	3.44	2.81
033	F	3;5	Secondary	Spanish only	Primarily Spanish	Spanish	2.36	2.85	4.08	2.63
034	F	4;2	Primary	Spanish only	Primarily English	English	1.88	3.24	3.32	4.78
035	M	4;1	Primary	Spanish only	English only	Spanish	2.96	2.49	5.2	4.82

Table 3. Accuracy measures: English; SDs in parenthesis; *** = significant effect (Holm-Bonferroni adjusted).

	TIME 1	TIME 2	DIFFERENCE	ADJUSTED α-level
SINGLETONS				
PCC-R	71.11 (14.60) *n*= 1633	82.18 (9.36) *n*= 1688	***$F(1,34) = 19.218$, $p < .001$, $\eta^2 = .361$	.0083
PCC-R shared	80.08 (14.35) *n*=1249	89.56 (7.38) *n*= 1279	***$F(1,34) = 20.256$, $p < .001$, $\eta^2 = .373$	.0071
PCC-R unshared	47.77 (15.65) *n*= 384	59.74 (12.65) *n*= 409	***$F(1,34) = 11.654$, $p = .002$, $\eta^2 = .255$	.0167
PCC-R stops	84.96 (15.17) *n*= 578	90.56 (8.70) *n*= 592	***$F(1,34) = 5.867$, $p = .021$, $\eta^2 = .147$	.025
PCC-R fricatives	47.19 (15.30) *n*= 381	59.22 (13.25) *n*= 405	***$F(1,34) = 14.735$, $p = .001$, $\eta^2 = .302$	.01
PCC-R affricates	71.90 (17.04) *n*= 103	80.47 (29.84) *n*= 100	$F(1,34) = 1.452$, $p = .236$, $\eta^2 = .041$	.05
PCC-R nasals	82.45 (17.04) *n*=303	93.96 (6.79) *n*= 313	***$F(1,34) = 12.989$, $p = .001$, $\eta^2 = .276$	.0125
PCC-R approximants	69.07 (21.29) *n*= 268	86.68 (15.23) *n*= 278	***$F(1,34) = 27.344$, $p < .001$, $\eta^2 = .446$	.0056
CLUSTERS				
PClC	46.95 (23.19) *n*= 565	74.37 (15.79) *n*=587	***$F(1,34) = 61.515$, $p < .001$, $\eta^2 = .644$	.005
PClC shared	69.76 (29.09) *n*= 137	92.62 (17.0) *n*= 140	***$F(1,34) = 23.272$, $p < .001$, $\eta^2 = .406$	.0063
PClC unshared	39.97 (23.48) *n*= 428	68.64 (17.95) *n*= 447	***$F(1,34) = 62.2$, $p < .001$, $\eta^2 = .647$	.00455

Table 4. Accuracy measures: Spanish; SDs in parenthesis; *** = significant effect (Holm-Bonferroni adjusted).

	TIME 1	TIME 2	DIFFERENCE	ADJUSTED α-level
SINGLETONS				
PCC-R	76.19 (11.88) *n*= 1847	80.64 (9.18) *n*= 1848	***$F(1,34)= 13.011$, $p= .001$, $\eta^2= .277$	.005.
PCC-R shared	85.93 (12.26) *n*= 1234	89.61 (8.65) *n*= 1274	$F(1,34)= 6.712$, $p= .014$, $\eta^2=.165$	.0063
PCC-R unshared	48.08 (18.37) *n*= 578	53.12 (16.97) *n*= 574	$F(1,34)= 3.7$, $p= .063$, $\eta^2=.098$	.01
PCC-R stops	87.38 (13.51) *n*= 480	90.75 (10.76) *n*= 480	$F(1,34)= 2.918$, $p= .097$, $\eta^2=.079$	.0167
PCC-R fricatives	75.11 (13.45) *n*= 581	76.68 (10.14) *n*= 583	$F(1,34)= 0.5$, $p= .484$, $\eta^2=.014$	.05
PCC-R affricates	97.14 (16.90) *n*= 35	100.0 (0.0) *n*= 35	$F(1,34)= 1.0$, $p= .324$, $\eta^2=.029$	.025
PCC-R nasals	95.51 (9.66) *n*= 243	98.37 (5.77) *n*= 243	$F(1,34)= 4.959$, $p= .033$, $\eta^2=.127$	.0071
PCC-R approximants	80.57 (21.67) *n*= 206	90.38 (13.64) *n*= 208	$F(1,34)= 7.448$, $p= .01$, $\eta^2=.18$	.0056
PCC-R flaps	69.52 (33.45) *n*= 102	78.57 (33.72) *n*= 98	$F(1,34)= 4.141$, $p= .05$, $\eta^2=.109$	.0083
PCC-R trills	24.57 (25.72) *n*= 200	33.81 (36.37) *n*= 201	$F(1,34)= 2.981$, $p= .093$, $\eta^2=.081$	.0125
CLUSTERS				
PClC	49.34 (33.42) *n*= 274	75.71 (22.88) *n*= 280	*** $F(1,34)= 27.291$, $p< .001$, $\eta^2=.445$	.0038
PClC shared	65.95 (41.39) *n*= 136	93.57 (19.50) *n*= 140	*** $F(1,34)= 18.066$, $p< .001$, $\eta^2=.347$	.0042
PClC unshared	32.86 (33.08) *n*= 138	57.86 (39.65) *n*= 140	*** $F(1,34)= 17.50$, $p<.001$, $\eta^2=.340$	.0045

Table 5. Error patterns: English.

