CHAPTER 5

Children's Speech Sound Disorders: An Acoustic Perspective

RAYMOND D. KENT, LUCIANA PAGAN-NEVES, KATHERINE C. HUSTAD, AND HAYDEE FISZBEIN WERTZNER

Introduction

Speech sound disorders are almost always identified and described by their auditory-perceptual properties, as determined by adult listeners. Given that auditory-perceptual properties are extracted from the acoustic signal of speech, it follows that acoustic methods should be highly suited to the study of these disorders. But the capability for doing something is not the same as the necessity or even desirability for doing so. What does acoustic analysis offer for the assessment, treatment, and understanding of developmental speech disorders? This chapter takes the view that acoustic analysis is a valuable complement and co-referent to perceptual analysis. The advantages that acoustic analysis offers to the understanding of children's speech sound disorders are primarily objectivity, quantification, and sensitivity. Each of these advantages is discussed, with examples from the literature. Also included are suggestions to (a) improve the acoustic analysis of children's speech, and (b) apply acoustic methods to clinical assessment and treatment.

An Envisioned Future

A hopeful view of the future application of acoustic analysis to the clinical assessment and treatment of speech disorders includes the routine use of computer-based methods to record, display, analyze, and store information about speech sound patterns (also see Chapter 6). In fact, these functions have been available for some time, so this view of the future is not especially bold or revolutionary. But the operative word is "routine." Despite the general availability of computer-based acoustic analysis at relatively low cost, the use of such analysis tools is by no means routine. Are there real prospects for routine clinical application? And what needs to be done to bring these prospects to reality? This chapter addresses these questions and, in so doing, reviews major accomplishments in the acoustic analysis of speech disorders. The emphasis is on speech disorders in children, but occasional reference is made to disorders in adults as they help to reveal potential clinical tools for children's speech.

The pivotal technology is digital signal processing, which enables a user to record samples of speech as a digital file, display this file as a waveform or other pattern, select and edit parts of the saved file, conduct various types of analysis (e.g., waveform, spectrogram, spectrum, fundamental-frequency contour, intensity envelope, some of which are shown in Figure 5–1), play all or selected parts of the file, and save the results of analysis. The basic methods are found in several different systems that are available commercially at varying costs (Ingram, Bunta, & Ingram, 2004; Read, Buder, & Kent, 1990) or as free downloads (such as the computer analysis program *Praat* [Dutch for *talk*] developed by Paul Boersma and David Wee-

