LSHSS

Research Note

Diagnostic Accuracy of Traditional Measures of Phonological Ability for Bilingual Preschoolers and Kindergarteners

Leah Fabiano-Smith[a] and Katherine Hoffman[a]

Purpose: Bilingual children whose phonological skills are evaluated using measures designed for monolingual English speakers are at risk for misdiagnosis of speech sound disorders (De Lamo White & Jin, 2011).
Method: Forty-four children participated in this study: 15 typically developing monolingual English speakers, 7 monolingual English speakers with phonological disorders, 14 typically developing bilingual Spanish–English speakers, and 8 bilingual children with phonological disorders. Children's single-word speech productions were examined on Percentage Consonants Correct–Revised (Shriberg, Austin, Lewis, McSweeny, & Wilson, 1997a) and accuracy of early-, middle-, and late-developing sounds (Shriberg, 1993) in English. Consonant accuracy in English was compared between monolinguals and bilinguals with and without speech sound disorders. Logistic regression and receiver operating characteristic curves were used to observe diagnostic accuracy of the measures examined.
Results: Percentage Consonants Correct–Revised was found to be a good indicator of phonological ability in both monolingual and bilingual English-speaking children at the age of 5;0. No significant differences were found between language groups on any of the measures examined.
Conclusions: Our results suggest that traditional measures of phonological ability for monolinguals could provide good diagnostic accuracy for bilingual children at the age of 5;0 years. These findings are preliminary, and children younger than 5;0 years should be examined for risk of misdiagnosis.

In 2010, 64% of school-based speech-language pathologists (SLPs) reported working with bilingual students (American Speech-Language-Hearing Association, 2012); however, the literature points to a lack of cultural competence for clinicians attempting to diagnose speech sound disorders in bilingual children (e.g., Kritikos, 2003; Skahan, Watson, & Lof, 2007). This is largely due to the absence of phonological normative data, a paucity of research on bilingual phonological acquisition (Hambly, Wren, McLeod, & Roulstone, 2013), and a subsequent shortage of SLPs trained in evidence-based procedures for the diagnosis of speech sound disorders in bilingual children (American Speech-Language-Hearing Association, 2012). In the absence of norm-referenced standardized tests for bilingual children, 67% of clinicians have reported using informal measures of phonological assessment (Skahan et al., 2007, p. 251). Researchers have also categorized bilingual children's phonological skills into typical and disordered categories based on informal measures of phonological ability developed for monolingual English-speaking children (e.g., Fabiano-Smith & Goldstein, 2010b; Gildersleeve-Neumann, Kester, Davis, & Peña, 2008).

The use of informal measures comes with many challenges, however. The differences found in previous studies between typically developing monolingual and bilingual preschoolers on measures of phonological ability argue that bilingual children develop similarly, but not identically, to their monolingual peers in the realm of phonology (in both English and Spanish; e.g., Fabiano-Smith & Goldstein, 2010a, 2010b; Gildersleeve-Neumann et al., 2008). Bilingual preschoolers are at high risk for misdiagnosis because they demonstrate lower levels of consonant accuracy, albeit within the normal range, on overall measures of accuracy and certain manner classes of sounds as compared to their age-matched monolingual peers (Fabiano-Smith & Goldstein, 2010a, 2010b), as well as high variability in production

[a]The University of Arizona, Tucson

Correspondence to Leah Fabiano-Smith: leahfabianosmith@email.arizona.edu

Editor-in-Chief: Shelley Gray
Editor: Ignatius Nip

Received May 16, 2017
Revision received August 1, 2017
Accepted August 29, 2017
https://doi.org/10.1044/2017_LSHSS-17-0043

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

accuracy (Fabiano-Smith & Goldstein, 2010a). These differences, however, appear to resolve at the age of 5;0 (Goldstein, Fabiano, & Washington, 2005). Because the mean age of referral for children with phonological disorders is 4;3 years (Shriberg & Kwiatkowski, 1994), preschoolers are at high risk for misdiagnosis.

To illustrate how bilingual children are performing on traditional measures of speech sound ability developed for monolinguals, the current study focuses on the use of a common criterion-referenced measure, Percentage Consonants Correct–Revised (PCC-R), as part of a comprehensive evaluation of bilingual preschoolers with suspected speech sound disorders (Shriberg, Austin, Lewis, McSweeny, & Wilson, 1997b). This measure has been found to correlate strongly with the perceptual severity of a phonological disorder (Garrett & Moran, 1992; Miccio, 2002; Shriberg & Kwiatkowski, 1982) and also correlates positively with speech intelligibility, an aspect of speech sound production often evaluated in preschoolers suspected of having a phonological disorder (Skahan et al., 2007). Percent Consonants Correct (PCC) was first described in a series of articles that outlined the Speech Disorders Classification System (Shriberg, 1993; Shriberg & Kwiatkowski, 1982). The authors then modified the measure to include substitution and omission errors but exclude distortion errors from analysis, renaming the updated measure Percent Consonants Correct–Revised. To calculate PCC-R, the total number of correct consonants in a speech sample is divided by the total number of opportunities for consonant production in that sample and multiplied by 100. The resulting percentage reflects how accurately the child produced the consonant phonemes of his or her language. Shriberg et al. (1997b) recommended using PCC-R to assess severity for 3- to 6-year-old monolingual English-speaking children with speech delays. In the interest of using the most current measure, we elected to examine PCC-R instead of the original PCC measure.

PCC-R has been used in research studies examining the phonological skills of bilingual, as well as monolingual, children (see Table 1). Gildersleeve-Neumann and colleagues (2008) examined the English skills of bilingual Spanish–English–speaking preschoolers using PCC-R and found significantly lower accuracy of speech sound production in bilinguals when compared to their typically developing monolingual English-speaking peers. Specifically, mean PCC-R in English was 77.81% for the monolingual English group, 70.18% for the bilingual but predominantly English-speaking group, and 53.94% for the English productions of the balanced bilingual group. Subsequently, two studies were conducted by Fabiano-Smith and Goldstein (2010a, 2010b) examining eight bilingual Spanish–English–speaking 3-year-olds on measures of consonant accuracy in both their English and Spanish productions. Fabiano-Smith and Goldstein (2010b) found that the English productions of bilinguals were lower on measures of PCC-R than those of their monolingual peers, but not significantly lower (84.10% for monolinguals vs. 72.31% for bilinguals). Mean consonant accuracy for the Spanish productions of bilingual children were, however, significantly lower than their monolingual Spanish-speaking peers (75.58% for monolinguals vs. 65.77% for bilinguals).

In a follow-up study, Fabiano-Smith and Goldstein (2010a) examined the same group of children and categorized consonant accuracy into early-, middle-, and late (EML)-developing phoneme categories in both English and Spanish (see Table 2). Recall that these are the same monolingual English-speaking children in Fabiano-Smith and Goldstein (2010b) who exhibited overall PCC-R scores that were not significantly higher than that of the English productions of their bilingual peers; however, when their speech sounds were analyzed by EML-developing phoneme categories, monolingual children were significantly more accurate on early-developing consonant phonemes as compared to their bilingual counterparts. This was not the case for middle- or late-developing consonant phonemes. Although monolingual Spanish-speaking children and the Spanish productions of bilingual children differed significantly on overall PCC-R in Fabiano-Smith and Goldstein (2010b), the two groups did not differ on accuracy of EML-developing categories of phonemes.

