Title: Transcription of the speech of multilingual children with speech sound disorders

Authors: Jan Edwards[1] and Benjamin Munson[2]

Phonetic transcription of authors' names:

Affiliations: [1]Department of Communicative Disorders, University of Wisconsin-Madison, Madison, WI, USA and [2]Department of Speech, Language, and Hearing Sciences, University of Minnesota, Minneapolis, MN, USA

Email addresses: jedwards2@wisc.edu, munso005@umn.edu

Biographies:

Jan Edwards is a professor of Communicative Disorders at the University of Wisconsin, Madison. Her research focuses on lexical and phonological development of children learning different first languages, including children with both typical and atypical language development.

Benjamin Munson is an Associate Professor of Speech-Language-Hearing Sciences at the University of Minnesoata, Minneapolis. His research focuses on speech production-speech perception dynamics, relationships between word learning, and socially meaningful variation in speech in children with typical development, children with atypical language development, and normal adults.

Languages included: English, Greek, Japanese, Spanish

It goes without saying that our assessments of and treatments for children with speech-sound disorder rest on the reliability and validity of our transcriptions of their speech. An examination of the literature on multilingual phonological acquisition makes it clear that there is a need for research to develop evidence-based best practices for phonetic transcription of multilingual children, both those with typically development and those with speech sound disorders. The most common practice is for a bilingual speaker to transcribe phonetically children's productions in both languages (e.g., Fabiano-Smith & Goldstein, 2010; Faingold, 1996; Goldstein, Fabiano, & Washington, 2005; Holm & Dodd, 1999; Holm, Dodd, Stow, & Pert, 1999; Paradis, 2001; Vihman, 2002). All of the studies on multilingual phonological acquisition make several assumptions. They assume that transcription alone is adequate to describe the speech sound system of bilingual children and they assume that differences in children's phonological systems across languages can be described in terms of IPA phoneme categories. The purpose of this chapter is first, to question these assumptions based on the cross-linguistic research literature on first language phonological acquisition, and then to consider the clinical implications of these assumptions.

Phonemes are not Platonic Ideals, or, an /s/ by the Same Name is not really the Same

The first problem with the methodology described above is that there are cross-linguistic differences in the denotational values of the transcription system itself. That is, the same symbol does not necessarily denote the same sound across languages. We tend to think of IPA symbols as a universal denotational system — as if the same symbol reliably denotes the same sound across languages. After all, IPA does stand for the *international* phonetic alphabet, doesn't it? However, the same symbol does not always stand for the same sound across different languages. The voicing contrast for stops is probably the best-known example of this. Researchers have known for more than 40 years that there are three basic voicing categories for word-initial stop consonants that can be defined primarily in terms of voice onset time (VOT, Lisker & Abramson, 1964). These three categories are prevoiced stops (voicing begins prior to the stop release), short-lag stops (voicing begins at or almost immediately after the stop release), and voiceless aspirated stops (voicing begins considerably after the stop release, with a period of aspiration between the stop release and the onset of voicing). Languages with a two-way voicing contrast generally use two of these three categories – either prevoiced vs. short-lag (e.g., European French [Allen, 1985], Spanish [Macken & Barton, 1980b]) or short-lag vs. voiceless unaspirated (e.g., English [Macken & Barton, 1980a], Cantonese [Clumeck et al., 1981]). In languages which contrast prevoiced vs. short-lag stops, the symbols /b, d, ɡ / are used to represent the prevoiced stops and the symbols /p, t, k/ are used to represent the short-lag stops. In languages which contrast short-lag vs. voiceless aspirated stops, the symbols /b, d, ɡ/ can be used to represent the short-lag stops and the symbols /p, t, k/ represent the aspirated stops, to avoid the awkwardness of the aspiration diacritic. Thus, the short-lag stops can be represented by the symbols /b, d, ɡ/ in one set of languages and by the symbols /p, t, k/ in another set of languages. This can lead to confusion for English-speakers learning a second language, such as the 10-year-old American boy living in France who decided that the French word for *tag* was *douche* (*shower*) instead of *touche* (*touch*).

