Measuring Speech-Sound Learning Using Visual Analog Scaling

Benjamin Munson

Sarah K. Schellinger

Kari Urberg Carlson

Department of Speech-Language-Hearing Sciences, University of Minnesota
Twin Cities, MN

The ultimate goal for speech-language pathologists is to align the linguistic behaviors of the clients whom we serve with those of the ambient language of the community. In light of this goal, it is critical that change in speech production is measured accurately. In this article, we review the use of visual analog scaling as a measure of change in children's speech production. Following a discussion of this tool, the authors consider the clinical utility of this type of measurement.

One of the most critical challenges facing speech-language pathologists (SLPs) is facilitating change in speech and language behaviors. Regardless of our clients' starting states, our ultimate goal is to align our clients' linguistic behaviors with those of the ambient language community. Consequently, it is critically important that SLPs measure change accurately. In this brief article, we describe one method for measuring the accuracy of children's speech production, visual analog scaling (VAS). We argue that this method should be used to measure changes in speech production in children who are receiving therapy for speech sound disorders.

Change in Speech-Sound Development

Children's first words sound very much unlike adults' productions. Consider, for example, a child's early productions of the words *shoe* and *Sue*. A child might initially produce these two words in a manner that adults perceive to be identical to one another and with a sound that is broadly similar to the sound in adults' productions of the word *two*. Later in his/her development, this child might produce these words perceptually identically to one another, but with a sound that is similar to the adults' production of the first sound in *Sue*. Only later in life would this child produce *Sue* and *shoe* with initial sounds that are perceived to match those in the adult productions. This description of speech-sound development is consistent with findings of large-scale normative studies that use phonetic transcriptions, such as studies by Smit, Hand, Freilinger, Bernthal, and Bird (1990) and Edwards and Beckman (2008). This is also broadly consistent with many SLPs' observations of speech-sound learning by children with speech sound disorders. Our phonetic transcriptions of children's speech show a progression from productions that are transcribed as errors to productions transcribed with the same symbol as would be used to denote an adult's accurate production.

Detailed acoustic and articulatory studies paint a different picture of speech-sound acquisition. Consider, again, a child whose productions of *Sue* and *shoe* are perceived to be identical and to match the adult's production of *Sue*. Imagine that this child was followed from his/her first productions of the words *Sue* and *shoe* to well beyond the point at which these productions were transcribed as correct. In this hypothetical scenario, we would have access to high-quality recordings of the child's speech. Imagine that we were able to capture the child's productions of the word-initial fricatives /s/ in *Sue* and /ʃ/ in *shoe* with a single acoustic measure. What would the trajectory of acquisition look like? Would it occur in a step-wise fashion, with an "a-ha" moment corresponding to the sudden acquisition of the /ʃ/ sound?

Figure 1 shows one possible outcome of this study. In this graph, the y-axis represents a hypothetical acoustic measure differentiating /s/ from /ʃ/. Readers can think of it as analogous to the frequency in the fricative that has the most energy; this is higher in /s/ than in /ʃ/ in adults' speech (e.g., Jongman, Wayland, & Wong, 2000). The x-axis represents a hypothetical variable indicating developmental time. The solid line shows how the peak frequencies of /s/ might change across early development, and the dotted line shows how the peak frequencies of /ʃ/ might change during this same time. The data in this figure suggest that development involves the gradual differentiation of /s/ from /ʃ/. That is, both /s/ and /ʃ/ become gradually more different from one another over the course of development. The /s/ reaches adult-like levels by time 12, while /ʃ/ does not achieve adult-like values until time 30.

Figure 1. A hypothetical trajectory of acquisition of /s/ and /ʃ/.