What we can conclude from these studies examining the speech of typically developing bilingual preschoolers is that (a) monolinguals and bilinguals might differ on measures such as PCC-R in one language but not in the other; (b) monolinguals and bilinguals might not differ on gross measures of speech sound accuracy but might differ on discrete measures of accuracy (and vice versa); and (c) across bilingual children, speech sound production is extremely variable. These differences place bilingual children in danger of misdiagnosis when clinicians are trained for judgment of speech sound accuracy based on a monolingual model. Interestingly, when we look beyond the preschool years, we see that the phonological skills of bilingual children begin to normalize at the age of approximately 5 years (Goldstein et al., 2005), which is what we also observe in the literature on monolingual English-speaking children (e.g., Kamhi, 2006). By the time bilingual children enter kindergarten, it appears that, generally speaking, they hit a proverbial ceiling in terms of consonant accuracy. Based on the findings of Goldstein et al. (2005), it is assumed that 5-year-old bilingual children's mean English PCC-R would be similar to Shriberg and Austin's (1997) normative data of monolingual English-speaking children; however, further data are needed to support or challenge this assumption. What is of particular interest to clinicians is that younger bilingual children remain in danger of being misdiagnosed with a speech sound disorder due to (a) the absence of standardized assessments of phonology for bilingual children younger than 4;0 years, (b) their lower performance on the informal measures used in place of those standardized measures when compared to monolingual criteria during the preschool years (e.g., between 3;0 and 6;0 years), and (c) varying degrees of bilingualism across children. How do we know the difference between a speech sound disorder and a child who is still in the process of organizing two sets of speech sounds? We do not yet have empirically derived

Table 1. Summary of previous work examining PCC-R and accuracy of early-, middle-, and late-developing sounds in the English of bilingual Spanish–English–speaking preschoolers.

Publication	*N* (total subjects, bilingual and monolingual)	Chronological age (years;months)	Findings
Goldstein, Bunta, Lange, Rodríguez, & Burrows (2010)	50	4;3–7;1	Average English PCC-R scores were approximately 92% for bilingual children with differing amounts of language exposure and use.
Goldstein, Fabiano, & Washington (2005)	15	5;0–5;5 (x = 5;2)	No significant difference between monolinguals (96.54%) and bilinguals (94.81%) on PCC in English.
Goldstein & Washington (2001)	12	4;0–4;11 (x = 4;7)	English PCC for place of articulation was 94%, with place and manner approaching mastery. No comparison to monolingual English speakers.
Gildersleeve-Neumann et al. (2008)	29	3;1–3;10 (x = 3;6)	Monolingual English speakers produced significantly higher consonant accuracy (PCC = 77%–78%) than predominantly English speakers (PCC = 70%–71%), who demonstrated significantly higher consonant accuracy than balanced bilinguals (PCC = 53%–55%).
Bunta, Fabiano-Smith, Goldstein, & Ingram (2009)	8	3;0–3;11 (x = 3;3)	Monolingual English speakers demonstrated significantly higher PCC (84.88%) than bilingual children (72.85%).
Fabiano-Smith & Goldstein (2010a)	24	3;0–4;0	Accuracy of English early-developing sounds was 93.3%, accuracy of middle-developing sounds was 86.53%, and accuracy of late-developing sounds reached 74.1%. The English productions of monolingual English-speaking children and bilingual children differed significantly on accuracy of early-developing sounds but did not differ in English on middle- or late-developing sounds.
Fabiano-Smith & Goldstein (2010b)	24	3;0–4;0	Overall PCC in English for monolinguals was 84.1%, and overall PCC for bilinguals reached 72.31% (no significant difference).
Burrows & Goldstein (2010)	24	3;1–5;2	PCC was higher in monolingual English speakers with phonological disorders (70.59%), but not significantly higher than bilingual children with speech sound disorders in English (60.55%).

Note. PCC-R = Percentage Consonants Correct–Revised; PCC = Percentage Consonants Correct.

Table 2. Early-, middle-, and late-developing sounds in English and Spanish.

Language	Early	Middle	Late
English[a] (Shriberg, 1993)	/m, b, j, n, w, d, p, h/	/t, ŋ, k, g, f, v, ʧ, ʤ/	/ʃ, ð, s, z, Ɵ, l, ɹ/
Spanish[b] (Fabiano-Smith & Goldstein, 2010a)	/ɲ, t, m, n, k, x/	/s, f, p, ʧ, β, ɣ/	/l, ð, r, ɾ/

[a]The sound /ʒ/ was excluded from analysis by Shriberg (1993). [b]The spirants /β, ɣ, ð/ were analyzed in place of the voiceless stops because they are more likely to be the phoneme, rather than the allophone, in the stop–spirant alternation (Barlow, 2003b).

criteria that would indicate if a bilingual preschooler's speech is typical or disordered, in either English or Spanish, for measures such as PCC-R. Some exploratory data for English are presented here that illustrate the need for normative data on bilingual children and suggest sensitive and specific cutoff scores for traditional measures of consonant accuracy.

Purpose

The purpose of this study was to observe the diagnostic accuracy of PCC-R and accuracy of EML-developing sounds for bilingual preschoolers in English. Goldstein et al. (2005) found that 5-year-old bilingual children performed similarly to their monolingual peers on measures of phonological ability; however, 3-year-olds examined in Fabiano-Smith and Goldstein (2010a, 2010b) demonstrated differences in performance on these measures, with typically developing monolinguals outperforming bilingual children in both languages. These differences place bilingual children at risk for misdiagnosis. To address this issue, the following research question was posed: Are bilingual preschoolers at risk for misdiagnosis of speech sound disorders in English when traditional measures of phonological ability developed for, and validated on, monolingual English speakers are used as part of a diagnostic battery?

Method

Forty-four children participated in this study, ages 3;3–6;6: 15 typically developing monolingual English speakers (mean age = 60.73 months), seven monolingual English speakers with phonological disorders (mean age = 53.14 months), 14 typically developing bilingual Spanish–English speakers (mean age = 65.29 months), and eight bilingual children with phonological disorders (mean age = 58 months). Demographic information on all participants can be found in Table 3. Children were administered the Kaufman Assessment Battery for Children–Nonverbal Scale (Kaufman & Kaufman, 1983), and all demonstrated cognitive abilities within the typical range. Age and nonverbal IQ score were included in our logistic regression model and did not have a significant effect on typical versus disordered categorization. All children were recorded in the border region of Tucson, Arizona, at a Title I preschool program. Because the effects of poverty do not appear to negatively impact phonological acquisition (e.g., Oller, Eilers, Basinger, Steffens, & Urbano, 1995), children from differing socioeconomic levels were all included in the study. Mother's level of education is coded in Table 3 as follows: 1 = *less than high school*; 2 = *high school diploma*; 3 = *post–high school certification*; 4 = *some college*; 5 = *bachelor's degree*; 6 = *master's degree or higher*.