We have also known for some time that these cross-language differences in the phonetics of the voicing contrast explain seemingly contradictory acquisition patterns across languages. Short-lag stops are acquired earliest across languages, regardless of whether they are the /b, d, ɡ/ of English or the /p, t, k/ of Spanish (e.g., Macken & Barton, 1980a, 1980b). Voiceless unaspirated stops are acquired next, and prevoiced stops are acquired last (e.g., Allen, 1985;

Davis, 1995; Gandour et al, 1986; Macken & Barton, 1980a, 1980b). As Kewley-Port and Preston (1974) point out, these acquisition patterns can be explained in terms of the relative difficulty of satisfying aerodynamic requirements for the different stop types. The buildup of oral air pressure during stop closure inhibits voicing even when the vocal folds are adducted, so producing prevoiced stops requires the child to perform other maneuvers, such as expanding the pharynx. The production of voiceless aspirated stops is not as complex, but it does require the child to keep the glottis open exactly long enough after the release of the oral closure to create an audible interval of aspiration during the first part of the following vowel.

If cross-linguistic phonetic differences were as simple as we have described thus far, then it would be relatively easy to capture them within IPA using the standard IPA conventions for differentiating "narrow" phonetic transcription from "broad" phonemic transcription. That is, [b, d, ɡ] could be used to denote voiced stops, [p, t, k] could denote voiceless unaspirated stops, and [pʰ, tʰ, kʰ] could denote voiceless aspirated stops, even if the phonemic transcription uses only the unadorned /b, d, ɡ / versus /p, t, k/. In fact, many phoneticians use "narrow" transcription in this way already. However, the phonetic differences are actually more complicated than this. For example, Canadian French is different from European French in having shifted the voiced-voiceless distinction slightly, but not completely, in the direction of the English one (Caramazza and Yeni-Komshian, 1974). Riney et al. (2007) show that VOT values for Japanese voiceless stops are similar to those in Canadian French, and Kong (2009) provides data showing that VOT is necessary but not sufficient to describe the two-way voicing contrast in Japanese. While VOT alone correctly categorizes 94% of the stop consonants produced by 2- to 5-year-old English-speaking children, it correctly categorizes only 80% of the stop consonants produced by Japanese-speaking children in the same age range. Adding H1-H2 of the following vowel at vowel onset (the amplitude difference between the first and second harmonic, an acoustic measure of breathiness of the onset of the vowel) is needed to improve classification for the productions of the Japanese-speaking children.

Several other results of a recent series of cross-linguistic studies of the acquisition of lingual obstruents reinforce the suggestion that differences in the production of what are ostensibly the "same" sounds across different languages (Cantonese, English, Greek, Japanese, Korean, and Mandarin) are both much more pervasive and much more complex than has been described previously (Arbisi-Kelm et al., 2009; Edwards & Beckman, 2008a, 2008b; Li et al., 2009; Kong et al., 2007). We will illustrate with two examples from the παιδολογος database, (http://www.ling.ohio-state.edu/~edwards). This database consists of single word productions of familiar words and nonwords from at least 20 adults and 100 children aged 2 through 5 years for each of the six languages. The productions were elicited by a combination of a picture and an auditory prompt. All words and nonwords contain word-initial lingual obstruents followed by one of the five vowels (/i, e, a, o, u/) and were transcribed by an adult native speaker who was also a trained phonetician.

One example of the complexity of these cross-linguistic differences is exactly the contrast that we have already discussed, the voicing contrast. Kong et al. (2007) observed that children acquiring Greek correctly produced prevoiced stops at a much younger age than had been described in the literature for children learning other languages with a contrast between prevoiced and short lag stops. On investigating this phenomenon further, Kong found that the word-initial prevoiced stops in Greek are optionally prenasalized in adult productions. This prenasalization essentially solves the problem of maintaining voicing during closure because the speaker can vent air through the nasal cavity. Thus, prevoiced stops are acquired earlier in Greek than in French because Greek-acquiring children have the option of prenasalization and French-acquiring children do not. Similarly, voiceless unaspirated stop allophones of phonemically voiced stops are acquired later in Japanese than in English because Japanese-speaking children have to learn to control two parameters (VOT and degree of breathiness as measured by H1-H2), while English-speaking children only have to learn to control VOT (Kong et al.,2009).