	Example	*TIME 1*	*TIME 2*
SINGLETONS			
Correct		1187 (72.7%)	1392 (82.5%)
Total errors		446 (27.3%)	296 (17.5%)
Deletion	/ˈklaʊn/ → [ˈklaʊ]	156 (9.6%)	91 (5.4%)
Stopping	/ˈfɛðɚ/ → [ˈfɛdɚ]	88 (5.4%)	76 (4.5%)
Devoicing	/ˈfɹɑg/ → [ˈfɹɑk]	54 (3.3%)	48 (2.8%)
Gliding	/ˈɹɪŋ/ → [ˈwɪŋ]	28 (1.7%)	11 (<1%)
Other	/kʌmˈpjuɾɚ/→[kʌmˈpjunɚ]	120 (7.3%)	70 (4.1%)
CLUSTERS			
Correct		267 (47.3%)	442 (75.3%)
Total errors		298 (52.7%)	145 (24.7%)
Reduction	/ˈpleɪt/ → [ˈpeɪt]	221 (39.1%)	111 (18.9%)
Simplification	/ˈpɹɛzənt/ → [ˈpwɛzənt]	66 (11.7%)	29 (4.9%)
Deletion	/ˈglæs/ → [ˈæs]	6 (1.1%)	3 (<1%)
Other	/ˈskɹuˈdɹaɪvɚ/ → [ˈskɹuˈstaɪvɚ]	5 (<1%)	2 (<1%)

Table 6. Error patterns: Spanish.

	Example	*TIME 1*	*TIME 2*
SINGLETONS			
Correct		1409 (76.3%)	1494 (80.8%)
Total errors		438 (23.7%)	354 (19.2%)
Deletion	/ˈfloɾ/ → [ˈflo]	170 (9.2%)	59 (3.2%)
Flap/ trill deviation	/aˈros/ → [aˈjos]	140 (7.6%)	134 (7.3%)
Stopping	/aˈɣuxa/ → [aˈguxa]	54 (2.9%)	89 (4.8%)
Spirantization	/ˈdientes/ → [ˈðientes]	9 (<1%)	19 (1.0%)
Other	/aˈɣuxa/ → [aˈɣuɣa]	65 (3.5%)	53 (2.9%)
CLUSTERS			
Correct		134 (48.9%)	210 (75.0%)
Total errors		140 (51.1%)	70 (25.0%)
Reduction	/ˈfɾio/ → [ˈfio]	109 (39.8%)	27 (9.6%)
Simplification	/ˈkɾus/ → [ˈklus]	19 (6.9%)	43 (15.4%)
Deletion	/ˈbloke/ →[ˈoke]	12 (4.4%)	-

Table 7. Correlations: Singleton consonant and cluster accuracy and MLUw by language and time; *** = significant effect (Holm-Bonferroni adjusted).

	Time 1			Time 2		
ENGLISH	*r-value*	*p-value*	*Adjusted α-level*	*r-value*	*p-value*	*Adjusted α-level*
PCC-R	.403	***.016	.05	.460	***.006	.025
PClC	.538	***.001	.025	.433	***.011	.05
SPANISH	*r-value*	*p-value*	*Adjusted α-level*	*r-value*	*p-value*	*Adjusted α-level*
PCC-R	.415	***.013	.05	.157	.368	.05
PClC	.462	***.005	.025	.265	.124	.025

Table 8. Cross-linguistic correlations: Accuracy of shared and unshared singletons and clusters; *** = significant effect (Holm-Bonferroni adjusted).

	Time 1			Time 2		
	r-value	*p-value*	*Adjusted α-level*	*r-value*	*p-value*	*Adjusted α-level*
Singletons						
PCC-R	.732	***<.001	.0125	.431	***.01	.0167
PCC-R shared	.797	***<.001	.0083	.606	***<.001	.0083
PCC-R unshared	-.061	.727	.05	.025	.888	.05
Clusters						
PClC	.770	***<.001	.01	.479	***.004	.0125
PClC shared	.642	***<.001	.0167	.573	***<.001	.01
PClC unshared	.598	***<.001	.025	.277	.107	.025