Parent report was used to obtain demographic information on each child, such as chronological age, birth and developmental history, percent language input and output in each language, and phonological status (i.e., any history of speech problems and parental concern). Children with a variety of language exposure levels were included in the study (a) to reflect the heterogeneous nature of the bilingual community and (b) because the children, as a group, demonstrated balanced phonological skills in both languages, indicating balanced input in both English and Spanish. Mean percent English exposure for typically developing bilingual children was 37.7% (SD = 13.9), and mean percent exposure in Spanish was 62% (SD = 14.2). Bilingual children with speech sound disorders had a mean exposure of 47% (SD = 20) in English and 53% (SD = 19.9) in Spanish. For some children, we were unable to reach their parents or caregivers to complete the parent interview portion of the study. In these cases, we classified children into bilingual categories based on teacher report, clinician report, and if the child was able to produce speech samples in both English and Spanish. Children who had exposure to Spanish but spoke only English were included in the monolingual categories. Educational instruction in the public schools is limited to English in the state of Arizona (Arizona Proposition 203, 2000). Bilingual children with older siblings who speak English receive more English exposure than children who are the oldest child in a Spanish-speaking household, and some children simply refuse to speak English, even if it is the primary language of the home. Therefore, there are many children who are growing up in a bilingual environment but, without enough Spanish exposure and use, are speakers of English only.

Children with disorders were identified using a converging concern approach: A parent or teacher and an SLP were required to demonstrate concern regarding the child's communication skills (Restrepo, 1998). Both teacher and parent reports have been found to be significantly correlated with language ability in bilingual children (Bedore, Peña, Joyner, & Macken, 2011). The combination of a parent report of a child's speech and language skills and the number of total errors on a speech or language probe has been found to significantly predict the presence of a

Table 3. Demographic information for study participants.

Child ID	Age (years; months)	Gender	% Spanish input	% Spanish output	% English input	% English output	Socioeconomic status[a]	Goldman-Fristoe Test of Articulation score	Bilingual English–Spanish Assessment score	Total number of errors
TM01[b]	5;8	F	20.00%	20.00%	80.00%	80.00%	3	111		2
TM02	5;8	F	0.00%	0.00%	100.00%	100.00%	3	109		6
TM03	5;7	M	0.00%	0.00%	100.00%	100.00%	2	113		4
TM04	5;11	F	0.00%	0.00%	100.00%	100.00%	5	107		9
TM05	5;11	M	0.00%	0.00%	100.00%	100.00%	4	103		10
TM06[b]	6;2	F	29.00%	29.00%	71.00%	71.00%	2	110		4
TM07	5;8	F	Unavailable	Unavailable	Unavailable	Unavailable	Unavailable	102		16
TM08	5;10	F	25.00%	25.00%	75.00%	75.00%	5	105		11
TM09	4;11	M	Unavailable	Unavailable	Unavailable	Unavailable	Unavailable	113		10
TM10	3;7	M	0.00%	0.00%	100.00%	100.00%	6	117		17
TM11[b]	5;5	M	26.00%	26.00%	73.00%	73.00%	2	110		16
TM12	4;11	F	0.00%	0.00%	100.00%	100.00%	4	110		14
TM13	4;9	F	Unavailable	Unavailable	Unavailable	Unavailable	Unavailable	117		8
TM14	5;6	M	Unavailable	Unavailable	Unavailable	Unavailable	Unavailable	103		9
TM15	4;2	M	0.00%	0.00%	100.00%	100.00%	5	113		7
TM16	3;9	M	0.00%	0.00%	100.00%	100.00%	5	102		7
TB01	5;8	F	34.60%	34.60%	65.40%	65.40%	3		120	0
TB02	5;7	F	56.00%	56.00%	43.00%	43.00%	4		100	6
TB03	5;7	F	56.00%	56.00%	43.00%	43.00%	5		100	25
TB04	5;4	M	28.50%	28.50%	71.40%	71.40%	4		100	20
TB05	5;8	M	Unavailable	Unavailable	Unavailable	Unavailable	Unavailable		115	9
TB06	5;9	F	48.00%	48.00%	52.00%	52.00%	3		115	3
TB07	5;5	M	Unavailable	Unavailable	Unavailable	Unavailable	Unavailable		110	9
TB08	6;3	F	28.50%	28.50%	71.40%	71.40%	2		100	20
TB09	5;5	F	19.00%	19.00%	81.00%	81.00%	2		115	6
TB10	6;1	F	Unavailable	Unavailable	Unavailable	Unavailable	Unavailable		90	16

(*table continues*)

Table 3. *(Continued).*

Child ID	Age (years; months)	Gender	% Spanish input	% Spanish output	% English input	% English output	Socioeconomic status[a]	Goldman-Fristoe Test of Articulation score	Bilingual English–Spanish Assessment score	Total number of errors
TB11	5;0	F	Unavailable	Unavailable	Unavailable	Unavailable	Unavailable		110	14
TB12	4;7	M	43.00%	43.00%	57.00%	57.00%	5		90	15
TB13	5;0	M	18.75%	18.75%	81.25%	81.25%	4		105	11
TB14	5;1	F	52.00%	52.00%	48.00%	48.00%	2		90	29
TB15	5;6	F	30.65%	30.65%	69.35%	69.35%	2		110	7
DM01[b]	3;3	M	25.00%	25.00%	75.00%	75.00%	5	94		74
DM02	6;6	M	0.00%	0.00%	100.00%	100.00%	1	77		29
DM03[b]	5;0	M	0.00%	0.00%	100.00%	100.00%	4	97		21
DM04	4;4	M	32.00%	32.00%	68.00%	68.00%	5	65		80
DM05	4;4	F	33.00%	33.00%	67.00%	67.00%	1	86		61
DM06	3;10	F	0.00%	0.00%	100.00%	100.00%	4	107		33
DB01	5;5	M	29.60%	29.60%	73.00%	73.00%	1		80	48
DB02	4;0	F	71.40%	71.40%	28.50%	28.50%	2		80	54
DB03	4;6	M	36.00%	36.00%	63.00%	63.00%	5		70	35
DB04	4;9	F	Unavailable	Unavailable	Unavailable	Unavailable	Unavailable		80	47
DB05	4;4	M	Unavailable	Unavailable	Unavailable	Unavailable	Unavailable		105	14
DB06	4;8	F	42.00%	42.00%	58.00%	58.00%	2		110	25
DB07	5;8	M	29.60%	29.60%	73.00%	73.00%	3		90	23
DB08[c]	5;4	F	Unavailable	Unavailable	Unavailable	Unavailable	Unavailable		105	2

Note. F = female; M = male; TM = typically developing monolingual; TB = typically developing bilingual; DM = monolingual child with speech sound disorder; DB = bilingual child with speech sound disorder.