Another example of a cross-linguistic difference in sound production concerns the most commonly occurring fricative in the world's languages, /s/. Typical descriptions of English /s/ are that it has a relatively long interval of aperiodic noise, with a concentration of energy in the higher frequencies. Cross-linguistic differences in the acoustic characteristics of /s/ were the subject of a recent study by Li, Edwards, and Beckman (2009). Li et al. examined Japanese- and English-speaking adult and children's productions of /s/ and the corresponding post-alveolar fricative. In descriptions of Japanese in the English-language literature, it is typical to equate the two post-alveolar sounds as well as the alveolar/dental sounds, reflecting the cross-language assimilation patterns that we have already noted in loan words such as *sushi*, although the Japanese post-alveolar fricative has a higher second-formant frequency at vowel onset than the English /ʃ/, as well as a concentration of energy in the higher frequencies overall than /ʃ/ (Li et al.). Somewhat surprisingly, Li and colleagues also found that the acoustics of /s/ differed across the two languages. The /s/ of English was much louder and had a more-compact spectrum than Japanese /s/. Li et al. showed that the two fricatives in the adult English speakers could be discriminated with high accuracy using just one parameter, centroid frequency. In Japanese, two parameters were needed: centroid, and the frequency of the second formant at the onset of the following vowel.

Again, these cross-linguistic differences appear to explain a cross-language asymmetry. English- and Japanese-acquiring two- and three-year-old children produce /s/ with very different accuracy rates. As described by Li et al. (2009), Japanese-acquiring 2-year-old children produced /s/ with an accuracy rate just over 30%, while English-acquiring children produced it with over 70% accuracy rates. More surprisingly, however, the two posterior fricatives, whose articulatory characteristics differ so much more sharply across these languages, were produced with very similar accuracy rates. To examine why this is so, Li, Munson, Edwards, Beckman, Yoneyama, and Hall (2010) conducted a cross-linguistic perception study in which English listeners (tested in Minneapolis, US) and Japanese listeners (tested in Tokyo, Japan) were presented with children's productions and asked to determine (in one block) whether they were instances of correct /s/, and in the other whether they were instances of correct /ʃ/ (for English listeners. Responses were pooled over the listeners and were categorized as either instances of /s/, /ʃ/, or neither (a category for sounds that reliably received 'no' answers in both blocks of questions). The fricatives labeled as /s/ by the English-speaking adult listeners covered a larger part of the two dimensional centroid-by-onset-F2 space than did the fricatives labeled as /s/ by the Japanese adults. Similarly, the fricatives labeled as /ʃ/ by English adults occupied a smaller area in the two-dimensional space than did those labeled as /ʃ/ by the Japanese adults, though this difference was smaller than the difference in /s/. This finding suggests that the cross-linguistic difference in acquisition is the result in part of the greater willingness to label an ambiguous sound as /s/ on the part of the English listeners versus as /ʃ/ on the part of the Japanese listeners.

Critically, Li et al. show that cross-language differences in order of acquisition of phonemes may not be explained solely by the children's productions and the articulatory-motor demands of particular sounds (e.g., Kewley-Port & Preston, 1974). Rather, differences may also be related to the different ways that listeners in the ambient language perceive children's productions. Such a finding is potentially very powerful, as it suggests that something as seemingly objective as the perception of sounds that are ostensibly shared by languages might not be as objective as it seems.

Intermediate productions, Multilingualism, and Speech Sound Disorders

Another problem with relying solely on transcription is that it assumes that children proceed directly and categorically from incorrect productions to correct productions. Both researchers and clinicians have known for many years that this assumption is not correct. As early as 1980, Macken and Barton described the existence of covert contrast, a subphonemic difference between two phoneme categories that can be observed acoustically. In a longitudinal study of three children, Macken and Barton (1980a) found evidence of a covert contrast in voicing for stop consonants. While the VOTs for the target voiceless stop consonants produced by these children were systematically longer than the VOTs for the voiced stop consonants, all VOTs were within the voiced range and so all consonants were transcribed as voiced. Since then, covert contrast has been observed for many contrasts including place of articulation for stops (Forrest & Rockman, 1988), place of articulation for fricatives (Baum & McNutt, 1990; Li et al., 2009), and voicing for stops (Macken & Barton, 1980a; Maxwell & Weismer, 1982; Gierut & Dinnsen, 1986). Covert contrast has been observed in both children with typical development and children with speech sound disorders (e.g., Forrest et al., 1994; Hewlett, 1988;) and in children learning languages other than English (e.g., Li et al., 2009). Tyler and colleagues (Tyler et al., 1993) found that covert contrast was clinically significant. Children who produced a covert contrast made faster progress in therapy than children who produced no contrast at all.