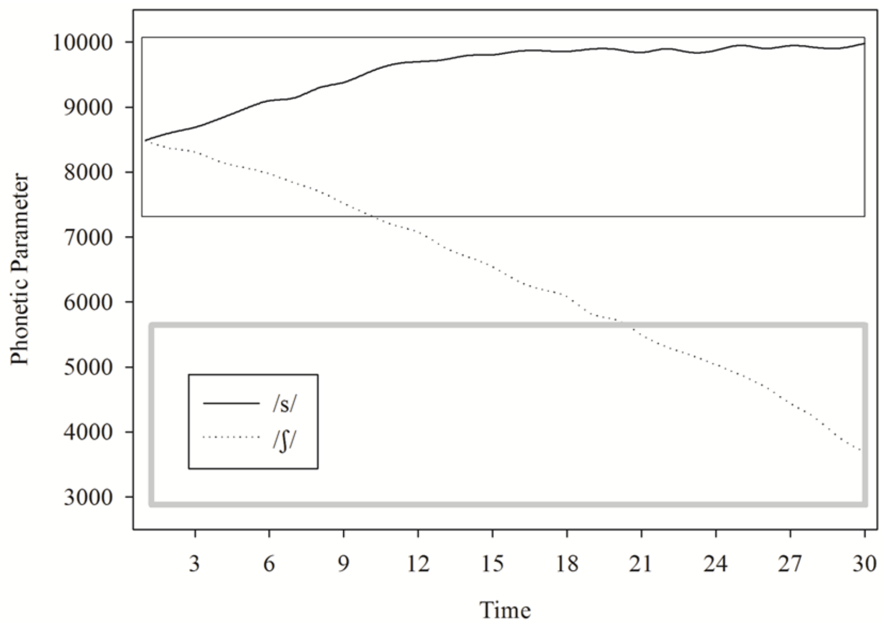


Table 1. Transcriptions of data in Figure 1.

	Time									
	3	6	9	12	15	18	21	24	27	30
Target /s/	[s]	[s]	[s]	[s]	[s]	[s]	[s]	[s]	[s]	[s]
Target /ʃ/	[s]	[s]	[s]	[s?]	[s?]	[s?]	[ʃ]	[ʃ]	[ʃ]	[ʃ]

The scenario in Figure 1 is supported by recent acoustic studies of children's speech using cross-sectional data, such as Li, Edwards, and Beckman (2009) and Li (2008). How does this square with studies of speech-sound development that use phonetic transcription? Consider how this child's fricative productions might be categorized using phonetic transcription. Imagine that the area demarcated with a thin black line contains adults' productions of /s/ and the area demarcated with a thicker gray line contains adults' productions of /ʃ/. We would expect the phonetic transcriptions shown in Table 1. The transcriptions of the child's productions that are intermediate between the adults' /s/ and /ʃ/ are marked with an [s?] precisely because they fall outside of the adult norms. We might make several predictions about how these productions would be transcribed. First, we might expect that these transcriptions would be more variable across different listeners than would transcriptions of the productions that fall clearly into the adult categories. Second, a recent study by Li, Munson, Edwards, Yoneyama, and Hall (2011) found that English-speaking listeners tend to label children's fricative productions intermediate between /s/ and /ʃ/ as /s/. Therefore, we might predict that they would be labeled as /s/ more often than /ʃ/.

The transcriptions in Table 1 suggest the sudden acquisition of /ʃ/ at time 21. They do not show that productions of target /ʃ/ are gradually approximating adult-like values from times 1 to 21, nor do they show the refinement of /ʃ/ production that occurs after it has been transcribed as correct, as in Figure 1. Put simply, the transcriptions paint an incomplete portrait of the acquisition of these sounds. For a case in which phonetic transcription is potentially misleading, consider the scenario in Figure 2. In this case, the child's productions of target /s/ and /ʃ/ start out as both identical and unlike either the adult /s/ or the adult /ʃ/. Development once again involves the gradual differentiation of /s/ from /ʃ/, but this time from a production that is intermediate between /s/ and /ʃ/ to adult-like productions. Transcription (Table 2) is misleading in this case, because the initial transcriptions suggest an adult-like /s/ production. This scenario is consistent with studies of speech development using *electropalatography*, a technique in which contact between the tongue and the palate is measured during real-time speech production using pseudopalates equipped with special sensors. These studies have shown that many children produce *merged* articulations of sounds—that is, productions that are intermediate between two sounds and resemble no typical adults' productions (Gibbon, 1999). The probable transcriptions of these articulations shown in Table 2 are similar to those in Table 1. Phonetic transcription misses the important differences between these two acquisition scenarios.

Figure 2. A hypothetical trajectory of acquisition of /s/ and /ʃ/.