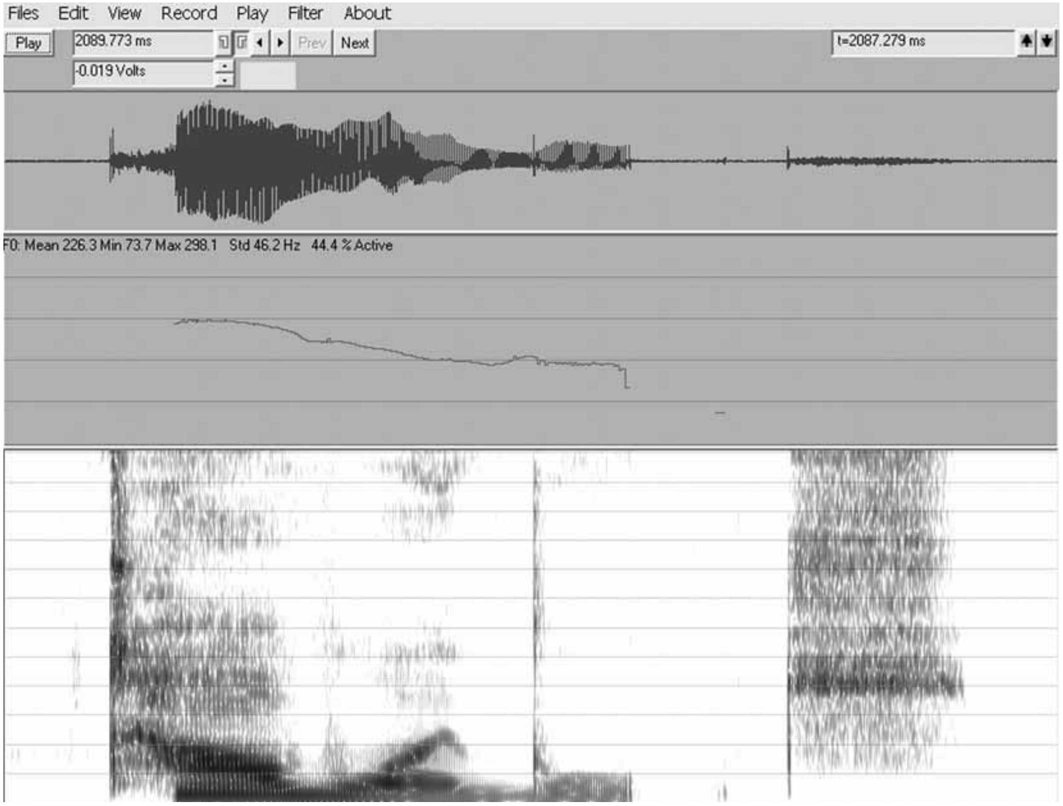


Figure 5–1. Screen display of a waveform (Panel A), pitch trace (Panel B), and spectrogram (Panel C) in TF32. Speech sample is from a three-year-old typically developing child producing the phrase "cowboy boots."

nink at the University of Amsterdam). Certainly, cost is not an obstacle to performing fairly sophisticated operations in the acoustic analysis and synthesis of speech. Perhaps a greater obstacle is a limited understanding of how these analyses can be used in the practice of speech-language pathology. Relevant discussions are available in articles and books explaining how these digital methods can be applied to speech and language disorders (Ingram et al., 2004; Kent & Read, 2002). Proficiency with acoustic methods may be the single most important factor that will lead to the increased use of these methods for clinical purposes. Acoustic analysis, like any laboratory tool, requires practice in its use (see Figure 5–1).

Speech Is More Than What Meets the Ear

The human auditory system is remarkable in its ability to segregate the speech signal from noise and to achieve a phonetic interpretation of that signal. The robustness of this process necessarily discards a fair amount of detail. A primary advantage of acoustic analysis is that it permits the detection of acoustic properties that may not be detected by auditory means. The ear is necessarily an informational filter that attends to certain aspects of sound and ignores others. Especially because of biases introduced through phonetic experience with a given language, human audition discards or neglects much of the acoustic signal of speech. Speech-language clinicians are taught to listen carefully to acoustic variations that the layperson may not hear at all. A recent study comparing the perception of correct and incorrect Brazilian Portuguese liquids /l/, /ɾ/, /ʎ/ by undergraduate and graduate students demonstrated a better performance of the undergraduate students, indicating that sometimes expert listeners can be more influenced by their knowledge of the pathology than by what they are really listening to (Pagan & Wertzner, 2007a). Furthermore, even the experts' ears fail in comparison to acoustic analysis, which is free of phonetic biases and other influences that inevitably affect perception. This conclusion has been reached for several aspects of speech (Kent, 1996).

Toward a Pediatric Speech Science

Most of the literature on acoustic theory, methods of analysis, and acoustic databases pertains to normal adult speech, and especially to the speech of men. Gradually, theory, methods, and databases are becoming more comprehensive to include women and children, that is, the community of speakers. The analysis of children's speech, in particular, needs to take account of various factors that can complicate the analysis task. Some of the major factors are as follows:

1. Because children have shorter vocal tracts than adults, children's speech sounds (both vowels and consonants) have energy at higher frequencies than those observed for adults. One consequence is that the total frequency range of analysis may need to be extended for satisfactory results with children's speech. For example, the spectral energy associated with infants' fricatives may reach as high as 16 kHz (Kent & Read, 2002). Fortunately, most contemporary systems for recording and analyzing speech permit a total bandwidth of about 20 kHz. Increases in computer memory accommodate such extended bandwidths.