[a]Socioeconomic status levels (mother's level of education): 1 = *less than high school*; 2 = *high school diploma*; 3 = *post–high school certification*; 4 = *some college*; 5 = *bachelor's degree*; 6 = *master's degree.* [b]Functionally monolingual, even with Spanish exposure. [c]Child originally classified as "typical"; however, the statistician identified her as an outlier whose scores were patterning with the disordered group (e.g., PCC-R = 60%). Thus, she was recategorized into the group with disorders.

communication impairment over measures derived from standardized tests (Restrepo, 1998). In order to observe children's performance on standardized tests, monolingual children were administered the Goldman-Fristoe Test of Articulation (GFTA; Goldman & Fristoe, 1986), and if older than 4;0 years, bilingual children were administered the English and Spanish phonology subtests of the Bilingual English–Spanish Assessment (BESA; Peña, Gutierrez-Clellen, Iglesias, Goldstein, & Bedore, 2014) to validate concern. Both the GFTA and the BESA are single-word articulation tests that are designed and scored in a similar fashion. A Mann–Whitney *U* test confirmed that monolingual English-speaking children with speech sound disorders exhibited a significantly higher number of total errors than their typically developing peers on the GFTA ($z = -3.54$, $p = .00$), and bilingual children with speech sound disorders exhibited a significantly higher number of total errors on the English BESA than our typically developing bilingual children ($z = -2.32$, $p = .02$); however, some children's standardized test scores did not confirm parent concern. It is well documented that children with speech sound disorders can demonstrate higher accuracy in single-word production than in conversation (e.g., Miller, Heise, & Lichten, 1951), resulting in higher scores on a standardized test, but low intelligibility of speech in conversation. Therefore, we further analyzed the children's speech (a) by age and (b) by speaker group (monolingual and bilingual) in order to determine that children were correctly classified into typical and disordered categories. Percent occurrence of phonological error patterns, by age, are found in Table 4.

The Assessment of English Phonology (Barlow, 2003a; see description below) was used to elicit speech in single words, and those productions were transcribed and analyzed for common phonological error patterns using the Logical International Phonetics Program (LIPP; Oller & Delgado, 2000). Mann–Whitney *U* tests were used to compare percent occurrence of common phonological error patterns between (a) typical monolingual English speakers and bilingual speakers with speech sound disorders and (b) typically developing bilingual speakers and bilingual speakers with speech sound disorders in English. Monolingual English-speaking children with speech sound disorders ($n = 6$) exhibited a significantly higher percent occurrence of cluster deletion ($z = -2.36$, $p = .018$), gliding ($z = -2.99$, $p = .003$), backing ($z = -3.09$, $p = .002$), and fronting ($z = -3.31$, $p = .001$), with stopping approaching significance ($z = -1.77$, $p = .077$), as compared to the children in the typically developing category ($n = 16$). Bilingual children with speech sound disorders ($n = 7$) exhibited a significantly higher percent occurrence, as compared to typically developing bilingual children ($n = 15$), of stopping ($z = -2.08$, $p = .037$) and fronting ($z = -2.57$, $p = .010$), with backing approaching significance ($z = -1.83$, $p = .66$). Overall, it appears that both monolingual and bilingual children with speech sound disorders may perform in the typical range on single-word standardized assessments of articulation; however, their phonological abilities are not fully captured by standardized test scores alone; their intelligibility in connected speech (the reason for parent, teacher, and clinician concern), the number of errors exhibited on an articulation test (rather than their standard score), and percent occurrence of phonological error patterns appeared to more accurately categorize children into typical and disordered speech ability categories.

Data Collection

Picture-naming tasks from the Assessment of English Phonology (Barlow, 2003a) was used to elicit single-word samples in English for all children, and data from these probes were used to calculate our dependent measures of PCC-R and accuracy of EML-developing sounds. The Assessment of English Phonology single-word assessment tool includes a total of 125 words. It elicits approximately 25–75 words for children of this age (older children may produce more words, and younger children may fatigue sooner and produce fewer). Children with speech sound disorders in this study produced an average of 65 words, and children in the typically developing group produced an average of 71 words. Target items reflect the type and frequency of sounds, syllable types, and clusters in English. Children are shown a PowerPoint presentation on a touchscreen tablet consisting of child-friendly pictures depicting culturally sensitive, common items. Children labeled each picture after the research assistant asked, "What is this?" Children's speech was recorded using a Zoom H4 Handy Recorder (Zoom, Inc.) with integrated condenser microphone (24-bit sound card).

Analyses

Phonetic Transcription

Bilingual Spanish–English–speaking undergraduate and graduate students trained in narrow transcription of the International Phonetic Alphabet transcribed all recordings using the LIPP (Oller & Delgado, 2000). Transcription was performed post–data collection in the laboratory setting. Dialectal differences of Mexican Spanish–influenced English and Southwestern American English were anticipated and not counted as errors (e.g., [ʃ] for /ʧ/ in Spanish-influenced English). Interjudge reliability analyses were completed on 100% of all samples included in the study. Mean interjudge reliability for our database reached 98.48%. Transcription decisions were made between a set of two transcribers and the first author of the study, an English–Spanish speaker.

PCC-R

PCC-R was calculated after Shriberg (1993) for monolingual English-speaking children and the English productions of bilingual children using LIPP (Oller & Delgado, 2000).

Accuracy of EML-Developing Sounds

Accuracy of EML-developing sounds was calculated after Shriberg (1993).

Table 4. Mean percent occurrence of phonological patterns by age and speaker group.

Language group	*n*	Weak syllable deletion	Cluster deletion	Cluster reduction	Initial consonant deletion	Final consonant deletion	Gliding	Stopping	Backing	Fronting	Final consonant devoicing
Typical monolingual 3-year-olds	2	0	0.00	18.06	0.00	9.39	29.43	11.15	7.49	3.64	37.50
Monolingual 3-year-olds with disorders	2	0	0.00	33.33	0.00	6.53	15.45	12.35	6.97	4.94	26.98
Typical bilingual 3-year-olds	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Bilingual 3-year-olds with disorders	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Typical monolingual 4-year-olds	4	0	0.00	3.85	0.00	3.27	2.58	8.23	3.57	1.73	32.29
Monolingual 4-year-olds with disorders	2	0	4.06	54.79	0.00	11.91	10.66	34.12	7.66	7.89	13.39
Typical bilingual 4-year-olds[a]	1	0	0.00	3.77	2.33	10.81	53.57	1.20	8.42	1.49	70.00
Bilingual 4-year-olds with disorders	5	0	3.36	12.78	2.12	7.21	14.80	15.26	6.34	6.14	29.40
Typical monolingual 5-year-olds	9	0	0.00	6.07	0.00	2.69	0.71	6.53	2.27	1.88	8.82
Monolingual 5-year-olds with disorders[a]	1	0	0.00	9.62	0.00	2.70	34.48	11.11	8.82	4.90	0.00
Typical bilingual 5-year-olds	12	0	0.48	9.28	0.54	3.14	1.70	8.77	2.79	2.37	19.69
Bilingual 5-year-olds with disorders	2	0	1.17	19.87	0.00	5.63	10.34	17.83	6.35	5.02	33.33
Typical monolingual 6-year-olds[a]	1	0	0.00	0.00	0.00	5.41	0.00	18.31	4.55	2.53	14.29
Monolingual 6-year-olds with disorders[a]	1	0	0.00	3.92	0.00	2.70	27.59	5.00	6.37	6.86	0.00
Typical bilingual 6-year-olds	2	0	0.00	7.27	0.00	1.31	1.56	9.37	3.23	1.48	14.58
Bilingual 6-year-olds with disorders	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

[a]Individual data only due to one subject in this age group.