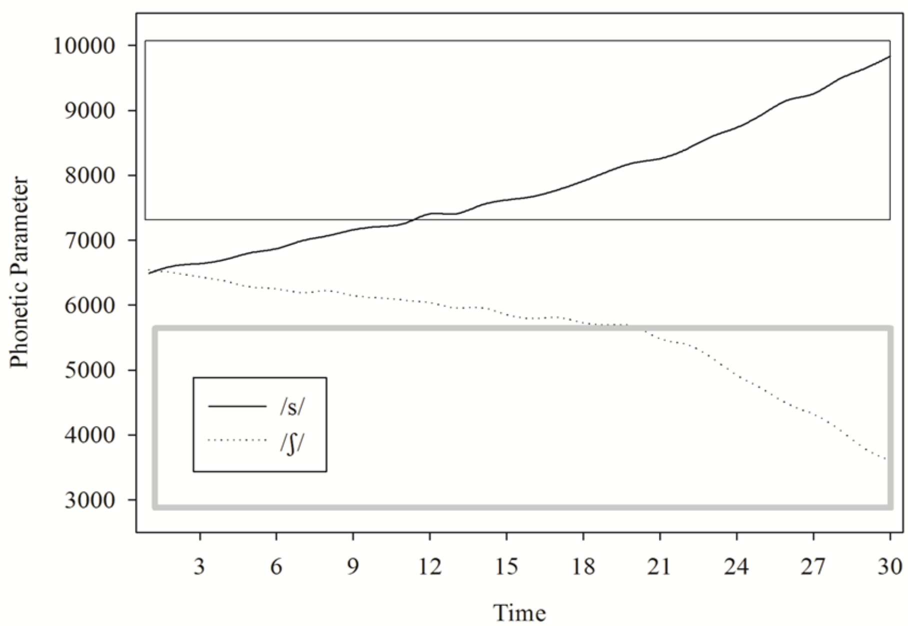


Table 2. Transcriptions of data in Figure 2.

	Time									
	3	6	9	12	15	18	21	24	27	30
Target /s/	[s?]	[s?]	[s?]	[s]	[s]	[s]	[s]	[s]	[s]	[s]
Target /ʃ/	[s?]	[s?]	[s?]	[s?]	[s?]	[s?]	[ʃ]	[ʃ]	[ʃ]	[ʃ]

How is this relevant for clinical practice? The goal of a child who produces an /s/-like sound at the beginning of *shoe* is to produce an adult-like /ʃ/. It is quite reasonable for a clinician to ask whether it is important to determine if a child produces a true adult-like /s/ or something that is intermediate between and /s/ and /ʃ/. After all, our overall goal for children with speech sound disorders is to sound like they produce speech correctly.

There are at least three reasons why we think it is clinically important to assess speech production in more detail than phonetic transcription allows—in essence, to assess whether a child fits into Scenario 1 or 2 and where the child is along the trajectory of development. The first reason relates to clinical outcomes. Tyler, Figurski, and Langsdale (1993) examined the progress through therapy of children whose productions of certain pairs of consonants were perceived to be identical. They found that the type of production error—a true substitution error (something like what we see at time 1 in Figure 1/Table 1) or a contrast that is not perceptible (something like what we see at time 9 or even 15 in Figure 1/Table 1)—has consequences for the child's progress through therapy. Children whose productions of sounds were perceived to be acoustically distinct progressed through therapy more quickly and generalized correct production more readily than did children whose productions were not acoustically distinct. Hence, clinicians can make better prognostic statements if they know the nature of a child's production. This could be particularly important when children have multiple errors and clinicians need to prioritize treatment. Clinicians might choose to treat true substitutions (i.e., "time 1" errors) on the assumption that they would require more time and effort to remediate. Clinicians might also choose a wait-and-see approach for children who are already making a distinction between two sounds and may be able to develop mature productions without intervention. They would choose to immediately treat a child who is truly producing the same two sounds identically.

The second motivation for assessing speech production in detail concerns the type of therapy approach that a clinician might use. Some therapy programs, such as Metaphon (Dean, Howell, Waters, & Reid, 1995), are based on the assumption that children's errors reflect conceptual difficulties with speech sounds, such as a lack of understanding that *Sue* and *shoe* should sound different. It is not surprising that these therapies involve numerous activities to enhance children's conceptual knowledge of speech sounds. Other therapy approaches, such as McDonald's (1964) sensory-motor therapy, are based on the assumption that incorrect productions arise primarily from perceptual-motor problems. These therapies focus on training the perpetual and articulatory characteristics of sounds. It is reasonable to assume that children with acoustically/phonetically undifferentiated /s/ and /ʃ/ productions may have more of a conceptual problem than do children who are able to produce a phonetic distinction between these two sounds. It is also reasonable to postulate that the clients whose speech is like those at time 1 in Figure 1 and Figure 2 would need different types of therapy: the child in Figure 1 needs to learn how to correctly produce /ʃ/, whereas the child in Figure 2 needs to learn how to produce /s/ and /ʃ/.

The final motivation for assessing speech production in detail concerns measuring change in production. SLPs are all too familiar with the following scenario. A child begins therapy and, after 4 weeks or so, appears to be making no progress, as assessed by phonetic transcriptions of productions of words on a list of probe words. The SLP carefully considers why no progress is being made and can find no obvious source: the child is engaged in therapy and attends regularly; his parents report dutifully implementing home-programming activities; and the clinician is confident that there isn't an undiagnosed mitigating factor, like a hearing impairment, ADHD, or a broader developmental disability. The fault must, therefore, be with the type of therapy being offered. The therapist changes tactics, perhaps shifting from a conceptual approach to a motor-based approach, and hopes that the child will make progress.