2. The precision of formant estimation varies with fundamental frequency (F_0). Voices with high F_0 (generally the case for children) are more challenging when it comes to estimating formant frequencies. The limitation is basically one of sampling. With higher F_0 values, the harmonics of the laryngeal source are farther apart, and this makes it more difficult to estimate the formant locations in the spectrum (Huggins, 1980: Kent & Read, 2002; Vorperian & Kent, 2007).
3. Children often are variable in their phonatory patterns, which may include transient or long-term features such as breathiness, roughness, pitch shift, and even register change (e.g., between chest and vocal fry registers). In contrast, adults tend to have fairly uniform phonatory patterns so that one set of analysis parameters generally is suitable for an entire utterance.
4. Velopharyngeal function may differ between children and adults. Although the precise maturational pattern is not well established, it appears that typically developing children may achieve speech-adequate control at about the same time as canonical babbling appears (Thom, Hoit, Hixon, & Smith, 2006). However, some children may show variable or unusual patterns of velopharyngeal function, which can complicate acoustic analysis.
5. Eccentric or idiosyncratic acoustic-phonetic patterns may appear. Because children are learning language, including its phonological and phonetic aspects, at the same time they are learning the motor skills of speech, they may exhibit behaviors that are seldom, if ever, observed in adult speech. Some of these behaviors may be highly transient, but others may persist over a substantial period of time.
6. The development of the vocal tract reflects a complicated interaction of the growth of its constituent structures, and this interaction is poorly understood (Kent & Hustad, 2009; Kent & Tilkens, 2007; Kent & Vorperian, 1995).
7. The acoustic database for children's speech is incomplete. The database is growing slowly, but it is not adequate for all purposes. Clinical interpretation depends critically on a secure knowledge of normative behavior.

These comments are not intended to discourage the use of acoustic analysis, but rather to forewarn those who attempt these analyses of the complications that lie in the path of discovery and application. Similar precautions could be issued on the use of phonetic transcription and physiologic analyses. These difficulties notwithstanding, there is no good reason why acoustics should not be a working partner with auditory-perceptual methods in the understanding of children's speech disorders.

Prosodic Patterns

Depending on the definition that is used, prosody can embrace a number of phenomena including intonation, tempo (pause and lengthening), vocal effort, and loudness. These are suprasegmental aspects of speech, meaning that their effects typically extend over two or more phonetic segments. It is not possible to offer an extensive review of prosody in this chapter, and the emphasis is on the tractability of an acoustic analysis of prosody in children. In one view of prosody that was designed expressly for application to language development (Gerken & McGregor, 1998), prosody was conceptualized as three general types of phenomena in language: phrasal stress, boundary cues, and meter. Each of these is elaborated in the following.

Phrasal Stress

Phrasal stress is the phenomenon of word prominence in a phrase. Stress is conveyed by adjustments of duration, fundamental frequency, and intensity. Children begin to regulate the acoustic cues of stress (fundamental frequency, amplitude, duration) as early as 18 to 30 months of age (Kehoe, Stoel-Gammon, & Buder, 1995). In a study of linguistic stress produced by 5 children with suspected developmental apraxia of speech (sDAS) and 5 children with phonological disorder, Munson, Bjorum, and Windsor (2003) reported that the children with sDAS were judged to be less successful than the children with phonological disorder in producing target stress contours. However, acoustic studies showed that the children with sDAS produced acoustic differences between stressed and unstressed syllables that apparently were not consistently detected by the listeners who made the stress judgments.

Boundary Cues

Boundary cues are pauses, adjustments in duration, or variations in pitch that mark the ends of language units. A well-known example of a boundary cue is phrase-final lengthening, in which a word or syllable that precedes the end of a major syntactic unit is lengthened. Phrase-final lengthening often is accompanied by a falling tone, and the two of these features are effective cues for a major constituent unit. Figure 5–2 illustrates both final syllable lengthening and falling tone. They also appear relatively early in speech-language development (Snow, 1994) and are robust in the face of speech or language disorder (Snow, 1998; Wang, Kent, Duffy, & Thomas, 2005). According to Snow's (1994) data on children aged 16 to 25 months, intonation is acquired earlier than final syllable timing. As Snow pointed out, one implication of this result is that final lengthening is a learned prosodic feature.