Table 5. Means and standard deviations for Percent Consonants Correct–Revised (PCC-R) in English by group.

Measure	Typical monolingual English speakers (*n* = 15) *M* (*SD*)	Typical bilingual children (*n* = 14) *M* (*SD*)	Monolingual English speakers with speech sound disorders (*n* = 7) *M* (*SD*)	Bilingual children with speech sound disorders (*n* = 8) *M* (*SD*)
PCC-R English	90.68 (5.56)	88.17 (7.01)	71.26 (11.25)	66.72 (14.00)
Early-developing sounds–English	95.15 (3.08)	93.99 (7.12)	77.30 (7.88)	76.57 (14.78)
Middle-developing sounds–English	91.00 (5.15)	90.85 (5.99)	72.30 (18.26)	65.82 (15.96)
Late-developing sounds–English	87.32 (7.48)	87.13 (11.23)	55.17 (15.21)	56.48 (14.52)

Statistical Analyses

Children with typical and disordered phonological skills were compared using simple logistic regression and receiver operating characteristic (ROC) curves in order to observe the diagnostic accuracy of PCC-R and accuracy of EML-developing sounds for the English productions of monolingual and bilingual children. Area under the curve (AUC) values are interpreted as follows: A value of 1 indicates a perfect separation between typical and disordered children on a given measure; values less than 1 indicate overlap between typical and disordered phonological skills. Cutoff scores for PCC-R that differentiate children with and without disorders were derived using logistic regression. Support vector machine (SVM) curves were derived from a multivariate regression analysis due to high correlations between PCC-R and accuracy of EML-developing sounds. SVMs minimize interclass correlations and maximize binary classifications between two sets of data (Chandaka, Chatterjee, & Munshi, 2009), providing a more nuanced ROC curve. Hanley and McNeil (1982) developed a method for analyzing the difference between AUC values when comparing two ROC curves (online calculator can be found at http://vassarstats.net/roc_comp.html), and this method was used to compare monolingual and bilingual children

Figure 1. Boxplots representing mean accuracy on Percent Consonants Correct–Revised (PCC-R) and percent accuracy of early-, middle-, and late-developing sounds for monolingual English-speaking children.

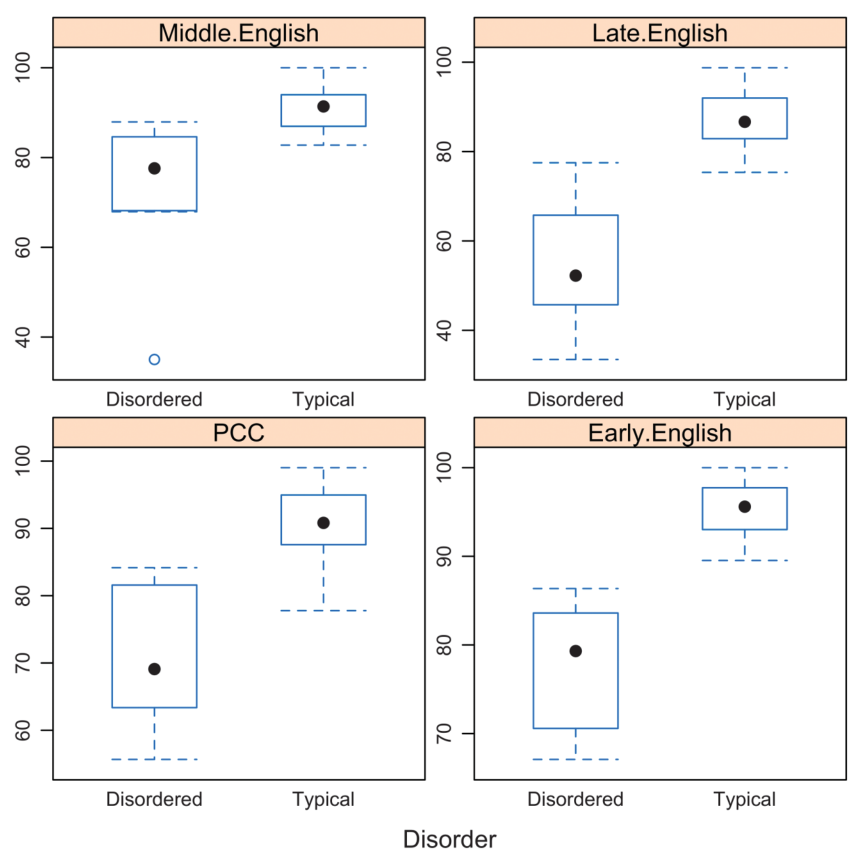
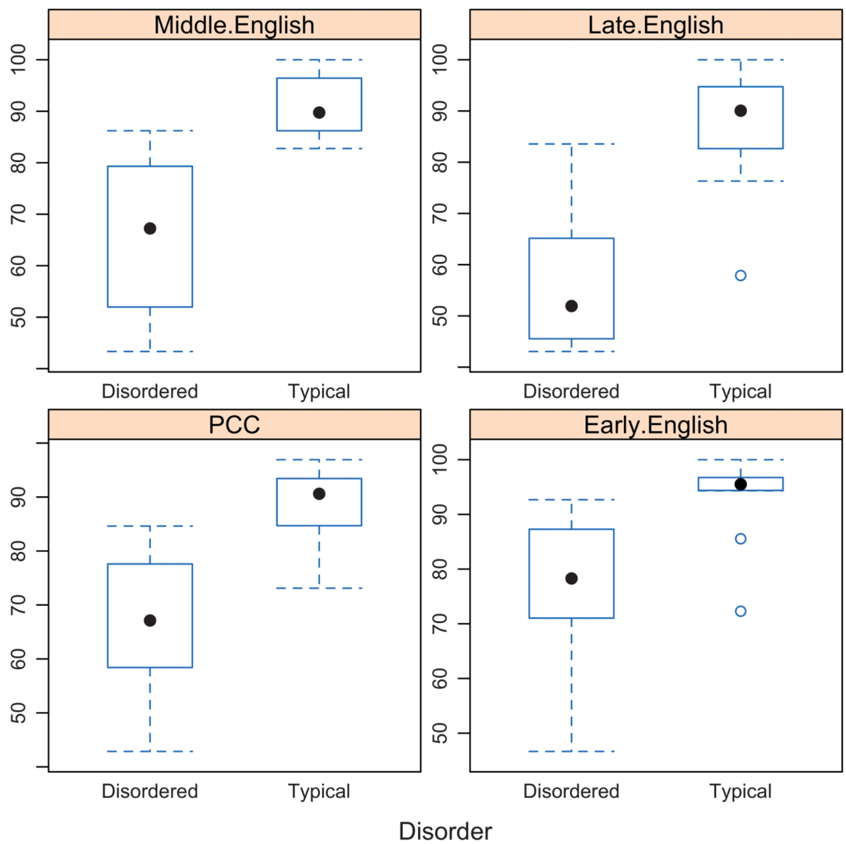


Figure 2. Boxplots representing mean accuracy on Percent Consonants Correct–Revised (PCC-R) and percent accuracy of early-, middle-, and late-developing sounds for the English productions of bilingual children.



on each measure. This was done to determine if these measures were equally accurate in identifying monolingual and bilingual children with speech sound disorders.