Surely, there are many cases in which children don't progress through therapy precisely because the therapy modality is not optimal. Consider, however, the data in Figures 1 and 2 as representing not time in development for a child with typical speech and language, but time in therapy for a child who is receiving treatment for a speech sound disorder. From time 1 to time 18, both of these children are making progress in learning to produce /s/ and /ʃ/ that wouldn't be reflected in phonetic transcriptions of their speech. If a clinician were able to assess a child's progress at this level of detail, then he/she might see that the therapy is, in fact, effecting change, even though this change isn't reflected in phonetic transcriptions.

Visual Analog Scaling

Clinicians don't have many choices of tools to measure gradual change in speech-sound learning. The most obvious tool, based on previous research, is acoustic analysis. At first glance, this option seems attractive. We live in the "golden age" of affordable and accessible acoustic analysis tools. High-quality microphones cost no more than most of the standardized assessments that SLPs currently use. These can be plugged into laptop computers, which can serve as high-quality digital recorders with built-in storage space. Many different acoustic analysis software packages, like Praat (Boersma, 2001), are free and supported by a world-wide community of users.

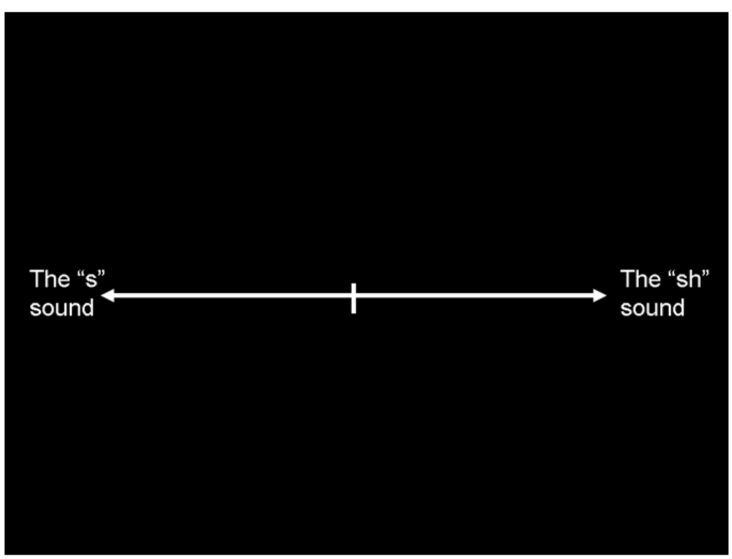
Unfortunately, there are challenges—some of them truly intractable—to implementing acoustic analysis in the clinic. First, a large-scale normative study of the acoustic characteristics of children's speech would need to be conducted before we can fully understand which acoustic characteristics relate to articulation and which relate to other linguistic and non-linguistic factors, such as the size of a child's vocal tract or the unique phonetic characteristics of the speaker's dialect. A normative study like this wouldn't be impossible, but our discipline is many years away from such a study being completed. More important, many of the acoustic measures needed to characterize speech sounds are highly sensitive to background noise. It would be impossible to get a valid acoustic measure of the difference between /s/ and /ʃ/ with a recording made in background noise, as the spectrum of the background noise would potentially obscure important parts of the spectra of /s/ and /ʃ/. Though microphones, computers, and acoustic analysis software are relatively cheap and portable, double-walled sound-treated booths are not, and without them a clinician would be left with hopelessly messy measures. Another promising set of tools comprises direct articulatory measures like electropalatography, which shows patterns of tongue-palate contact, and ultrasound, which can show movement of the tongue. Systems for electropalatography and ultrasound are becoming increasingly more affordable. It may not be long before these systems are in reasonably widespread use in clinical settings. However, as of the writing of this article, their prices remain in the four- to five-figure range.

Recently, our lab group began to investigate whether any of the many published techniques for measuring perception can be used to measure the kind of developmental progression illustrated in Figures 1 and 2. We considered many different paradigms. We had four criteria for an optimal measure. First, the response had to be continuous; that is, listeners' responses couldn't be simply "x" or "y," but had to index the degree of "x"-ness or "y"-ness of each stimulus. In the case of /s/ and /ʃ/, the measure could not simply be a forced choice judgment of "s" or "sh," but had to indicate how much each stimulus was like an ideal /s/ or /ʃ/. Second, the measure had to have good intra-rater reliability. Third, the measure had to correlate well with key acoustic characteristics that differentiate the sounds being rated.