Meter or Rhythm

Meter (or rhythm) is the pattern of stressed and unstressed syllables for words and phrases. In American English, syllables usually have a strong-weak (SW) alternation, and this alternation defines the rhythm of the language. The SW pattern is linked to a stress unit called the foot, which is a SW syllable pair. Low, Grabe, and Nolan (2000) introduced a measure called the Pairwise Variability Index (PVI), which seems to be a useful measure of a speaker's adherence to the normal stressed-unstressed alternation in English. PVI is an index of changes in successive vowel length over an utterance, and it is not affected by speaking rate. It is computed as follows:

$$\mathrm{PVI} = 100 \times [\sum |(d_k - d_{k-1})/d_k + d_{k-1})/2| \;/\; (m-1)]$$

where m equals the number of vowels (or syllables) in an utterance and d is the duration of the k^{th} vowel (syllable).

PVL has only recently been applied to the study of speech disorders (Henrich et al., 2006; Wang, Kent, Duffy, Thomas, & Fredericks, 2006), and, to our knowledge, has not been applied to the study of typical speech development.

The main conclusion is that acoustic correlates exist for prosodic constituents,

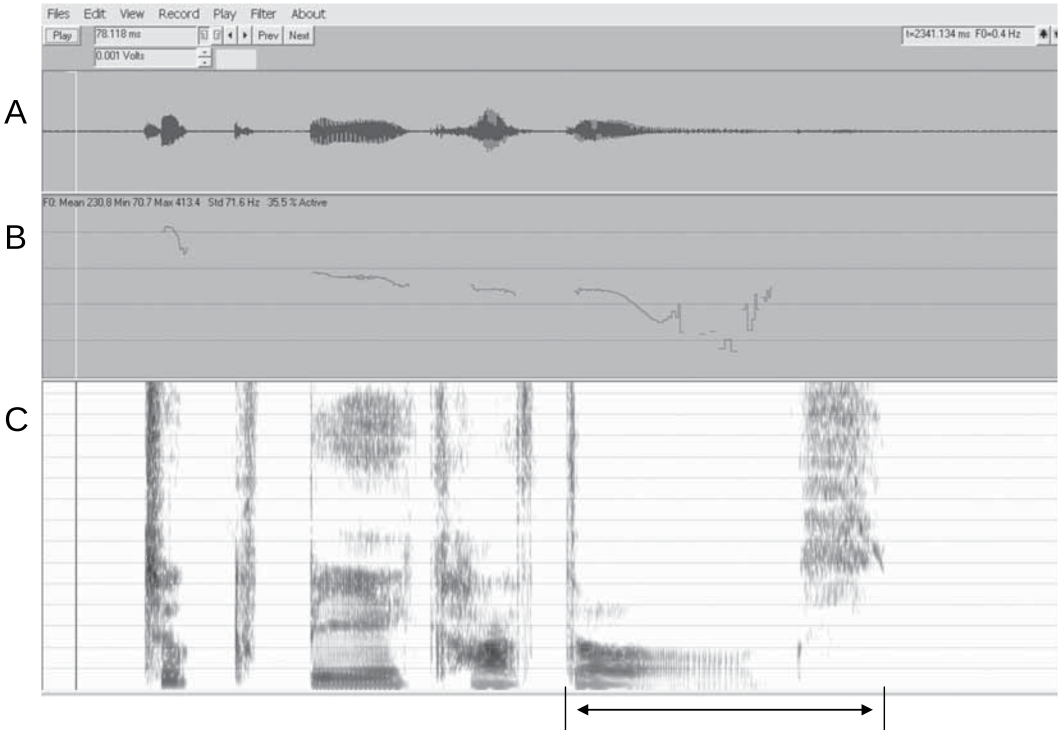


Figure 5–2. Screen display of a waveform (Panel A), pitch trace (Panel B), and spectrogram (Panel C) in TF32. Speech sample is from a three-year-old typically developing child producing the phrase "cook big hot dogs." Note the falling intonation pattern shown on the pitch trace (Panel B) and the syllable-final lengthening indicated by the arrow on the spectrogram (Panel C).

and these correlates are appropriate means for the study of speech and language development in children, or for disorders in development.