In our own research, we have found that covert contrast is even more widespread than had been shown in previous studies and that even naïve adults can identify covert contrast, given the appropriate task. The παιδολογος database was transcribed by trained native-speaker phoneticians for each language. In addition to coding initial consonant productions as correct or incorrect, the transcribers were trained to code clear substitutions, distortions, and intermediate productions (Stoel-Gammon, 2001). For example, "s:θ" means ambiguous between [s] and [θ] but closer to [s]. We then asked naïve adult native speakers of English to rate children's productions of target /s/ and /θ/ using visual analog scaling (VAS). In VAS rating tasks, participants are asked to scale a psychophysical parameter by indicating their percept on an idealized visual display (e.g., Urberg-Carlson, Munson, & Kaiser, 2009). In one experiment (Schellinger, Edwards, & Munson, 2010), the stimuli were initial /s/-vowel and /θ/-vowel sequences extracted from English-speaking children's productions from the παιδολογος database. The stimuli included roughly equal numbers of tokens from six transcription categories: correct productions of "s" for target [s] to correct productions of "θ" for target [θ], with substitutions of "s" for target [θ], the two intermediate categories, and substitutions of "θ" for target [s] in between. The naïve listeners' task was to rate the stimuli along a scale from "the 'th' sound" at one end to "the 's' sound" at the other. We found significant differences between mean VAS ratings for each of the six transcription categories. In other words, naïve listeners even rate substitutions as less close to the target than correct productions. This result suggests that covert contrast is not an isolated phenomenon observed in a few research studies using acoustic analyses. Rather, covert contrast is pervasive and can be measured in native speaker responses.

To our knowledge, there is no research on the existence of covert contrast in multilingual children with or without speech sound disorders. In fact, there are no systematic studies of whether covert contrast, as measured perceptually by VAS, is more widespread in monolingual children with speech sound disorders relative to monolingual children with typical speech development. We might predict, however, that children with speech sound disorders will have more intermediate productions as their speech is generally transcribed as having more substitution errors and more protracted development than children with typical development. Are intermediate productions also more prevalent in the productions of young multilingual children? One might expect this to be the case if young children learning two or more languages have non-autonomous phonological systems, as Paradis (2001) suggested, but we don't yet have enough data to make a strong prediction either way.

Problems with Transcription and Clinical Practice

In this final section, we present several clinical stories that describe how these problems with phonetic transcription described above may influence clinical practice for speech sound disorders with bilingual children. We also describe some alternative strategies that clinicians may take.

Story 1: The setting is Marysville, OH, home of the first Honda plant in the United States. A four-year-old boy, Tetsuo, speaks Japanese at home and English at preschool. His father is fluent in both English and Japanese, while his mother relies primarily on Japanese. The parents bring their son, Tetsuo, to the speech-language-hearing clinic at Ohio State University in Columbus, OH. The mother is very concerned that her son doesn't produce some sounds correctly at the age of 4, while his older sister, Kiyoko, had "perfect speech" at the same age. The clinician decides to test Tetsuo in English, as he speaks both languages well and she does not speak Japanese. She gives him the *Goldman-Fristoe Test of Articulation-2* (Goldman & Fristoe, 2000) and also analyzes his phoneme production in a 50-utterance language sample. She finds that he scores above the mean on the GFTA-2 and has only a few errors on late-developing sounds (/r/, /l/, /z/, /ð/) on both the standardized test and the language sample. When she looks up the sound system of Japanese on Wikipedia, she realizes that Tetsuo's errors on /l/ and /ð/ may be due to the fact that these sounds do not exist in Japanese. However, when she explains her findings to the father (who then translates for the mother), the mother is upset and insists that her son has difficulty on other sounds also, such as the consonant in the first syllable of "sushi". The clinician then asks the father to elicit some Japanese/s/-initial words from Tetsuo, thinking that perhaps he produces /s/ correctly in English, but not in Japanese. However, when she transcribes the Japanese words, she also codes Tetsuo's productions of /s/ as correct.

What is going on here?