In the case of /s/ and /ʃ/, the acoustic measure that best differentiates these sounds in English is the peak frequency of the frication noise. Finally, these measures should be straightforward to collect and interpret. That is, it should be no more difficult to use these tools than it is to use phonetic transcription. We considered a number of techniques that are well-established in the perception literature, including reaction times in forced-choice categorization (Whalen, 1991), Likert-type equal-appearing interval scales (Urberg-Carlson, Munson, & Kaiser, 2009), and direct magnitude estimation (Urberg-Carlson, Kaiser, & Munson, 2008). The technique we found most useful was visual analog scaling (VAS). In VAS, the auditory perceptual space is made analogous to a visual space. Listeners hear a stimulus and respond by indicating in this visual space where the stimulus falls in relation to the criteria. VAS is used clinically in the assessment of voice disorders using the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V; Kempster, Gerratt, Abbott, Barkmeier-Kraemer, & Hillman, 2009), which was developed by ASHA's Special Interest Group 3, Voice and Voice Disorders, to standardize perceptual judgment of voice quality.

In our first studies using VAS, we examined listeners' perception of children's productions of minimal contrasts such as /s/–/ʃ/, /s/–/θ/, /t/–/k/, and /d/–/g/. These productions were taken from a large database of speech samples of 2- to 5-year-old children acquiring English (Edwards & Beckman, 2008). The contrasts were chosen because they represent contrasts that are typically later acquired by children learning English. Preliminary summaries of the results of these studies can be found in Arbisi-Kelm, Edwards, Munson, and Kong (2010); Munson, Edwards, Schellinger, Beckman, and Meyer (2010); and Urberg-Carlson et al. (2008). In these studies, we presented listeners with displays like that in Figure 3: a double-headed arrow with the text "the {s, s, t, d} sound" at one end and "the {sh, th, d, g} sound" at the other end. Listeners heard children's productions—consonant-vowel sequences excised from productions of real words and nonwords—and were asked to rate the initial consonant. They were told to click on the line wherever they believed the child's production to fall. If, for example, they thought it was closer to adults' productions of one sound, like /s/, then they should click closer to the text marked "the 's' sound." They were encouraged to use the entire line. Critically, they were not told to base their ratings on anything specific, but rather were allowed to apply their own criteria for how close a production was to an ideal /s/, /ʃ/, /θ/, /k/, /t/, /d/, or /g/. Our first set of studies involved listeners who didn't have any specialized training in speech or language. Recently, we have extended this work to look at experienced SLPs (Munson, Johnson, & Edwards, 2011).

Figure 3. A visual-analog display used to elicit ratings of children's productions of /s/ and /ʃ/.



Our preliminary analysis suggests that VAS meets all of our criteria for an optimal measure. First, it showed acceptably high levels of intra-rater reliability, even among listeners who had no specialized training in speech or language. Second, with just a few exceptions, listeners used the entire line when making their responses. That is, listeners' responses suggested that they perceived degrees of how much a sound was like /s/ or /ʃ/. Third, the responses were well-correlated with the acoustic characteristics of the sounds being rated. To illustrate this, look at Figure 4, which shows the relationship between where one participant clicked on the line labeled "the 's' sound" to "the 'sh' sound" (shown on the y-axis) and the centroid frequency of the fricatives—a measure related to peak frequency—that he was rating (on the x-axis). Recall that centroid frequency characterizes the difference between /s/ and /ʃ/, with higher values associated with more /s/-like productions. As this figure shows, this person's click locations were well-correlated with click locations on the line. They also covered the entire line, rather than just clustering around the endpoints. Figure 5 shows another interesting fact about VAS ratings: In this figure, data from the same listener as in Figure 4 are plotted against a different characteristic of the stimuli, the proportion of times a different group of listeners (those from Li et al., 2011) perceived each stimulus to be "s" in a forced-choice classification task. As this figure shows, the stimuli that the listener rated as close to the center of the line (i.e., neither clearly /s/ nor clearly /ʃ/) were also those that the listeners in Li et al. (2011) disagreed about most.

Figure 4. Relationships between VAS click locations and centroid frequencies for one listener.