Segmental Analysis

Temporal Patterns

The temporal pattern of speech is determined by multiple influences, ranging from prosodic patterns (considered in the previous section) to intrinsic segment durations (Klatt, 1976). As children gain language proficiency and motor skill, the temporal pattern of their speech increasingly conforms to the adult standard in the language. At the segmental level, temporal measurement applies to the intrinsic duration of phonetic elements or to the effects of the immediate phonetic context. A number of generalizations have been established, including the following: (1) short or lax vowels have a briefer duration than long vowels; and (2) vowels preceding voiced consonants are longer than vowels before voiceless consonants. Other generalizations apply to segments in clusters or in word-sized units: (1) a singleton consonant has a longer duration than the same sound in a consonant cluster; (2) the base form of a word has a shorter duration as prefixes or suffixes are combined with it; and (3) new or novel words are produced with a longer duration than familiar words. These are regularities of American English, and children learn to incorporate them in their speech patterns. Their developmental appearance has clinical relevance. For example, Schwartz (1995) concluded that word familiarity is associated with shorter word duration, and he explained this outcome as evidence of word-specific motor maturation. An implication is that word duration can be used as a clinical index of familiarity or motor maturation.

Munson examined the mean duration of /s/ frication, and its variability in adults and in three groups of children (mean ages of 3;11; 5;04; and 8;04). Children had a larger temporal variability than adults. Weismer and Elbert (1982) studied the temporal characteristics of /s/ production in normally speaking adults, normal speaking children, and children with /s/ misarticulations. The /s/ durations of the misarticulating children were significantly more variable than those for the other two groups. This result was explained in terms of differences in speech motor control capabilities. It appears that temporal variability reflects both maturation and disorder (or perhaps only a single factor if it can be shown than disorder is equivalent to delayed maturation)

Figure 5–3 gives a comparison of typical and atypical (disordered) productions of a simple phrase. The atypical production is noticeably longer, with lengthening of phonetic segments and phrases.

One of the most frequently studied temporal features is the voicing contrast for word-initial stop consonants. These sounds are associated with a sequence of acoustic events, including a transient or burst (a pulse of energy that occurs with the initial release of the constriction), a frication interval (a period of turbulence noise generated as the constriction is progressively opened), and onset of voicing (the initiation of vocal fold vibration for the following vowel). An interval of aspiration typically occurs between the frication and the onset of voicing, so that word-initial voiceless stops in English are aspirated. The interval between the burst and the onset of voicing is called the voice onset time (VOT). VOT has a range of values that are often classified as voicing lead or prevoicing (voicing begins before the stop is released), simultaneous voicing (onset of voicing is simultaneous with the transient), short lag (onset of voicing begins shortly after the onset of voicing), and long lag (onset of voicing begins significantly after the onset of voicing. In short, VOT is a continuous variable on which various phonetic categories of voicing can be mapped, and these vary across languages. Perceptual studies have shown that listeners are generally oblivious to small differences within a voicing category. For example, a short-lag VOT of 5 msec cannot be distinguished from a short-lag VOT of 15 msec. As young children learn to control the production of VOT, they often begin with a preference for prevoicing or short-lag. Adults will tend to perceive both of these as voiced stops in American English. Macken and Barton (1980) reported that children produced small differences in VOT for voiced and voiceless cognates that were not perceived by adults. In an acoustic study of phonologically disordered children, Catts and Jensen (1983) concluded that some phonologically disordered children may have less mature speech timing control. A recent study with Brazilian Portuguese-speaking children aged between 6 and 10 years old (Gurgueira, 2006) demonstrated that voiced stops are always produced with prevoicing, which is also true