Japanese speakers have a smaller acoustic space for /s/, as compared to English speakers (Li et al., 2010). Thus, when the Japanese-speaking mother and the English-speaking clinician are listening to the *same* productions, the English speaker hears correct /s/ and the Japanese speaker hears [ʃ] for /s/ substitutions.

What could this clinician do?

Tetsuo does not make enough speech sound errors to warrant a diagnosis of speech sound disorder by most clinical criteria. However, the fact that his mother has expressed concern about his Japanese abilities indicates that it is of great cultural importance for him to produce /s/ with articulatory and acoustic characteristics that are appropriate for Japanese. This clinician has two invaluable resources at her disposal: an aware and engaged mother, and a sibling to serve as a peer model. The clinician could train the mother on the differences between English and

Japanese /s/, and could enlist Tetsuo's sister Kiyoko to model the Japanese and English /s/ tokens. The clinician could provide Tetsuo's mother with lists of cognate words like *sushi*, *sport* and *soccer* so that he can practice producing different types of /s/ in words that are otherwise similar in form.

Story 2: Two Spanish-English bilingual boys are five-years-old. They are best friends and they both have velar fronting (e.g., [d] for /ɡ/and [t] for /k). Both boys are enrolled in speech therapy at their dual-immersion elementary school in Madison, WI with two different bilingual therapists. Jose is in therapy for two months with Juanita and the problem resolves completely. Pedro is in therapy for three months with Maria and he is just beginning to produce alveolar stops in isolation. Pedro’s mother demands that he start seeing Juanita instead of Maria. However, even after Pedro has worked with Juanita for two months, he is only starting to produce alveolar stops consistently at the word level.

What is going on here? It is quite likely that Pedro and Jose had very different amounts of knowledge of alveolar stops when they begin therapy. It may be the case that Jose had a covert contrast between velar and alveolar stops before therapy began, but that Pedro did not.

What could the clinicians do? Juanita and Maria have several options. First, they could do a more fine-grained transcription by including intermediate categories such as “in between /t/ and /k/.” Juanita and Maria could also use VAS to rate their clients’ productions to determine if Pedro and Jose were producing a covert contrast. This information is critical for therapy, both for prognosis, and for planning treatment. If Jose has a covert contrast between velar and alveolar stops, then he needs a different approach to treatment than Pedro, who is neutralizing the contrast and may not even be able to perceive the difference between /t/ and /k/.

Story 3: The last scenario takes place in your own place of employment, and is a scenario that you probably have encountered many times before. You have been working with Evan, a 4-year-old boy with a severe speech-sound disorder for three months. You are using the cycles approach (Hodson & Paden, 1991), as Evan's speech was transcribed to have numerous substitution and deletion patterns resulting in similar sounding words (i.e., velar fronting, making *tea* and *key* sound similar, depalatalization, making *ship* and *sip* sound similar, final consonant deletion, making *keep* and *key* sound similar, and combinations like final consonant deletion and velar fronting, making *keep* and *key* and *teach* and *ti* sound similar). As is so often the case with children with multiple speech sound errors and numerous neutralizations, Evan's progress is slow. Your transcription of lists of probe words show no changes from baseline. Evan's parents perceive his speech as improving—extended family members and other caregivers report impressionistically that he's slightly more intelligible—but this is not reflected in your assessments. You're worried that he won't make the progress needed to continue on in therapy. You begin to worry that your approach to therapy isn't the best for her, and wonder whether you should use another, less-tested therapy, like non-speech oral-motor exercises.

What do you do? Before you make a big change in Evan's therapy plan you need to make sure that you're using the right assessment to measure his abilities. You remember from your graduate school coursework articles—including many referenced in this paper—illustrating that speech-sound acquisition is gradual. You wonder whether Evan is demonstrating the kind of gradual acquisition that is often obscured by transcriptions. You decide to measure his progress by taking pairs of productions that you transcribe as the same (like *tea* and *key*, or *sip* and *ship*), and rate the productions using a VAS scale similar to that used by Schellinger et al. (2010) and Urberg-Carlson et al. (2009). After doing this for 3 weeks (which ends up being six data points, as you see him twice a week), you see that indeed his productions are moving closer to the end-points of the VAS scale. You record some more productions and this time have another rater, a fellow clinician, do the VAS task. Your predictions are confirmed—he's making gradual progress toward the /t/, /k/, /s/, and /ʃ/ endpoints. You continue with therapy, and as you would predict, his productions soon become distinct enough that your transcriptions reflect the progress seen in VAS and noted by her family and caregivers. You are happy that you have saved Evan from the potential setback of changing the way you go about therapy. As you continue along in your therapy, you begin to incorporate VAS more and more, and find that it is very useful for demonstrating the small but the nonetheless important differences that are sometimes present across different dialects or different languages, like the differences between /t/ and /k/ that you encountered in bilingual Spanish-English acquiring children during a job you once had in Madison, WI, or the differences between Japanese and English /s/ that you encountered in a job in Columbus, OH.