Results

Means and standard deviations for all variables across language ability groups can be found in Table 5. Boxplots illustrating mean accuracy for PCC-R and accuracy of EML-developing sounds can be found in Figures 1 and 2.

Simple logistic regression models showed that, for monolingual English speakers, PCC-R was accurate in identifying speech sound disorder with an area under ROC curve of greater than .97 (see Figure 3). This measure was also accurate for the English of bilingual children, as the area under the ROC curve reached .95 (see Figure 4). The difference between the diagnostic accuracy of PCC-R for monolinguals and bilinguals was not significant ($z = 0.27$, $p = .78$). The cutoff score that reached maximum sensitivity and specificity for PCC-R was 89.18% for monolingual English speakers and 84.62% for the English productions of bilingual speakers.

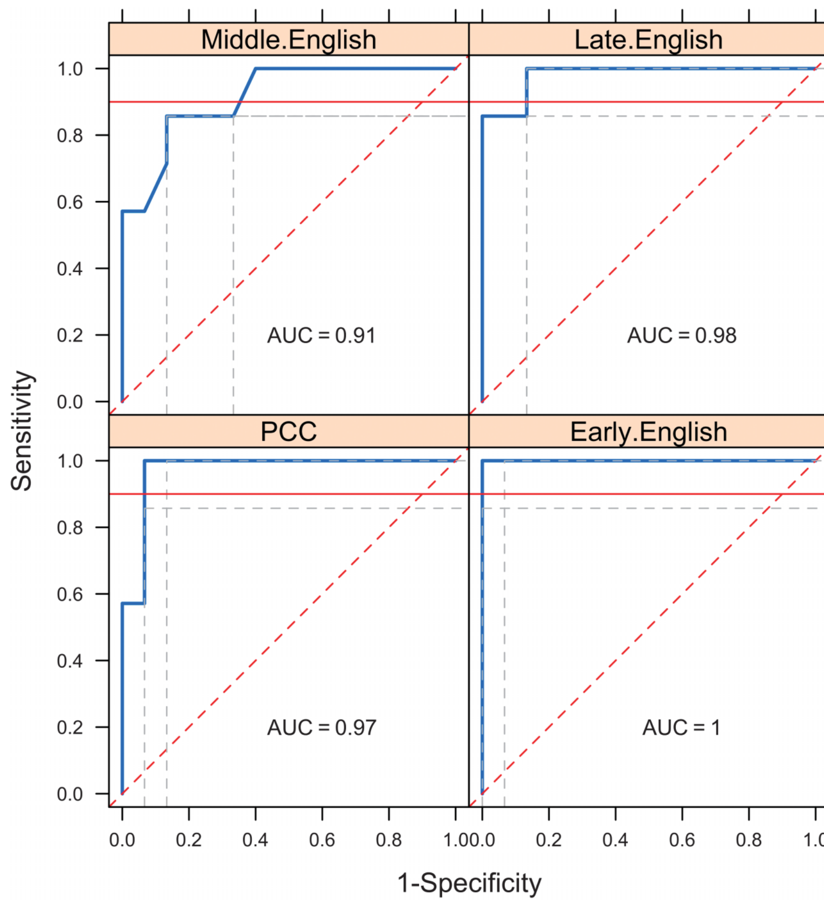
Cutoff scores for accuracy of EML-developing sounds for both groups can be found in Table 6. For the analysis of accuracy of EML-developing sounds, monolingual English speakers displayed AUC values of 1 for accuracy of early-developing sounds, .91 for accuracy of middle-developing sounds, and .98 for accuracy of late-developing sounds (see Figure 3). The analysis of the English productions of bilingual children resulted in AUC values of .93 for early-developing sounds, .96 for middle-developing sounds, and .96 for late-developing sounds (see Figure 4). Comparisons between monolinguals and bilinguals on PCC-R and accuracy of early-developing sounds ($z = 1.04$, $p = .29$), middle-developing sounds ($z = -0.52$, $p = .59$), and late-developing sounds ($z = 0$, $p = 1$) yielded no significant differences.

Due to high correlations among measures, a multivariate analysis was subsequently performed. A linear SVM examined overall PCC-R, and ROC curves were derived (see Figures 5 and 6). The productions of monolingual English-speaking children yielded an AUC of 1, and the English productions of bilingual children yielded an AUC of .94. The difference in AUC values between groups was not significant ($z = 0.96$, $p = .33$).

Discussion

Traditional measures of phonological ability developed for and validated on monolingual English-speaking children appear to have good diagnostic sensitivity and

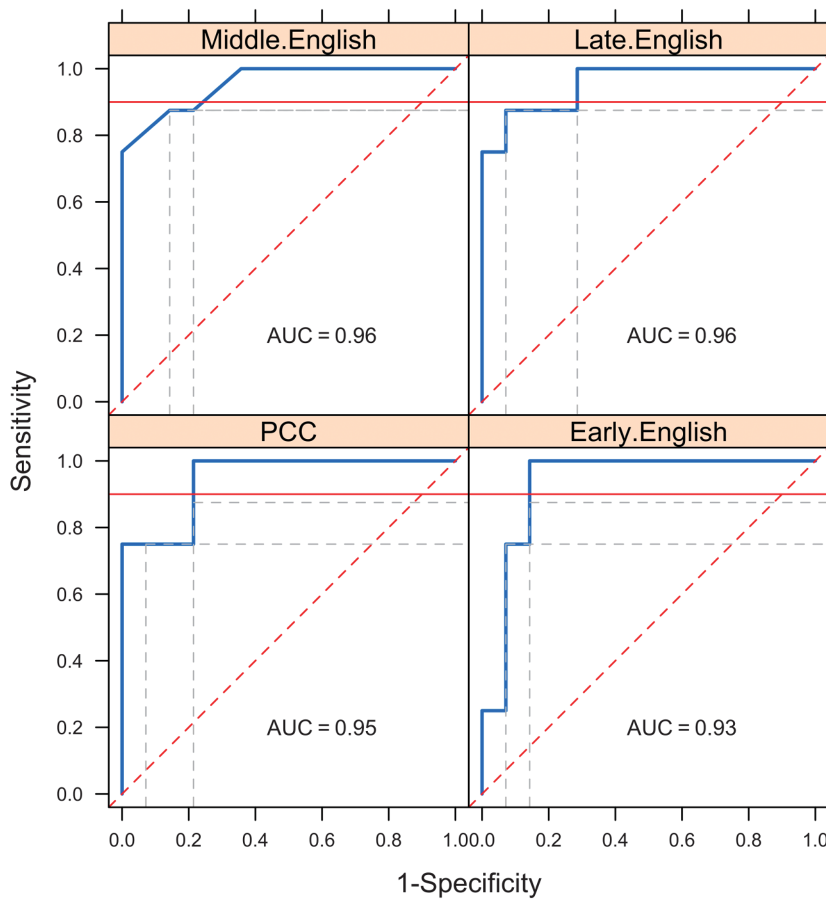
Figure 3. Receiver operating characteristic curves for Percent Consonants Correct–Revised (PCC-R) and percent accuracy of early-, middle-, and late-developing sounds in English for monolingual English-speaking children. AUC = area under the curve.