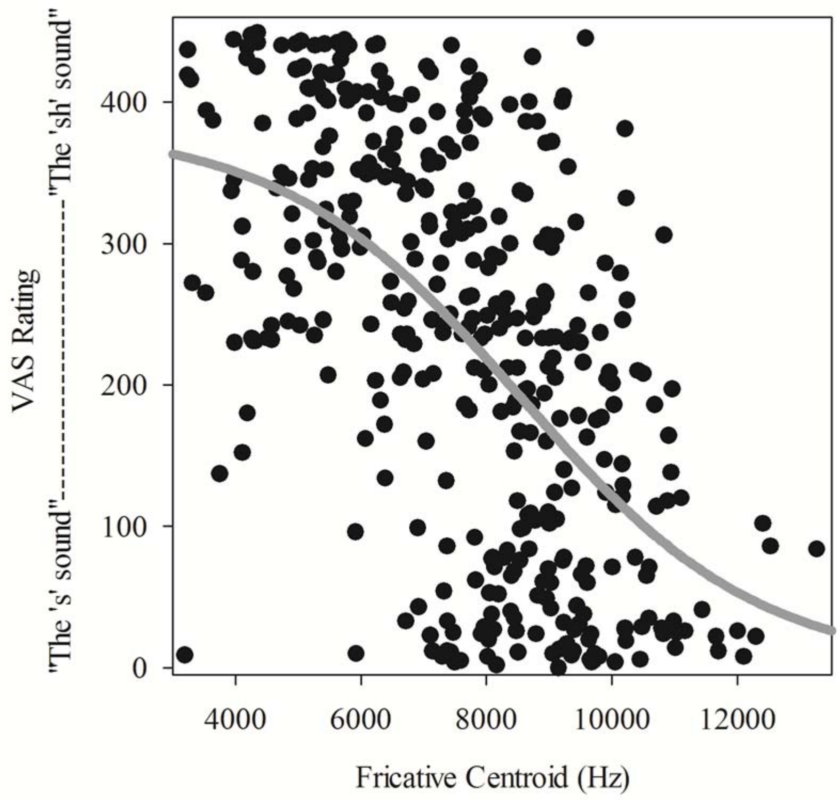
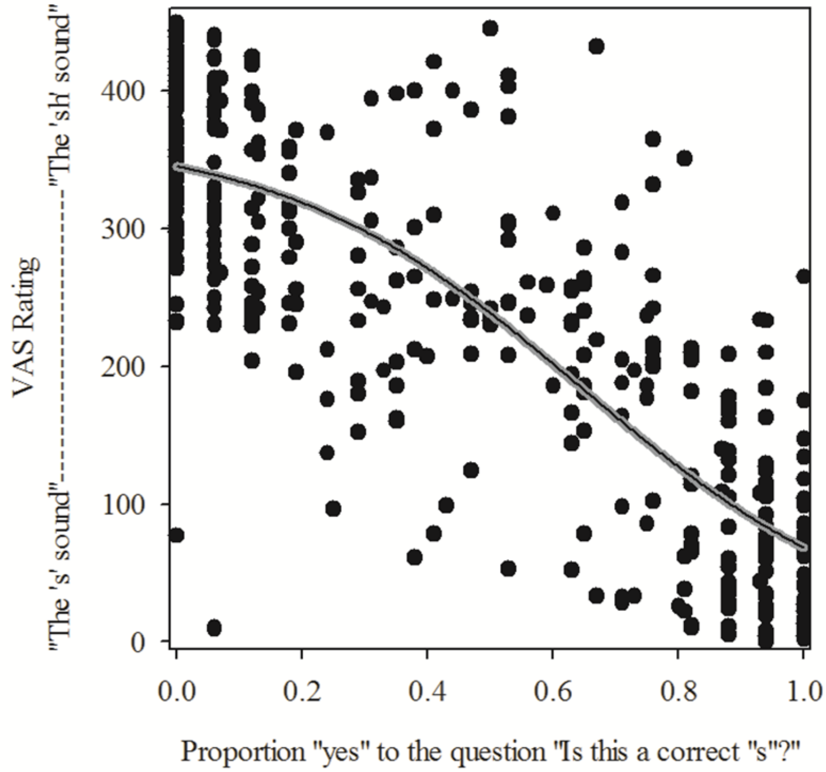


Figure 5. Relationships between VAS click locations and the proportion of "yes" responses to the question, "Is this a correct 's'?" for one listener.



Clinicians and researchers are not always interested in studying the acquisition of contrasts between two sounds, like /s/ and /ʃ/. A VAS display like that in Figure 3 simply isn't adaptable to measuring children's acquisition of a three-way contrast, like the contrast among /s/, /ʃ/, and /θ/ in a child who produces all three as /t/. Urberg-Carlson et al. (2009) circumvented this problem by asking listeners to first categorize sounds using a forced-choice task ("Did the child say 's' or 'sh'?"), then provide a goodness rating using a VAS anchored by the text "Perfect" and "Bad." For example, if someone identified a child's production as /s/, he/she would then be asked to rate how good an example of /s/ it was, using the VAS. Urberg-Carlson et al. examined listeners' ratings for the same /s/ and /ʃ/ stimuli that are shown in Figures 4 and 5. Again, listeners' VAS goodness ratings correlated well with both the acoustic characteristics of the sounds and with the proportion of times that the stimuli were rated as correct /s/ by a separate group of listeners. One representative listener's data are shown in Figures 6 and 7. These figures show the goodness ratings for the sounds judged to be /s/ (circles, solid black regression line) and those judged to /ʃ/ (stars, dashed gray regression line) plotted against the centroid frequency of the fricatives (Figure 6) or the proportion of times listeners in Li et al. judged them to be /s/ (Figure 7). As Figure 7 shows, goodness ratings were predicted by the typicality of their acoustic characteristics and by the proportion of times the stimuli were judged to be /s/.