Conclusion

To conclude, it is clear that the assumptions that clinicians and researchers make regarding phonetic transcription are problematic even for typically developing children learning only one first language. Transcription as the sole analysis tool is even more problematic when we are analyzing the speech of multilingual children or the speech of children with speech sound disorders, let along the speech of children with speech sound disorders who speak more than one language. The fine phonetic detail of a single sound based on IPA transcription differs from language to language. Furthermore, children don't proceed directly and categorically from incorrect to correct productions. These problems with transcription have real-life consequences for clinical assessment and treatment. We have suggested a few supplements to transcription that should be relatively easy for clinicians to implement and that should greatly improve their ability to describe the speech of multilingual children with speech sound disorders.

References

Allen, G. D. (1985). How the young French child avoids the pre-voicing problem for word-initial voiced stops. *Journal of Child Language*, *12*, 37-46.

Arbisi-Kelm, T., Beckman, M. E., Kong, E., & Edwards, J. (2009). Psychoacoustic measures of spectral properties of Cantonese, Greek, English, Japanese lingual stop bursts. Paper presented at the 2009 Linguistics Society of America Convention, San Francisco, 9-11 January 2009.

Baum, S. R. & McNutt, J. C. (1990). An acoustic analysis of frontal misarticulation on /s/ in children. *Journal of Phonetics, 18*, 51-63.

Caramazza, A. & Yeni-Komshian, G. H. (1974). Voice onset time in two French dialects. *Journal of Phonetics, 2*, 239-245.

Clumeck, H., Barton, D., Macken, M. A. & Huntington, D. A. (1981). The aspiration contrast in Cantonese word-initial stops: data from children and adults. *Journal of Chinese Linguistics, 9*, 210-224.

Davis, K. (1995). Phonetic and phonological contrasts in the acquisition of voicing: Voice onset time production in Hindi and English. *Journal of Child Language, 22*, 275-305.

Edwards J. & Beckman, M. E. (2008a). Methodological questions in studying consonant acquisition. *Clinical Linguistics and Phonetics, 22*, 937-956.

Edwards, J. & Beckman, M.E. (2008b). Some cross-linguistic evidence for modulation of implicational universals by language-specific frequency effects in phonological development. *Language, Learning, and Development*, *4*, 122-156.

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

Faingold, E. D. (1996). Variation in the application of natural processes: Language-dependent constraints in the phonological acquisition of bilingual children. *Journal of Psycholinguistic Research, 25*, 515-526.

Forrest, K. & Rockman, B. K. (1988) Acoustic and perceptual analyses of word-initial stop consonants in phonologically disordered children. *Journal of Speech and Hearing Research, 31*, 449-459.

Forrest, K., Weismer, G., Hodge, M., Dinnsen, D. A. & Elbert, M. (1990) Statistical analysis of word-initial /k/ and /t/ produced by normal and phonologically disordered children. *Clinical Linguistics and Phonetics, 4*, 327- 340.

Gandour, J., H., Petty, S. H., Dardarananda, R., Dechongkit, S., & Munkongoen, S. (1986). The acquisition of the Voicing Contrast in Thai: A Study of Voice Onset Time in Word-Initial Stop Consonants. *Journal of Child Language, 13*, 561-572.

Gierut, J. A. & Dinnsen, D. (1986) On word-initial voicing: Converging sources of evidence in phonologically disordered speech. *Language and Speech, 29*, 97-114.

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

Goldman, R. & Fristoe, M. (2000). *The Goldman-Fristoe Test of Articulation – Second Edition*. Circle Pines, MN: American Guidance Services.