specificity for preschoolers entering kindergarten. When the English performance of bilingual preschoolers was examined in previous studies, (a) typically developing 3-year-old bilingual children exhibited lower accuracy as compared to their monolingual, age-matched peers (Fabiano-Smith & Goldstein, 2010a, 2010b), and (b) bilingual children who are more proficient in English than in Spanish exhibited an approximate 10 percentage point lag behind Austin and Shriberg's (1996) published norms on 3-year-old monolingual English speakers (Gildersleeve-Neumann et al., 2008). These were younger preschoolers, however. The children in the current study (i.e., 5-year-olds, on average) mirror the skill set that was reported in Goldstein et al. (2005): Bilingual children, at the age of approximately 5;0 years, hit a proverbial ceiling in terms of phonological skills. As a result, they perform much like monolingual English-speaking children on measures of phonological abilities. The knowledge that many SLPs are assessing the phonological skills of bilingual children in English only (Skahan et al., 2007) paired with the current findings suggest that perhaps bilingual children entering kindergarten are at lower risk of misdiagnosis than their younger counterparts, when informal measures are used (misdiagnosis continues to be a problem when standardized tests of articulation, normed on monolingual English speakers, are employed by SLPs).

Another characteristic of the data that we observed was the high level of variability in performance among the bilingual children. Bilingual children are a heterogeneous group who experience differing levels of language input and output. Each bilingual child has a unique language environment at home, even if language exposure and use in school is consistent. These differing language experiences result in variability specific to bilingual populations. Ingram (1989) discussed the issue of variability in speech production for monolingual children as internally consistent but externally variable. For bilingual children, the same observation holds true. Previous studies have documented a high level of variability in the speech of typically developing bilingual children as compared to their monolingual peers (Fabiano-Smith & Goldstein 2010a, 2010b), and coupled with the presence of a speech sound disorder, we would expect even more variability in speech sound production (Shriberg, Gruber, & Kwiatkowski, 1994). High variability creates overlap between typical and disordered

Figure 4. Receiver operating characteristic curves for Percent Consonants Correct–Revised (PCC-R) and percent accuracy of early-, middle-, and late-developing sounds in English for the English productions of bilingual Spanish–English–speaking children. AUC = area under the curve.



groups, and this overlap impacts the validity of a criterion to differentiate between typical and disordered speech (Dollaghan & Campbell, 1998). These findings suggest that traditional measures and criteria developed for English-speaking children will most likely need to be adapted for use with bilingual children.

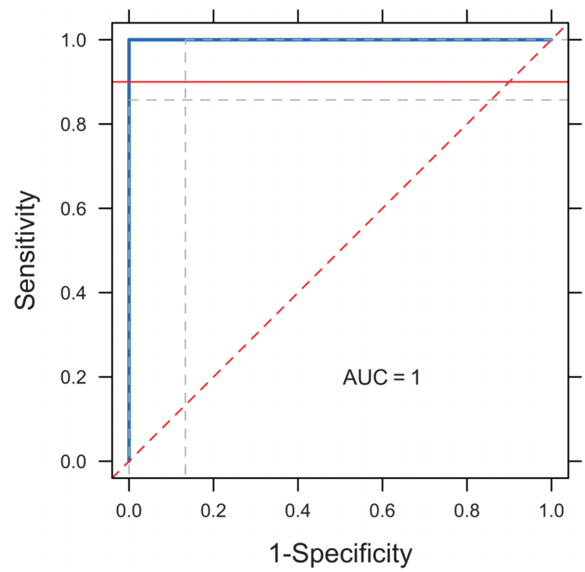
Some similarities do exist between monolingual and bilingual groups, however. For example, mean accuracy on PCC-R and EML-developing sounds were within a few percentage points for typical groups as well as for the two groups with disorders. It is important to note that our bilingual group with speech sound disorders exhibited a five percentage point difference as compared to the monolingual group with speech sound disorders on the gross measure of PCC-R. That difference is reflected in the cutoff scores for maximum sensitivity and specificity for PCC-R,

Table 6. Cutoff scores with maximum sensitivity and specificity for accuracy of early-, middle-, and late-developing sounds.

Measure	Estimate	*SE*	*z*	*p*	Cutoff
PCC-R monolingual English	−0.039	0.20	−1.89	0.06	89.18
Early sounds monolingual English	−13.530	12,338.48	−0.00	1.00	68.13
Middle sounds monolingual English	−0.410	0.23	−1.77	0.08	87.71
Late sounds monolingual English	−0.440	0.30	−1.45	0.15	85.71
PCC-R bilingual	−0.200	0.08	−2.46	0.01	84.62
Early sounds bilingual	−0.170	0.07	−2.39	0.02	90.70
Middle sounds bilingual	−0.400	0.25	−1.60	0.11	86.21
Late sounds bilingual	−0.130	0.05	−2.67	0.01	84.00

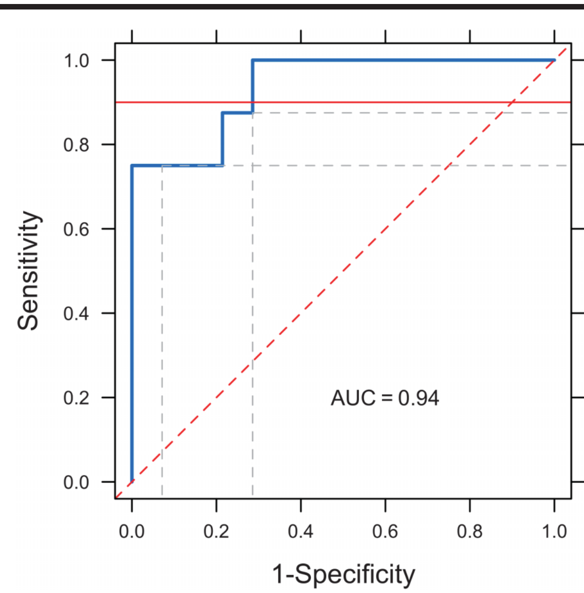
Note. PCC-R = Percent Consonants Correct–Revised.