Figure 6. Relationships between VAS goodness-rating click locations and centroid frequencies for one listener.

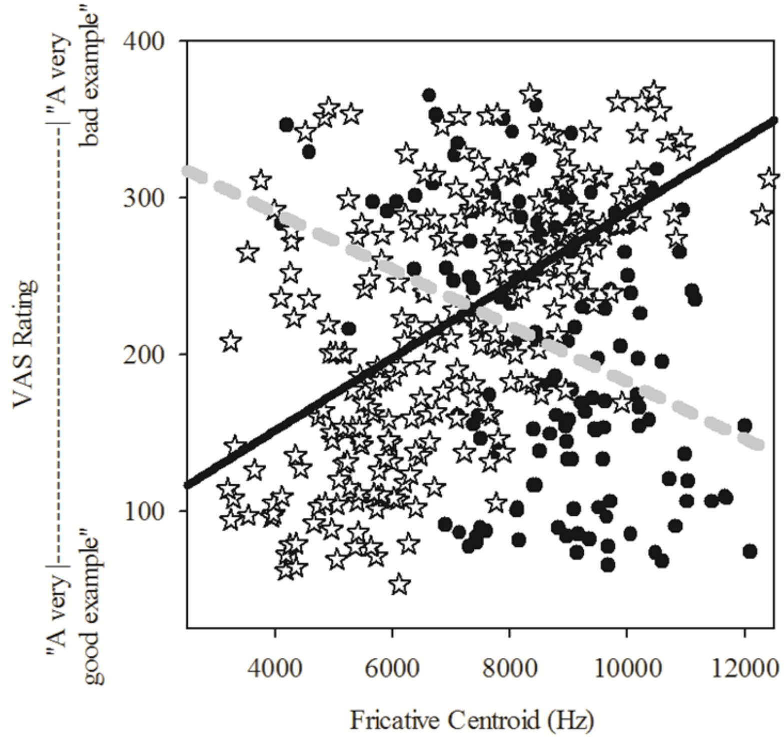
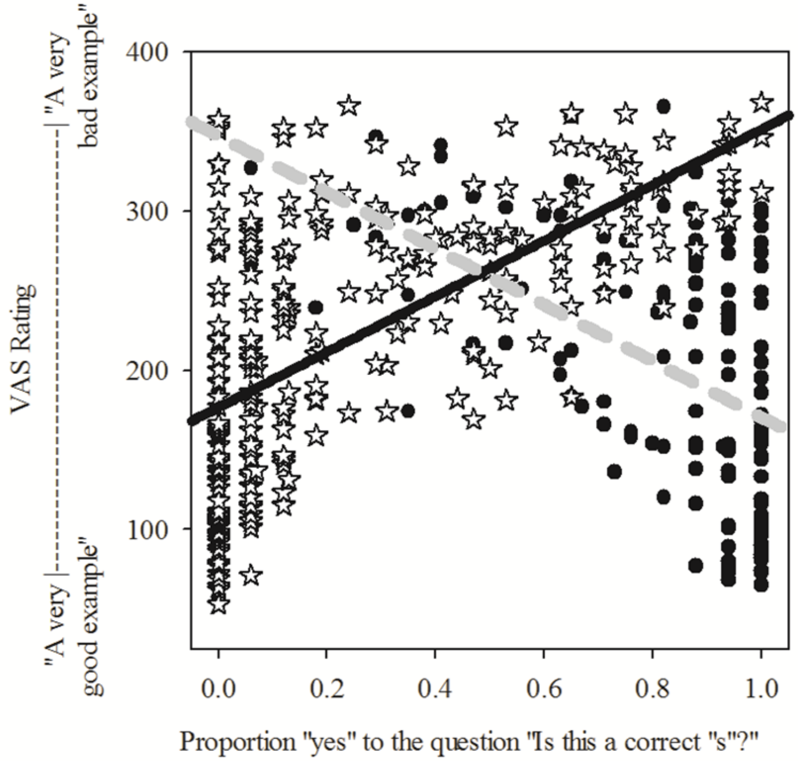


Figure 7. Relationships between VAS goodness-rating click locations and the proportion of "yes" responses to the questions "Is this a correct 's'?" for one listener.