Hewlett, N. (1988) Acoustic properties of /k/ and /t/ in normal and phonologically disordered speech. *Clinical Linguistics and Phonetics, 2*, 29-45.

Hodson, B. W. & Paden, E. P. (1991). *Targeting Intelligible Speech: A Phonological Approach to Remediation*. Austin, TX: Pro-Ed.

Holm, A. & Dodd, B. (1999). An intervention case study of a bilingual child with a phonological disorder. *Child Language Teaching and Therapy, 15*, 139-158.

Holm, A., Dodd, B., Stow, C., & Pert, S. (1999). Identification and differential diagnosis of phonological disorder in bilingual children. *Language Testing, 16*, 271-292.

Kewley-Port, D. & Preston, M. S. (1974). Early apical stop production: A voice onset time analysis. *Journal of Phonetics*, *2*, 195-210.

Kong, E. (2009). *The Development of Phonation-type Contrasts in Plosives: Cross-Linguistic Perspectives* (Unpublished Ph.D. Dissertation) Columbus, OH: Department of Linguistics, Ohio State University.

Kong, E., Beckman, M. E., & Edwards, J. (2007). Fine-grained phonetics and acquisition of Greek voiced stops. *Proceedings of the XVIth International Congress of Phonetic Sciences*, 6-10 August 2007, Saarbruecken.

Kong, E., Beckman, M. E., & Edwards, J. (2009). VOT is necessary but not sufficient for describing the voicing contrast in Japanese. Paper presented at the 2009 Linguistics Society of America Convention, San Francisco, 9-11 January 2009.

Li, F., Edwards, J., & Beckman, M.E. (2009). Contrast and covert contrast: The phonetic development of voiceless sibilant fricatives in English and Japanese toddlers. *Journal of Phonetics*. *37*, 111-124.

Li, F., Munson, B., Edwards, J., Beckman, M.E., Yoneyama, K., & Hall, K. (2010, in press) Language specificity in the perception of voiceless sibilant fricatives in Japanese and

English: Implications for cross-language differences in speech-sound development. *Journal of the Acoustical Society of America*.

Lisker, L. & Abramson, A. (1964) A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*, 384-422.

Macken, M. & Barton, D. (1980a). The acquisition of the voicing contrast in English: A study of voice onset time in word-initial stop consonants. *Journal of Child Language*, *7*, 41-74.

Macken, M. & Barton, D. (1980b) The acquisition of the voicing contrast in Spanish: a phonetic and phonological study of word-initial stop consonants. *Journal of Child Language, 7*, 433-458.

Maxwell, E. M. & Weismer, G. (1982) The contribution of phonological, acoustic and perceptual techniques to the characterization of a misarticulating child's voice contrast for stops. *Applied Psycholinguistics, 3*, 29-43.

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

Riney, T. J., Takagi, N., Ota, K., & Uchida, Y. (2007). The intermediate degree of VOT in Japanese initial voiceless stops. *Journal of Phonetics*, *35*, 439-443.

Schellinger, S., Edwards, J., & Munson, B.(2010, submitted). The role of intermediate productions and listener expectations on the perception of children's speech.

Stoel-Gammon, C. (2001) Transcribing the speech of young children, *Topics in Language Disorders*, *21*, 12-21.

Tyler, A. A., Figurskl G. R. & Langdale, T. (1993). Relationships between acoustically determined knowledge of stop place and voicing contrasts and phonological treatment progress. *Journal of Speech and Hearing Research, 36*, 746 - 759.

Urberg-Carlson, K., Munson, B., & Kaiser, E. (2009). *Gradient measures of children's speech production: Visual analog scale and equal appearing interval scale measures of fricative goodness*. Poster presented at the spring 2009 meeting of the Acoustical Society of America.

Vihman, M. M. (2002). Getting started without a system: From phonetics to phonology in bilingual development. *International Journal of Bilingualism, 6*, 239-254.

Acknowledgments

This research was supported by NIH grant R01 DC02932 and NSF grant BCS0729140 to Jan Edwards and by NSF grant BCS0729277 to Benjamin Munson. We generously thank Mary E. Beckman for her long-time collaboration on issues related to this work, as well as Sarah Schellinger and Kari Urberg-Carlson for their instrumental roles in developing alternative measures of children's speech production accuracy.