Figure 5. Receiver operating characteristic curve derived from the support vector machine analysis for monolingual English-speaking children. AUC = area under the curve.



as monolingual English speakers exhibit a cutoff of 89%, whereas the English productions of bilinguals exhibit a cutoff of 84%. For this reason, it is important to consider not only diagnosis but also differences between monolinguals and bilinguals on classification of severity for children with speech sound disorders. We strongly argue the need for large group studies that examine the phonological abilities in bilingual children, in both of their languages, which mirror the work that Shriberg (1993) and colleagues have performed for monolingual English-speaking children. This study examined PCC-R and accuracy of EML-developing sounds as examples of informal measures developed for monolingual English-speaking children that are being utilized in the assessment of bilingual children with suspected phonological disorders. These are just a few of many measures that could possibly contribute to the misdiagnosis of bilingual children based on the lack of available norms for children who speak both English and Spanish (e.g., phonetic inventory complexity, type and frequency of phonological patterns, substitution error patterns). Both researchers and SLPs should take caution in using current measures and criterion levels to diagnose bilingual preschoolers with speech sound disorders until empirically derived criteria are obtained for this population.

Figure 6. Receiver operating characteristic curve derived from the support vector machine analysis for the English productions of bilingual children. AUC = area under the curve.



Acknowledgments

This work was funded through the National Institutes of Health Eunice Kennedy Shriver National Institute of Child Health and Human Development (1-R21-HD081382-01A1, granted to Leah Fabiano-Smith, PI). The authors thank the National Institutes of Health Loan Repayment Program–Division of Health Disparities for funding the first author through the course of the project. The authors would also like to thank the children and families who participated in this study, Lingling An for help with statistical analyses, research assistants Vania Pantoja and Chelsea Privette for help with data collection, and Elena Plante and Mary Alt for providing input on an earlier version of the manuscript.

References

American Speech-Language-Hearing Association. (2012). *ASHA 2012 Schools Survey: SLP caseload characteristics report*. Rockville, MD: Author.

Austin, D., & Shriberg, L. (1996). *Lifespan reference data for ten measures of articulation competence using the Speech Disorders Classification System (SDCS): Phonology Project Technical Report Number 3*. Madison: University of Wisconsin. Retrieved from http://www.waisman.wisc.edu/phonology/techreports/trep03.pdf

Barlow, J. A. (2003a). *Assessment of English Phonology (AEP)*. Unpublished assessment tool.

Barlow, J. A. (2003b). *Assessment of Spanish Phonology (ASP)*. Unpublished assessment tool.

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

Bunta, F., Fabiano-Smith, L., Goldstein, B., & Ingram, D. (2009). Phonological whole-word measures in 3-year-old bilingual children and their age-matched monolingual peers. *Clinical Linguistics & Phonetics, 23*(2), 156–175.

Burrows, L., & Goldstein, B. A. (2010). Whole word measures in bilingual children with speech sound disorders. *Clinical Linguistics & Phonetics, 24*(4–5), 357–368.

Chandaka, S., Chatterjee, A., & Munshi, S. (2009). Cross-correlation aided support vector machine classifier for classification of EEG signals. *Expert Systems With Applications, 36*(2), 1329–1336.

De Lamo White, C., & Jin, L. (2011). Evaluation of speech and language assessment approaches with bilingual children. *International Journal of Language and Communication Disorders, 46*(6), 613–627.

Dollaghan, C., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41,* 1136–1146.

Fabiano-Smith, L., & Goldstein, B. A. (2010a). Early-, middle-, and late-developing sounds in monolingual and bilingual children: An exploratory investigation. *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.

Garrett, K. K., & Moran, M. J. (1992). A comparison of phonological severity measures. *Language, Speech, and Hearing Services in Schools, 23,* 48–51.

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.

Goldman, R., & Fristoe, M. (1986). *Goldman-Fristoe Test of Articulation*. Circle Pines, MN: American Guidance Service.

Goldstein, B., & 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*(3), 153–164.

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*(3), 238–247.

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

Hambly, H., Wren, Y., McLeod, S., & Roulstone, S. (2013). The influence of bilingualism on speech production: A systematic review. *International Journal of Communication Disorders, 48*(1), 1–24.

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology, 143*(1), 29–36.

Ingram, D. (1989). *First language acquisition: Method, description and explanation*. Cambridge: Cambridge University Press.

Kamhi, A. G. (2006). Treatment decisions for children with speech–sound disorders. *Language, Speech, and Hearing Services in Schools, 37*(4), 271–279.

Kaufman, A. S., & Kaufman, N. L. (1983). *Kaufman Assessment Battery for Children*. John Wiley & Sons, Inc.

Kritikos, E. P. (2003). Speech-language pathologists' beliefs about language assessment of bilingual/bicultural individuals. *American Journal of Speech-Language Pathology, 12,* 73–91.

Miccio, A. W. (2002). Clinical problem solving: Assessment of phonological disorders. *American Journal of Speech-Language Pathology, 11,* 221–229.

Miller, G. A., Heise, G. A., & Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology, 41*(5), 329–335.

Oller, K., & Delgado, R. (2000). Logical international phonetics program (Version 2.02) [Computer software]. Miami, FL: Intelligent Hearing Systems.

Oller, D. K., Eilers, R. E., Basinger, D., Steffens, M. L., & Urbano, R. (1995). Extreme poverty and the development of precursors to the speech capacity. *First Language, 15*(44), 167–187.

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

Proposition, Arizona. (2000). *203. English for Children*.

Restrepo, M. A. (1998). Identifiers of predominantly Spanish-speaking children with language impairment. *Journal of Speech, Language, and Hearing Research, 41*(6), 1398–1411.

Shriberg, L. D. (1993). Four new speech and prosody-voice measures for genetics research and other studies in developmental phonological disorders. *Journal of Speech and Hearing Research, 36,* 105–140.

Shriberg, L. D., & Austin, D. (1997). *Lifespan reference data for ten measures of articulation competence using the Speech Disorders Classification System (SDCS)(Phonology Project Technical Report No. 3)*. Madison, WI: University of Wisconsin-Madison, Waisman Center.

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

Shriberg, L. D., Austin, D., Lewis, B. A., McSweeny, J. L., & Wilson, D. L. (1997b). The speech disorders classification system (SDCS): Extensions and lifespan reference data. *Journal of Speech, Language, and Hearing Research, 40,* 723–740.

Shriberg, L. D., Gruber, F. A., & Kwiatkowski, J. (1994). Developmental phonological disorders III: Long-term speech-sound normalization. *Journal of Speech and Hearing Research, 37*(5), 1151–1177.

Shriberg, L. D., & Kwiatkowski, J. (1982). Phonological disorders III: A procedure for assessing severity of involvement. *Journal of Speech and Hearing Disorders, 47,* 242–256.

Shriberg, L. D., & Kwiatkowski, J. (1994). Developmental phonological disorders I: A clinical profile. *Journal of Speech, Language, and Hearing Research, 37*(5), 1100–1126.

Skahan, S. M., Watson, M., & Lof, G. L. (2007). Speech-language pathologists' assessment practices for children with suspected speech sound disorders: Results of a national survey. *American Journal of Speech-Language Pathology, 16,* 246–259.