Implementing VAS in the Clinic: A Proposal

Our studies on VAS suggest that it can be a useful tool for measuring gradual change in children's speech over the course of therapy. When a clinician works on a particular contrast or contrasts, like the /t/–/k/ and /d/–/g/ contrasts that must be learned by children with a fronting pattern, the unidimensional scales like that in Figure 3 can be used. When children are working on multiple sounds—as in the case of a child learning to produce /s/, /ʃ/, and /θ/ simultaneously—clinicians can categorize productions and then give a VAS judgment of how good a production it was. In our research, we elicited VAS ratings using special software that measures click locations automatically and reports them in pixels, which we could easily convert to the scales seen in Figures 4 through 7. A clinician could easily adapt this method by simply making paper-and-pencil measures and physically measuring the distance between the ratings and the endpoints of the scale. This is similar to the method used in the CAPE-V, in which distances along a 10 cm line are used.

The information presented in this article clearly documents the benefits of incorporating VAS in assessments of speech production. Clinicians can better determine whether sounds transcribed identically are truly identical, which would allow them to better select among different types of therapies and potentially to prioritize some targets over others. Given that VAS can show small changes with a category like /s/ or /ʃ/, clinicians who use VAS to measure children's progress in therapy might also be more likely to detect small changes during the course of therapy than would those who use phonetic transcription alone.. The clinical utility of VAS remains to be demonstrated empirically, and we are actively working to develop clinical research studies on this very topic. We encourage clinicians who choose to incorporate VAS into their clinical practice to share their experiences with us, so that together we can move our field forward in the important endeavor of improving the services we provide to our clients.

Acknowledgments

The stimuli used in the studies described in this article were supported by NIH grant DC02932 to Jan Edwards. The research developing VAS for the assessment of children's speech was supported by NSF grant BCS 0729277 to Benjamin Munson. That grant was part of the larger project using machine learning to model the interplay of production dynamics and perception dynamics in phonological acquisition. I thank the other principal investigators on that grant, Jan Edwards and Mary E. Beckman, for their important input in this endeavor. I gratefully acknowledge Fangfang Li for her work analyzing the acoustic characteristics of the fricatives used in the experiments described in this study and Eunjong Kong for her work analyzing the stop consonants.

References

Arbisi-Kelm, T., Edwards, J., Munson, B., & Kong, E.-J. (2010). Cross-linguistic perception of velar and alveolar obstruents: A perceptual and psychoacoustic study. Poster presentation at the Acoustical Society of America. Available in *Journal of the Acoustical Society of America, 127*, 1957.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International, 5*, 341–345.

Dean, E. C., Howell, J., Waters, D., & Reid, J. (1995). Metaphon: A metalinguistic approach to the treatment of phonological disorder in children. *Clinical Linguistics and Phonetics, 9*, 1–19.

Edwards, J., & Beckman, M. E. (2008). Methodological questions in studying phonological acquisition. *Clinical Linguistics and Phonetics, 12*, 937–956.

Gibbon, F. (1999). Undifferentiated lingual gestures in children with articulation/phonological disorders. *Journal of Speech, Language, and Hearing Research, 42*, 382–397.

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America, 108*, 1252–1263.

Kempster, G. B., Gerratt, B. R., Abbott, K. V., Barkmeier-Kraemer, J., & Hillman, R. E. (2009). Consensus Auditory-Perceptual Evaluation of Voice: Development of a standardized clinical protocol. *American Journal of Speech-Language Pathology, 18*, 124–132.

Li, F. (2008). *The phonetic development of voiceless sibilant fricatives in English, Japanese, and Mandarin Chinese* (Doctoral dissertation). Department of Linguistics, Ohio State University, Columbus, OH.

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

Li, F., Munson, B., Edwards, J., Yoneyama, K., & Hall, K. C. (2011). 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, 129*, 999–1011.

McDonald, E. T. (1964). Articulation testing and treatment: A sensory-motor approach. Pittsburgh, PA: Stanwix House.

Munson, B., Edwards, J., Schellinger, S. K., Beckman , M. E., & Meyer, M. K. (2010). Deconstructing phonetic transcription: Covert contrast, perceptual bias, and an extraterrestrial view of vox humana. *Clinical Linguistics and Phonetics, 24*, 245–260.

Munson, B., Johnson, J. M., & Edwards, J. (2011). *The role of clinical experience in speech-language pathologists' perception of subphonemic detail in children's speech.* Manuscript under consideration for publication.

Smit, A., Hand, L., Freilinger, J., Bernthal, J. E., & Bird, A. (1990). The Iowa articulation norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders, 55*, 779–798.

Tyler, A. A., Figurski, G. R., & Langsdale, 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., Kaiser, E., & Munson, B. (2008, November). *Assessment of children's speech production 2: Testing gradient measures of children's productions*. Poster presented at the 2008 ASHA Convention, Chicago, IL.

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. Available in *Journal of the Acoustical Society of America, 125*, 2529.

Whalen, D. (1991). Subcategorical phonetic mismatches and lexical access. *Perception and Psychophysics, 50*, 351–360.