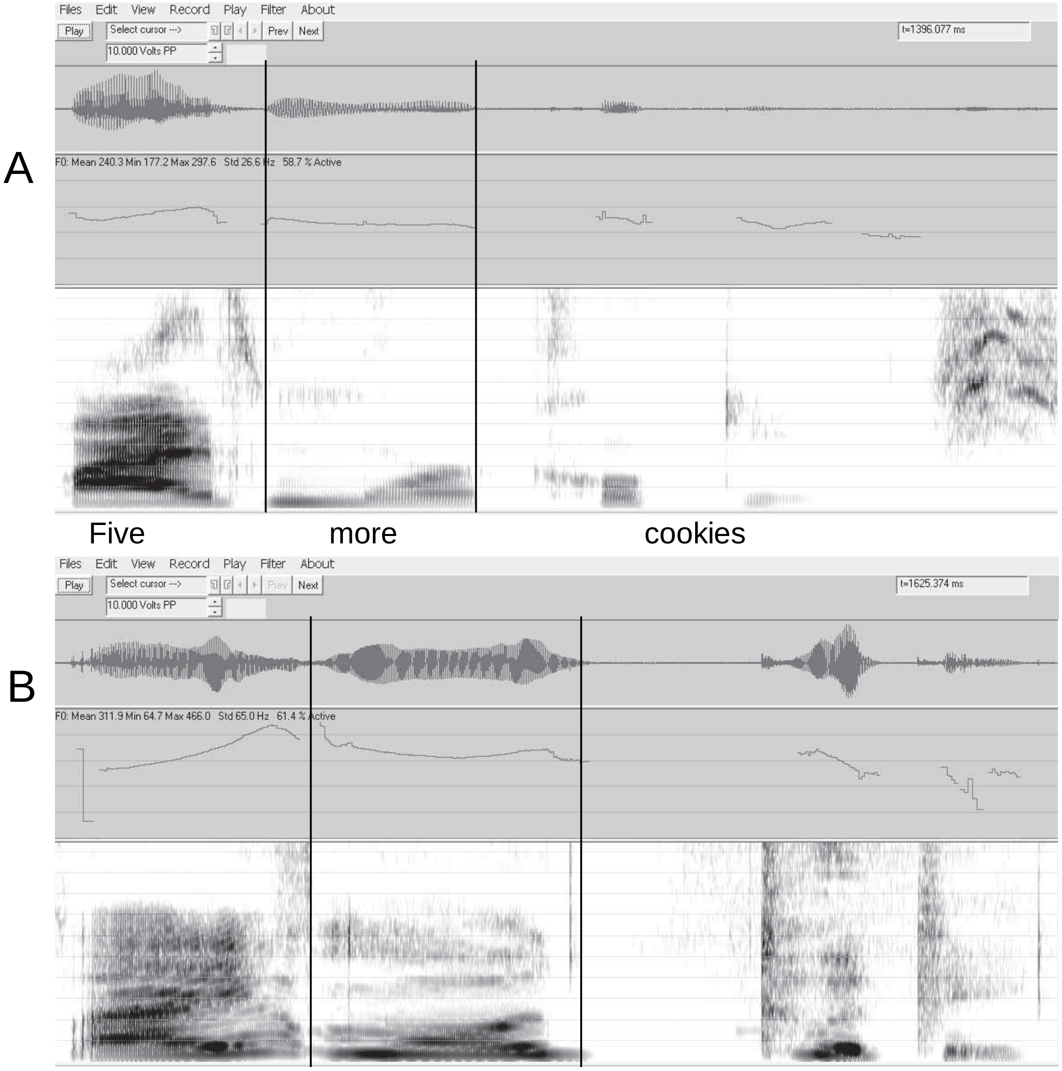


Figure 5–3. Screen displays of waveform, pitch trace, and spectrogram in TF32. Panel A shows the speech of a 5-year-old boy who is typically developing. Panel B shows the speech of a 5-year-old boy with apraxia of speech and mild dysarthria. Both boys are producing the phrase "five more cookies." Note the overall duration difference for the two productions and the increased length of individual words and pauses for the child with the speech disorder (Panel B).

for Spanish, Italian, and French (Borden, Harris, & Raphael, 1994).

The differences in VOT that can be registered by acoustic means could have implications for treatment. Tyler, Edwards, and Saxman (1990) used both phonological and acoustic analyses to describe the speech of four children with a phonological disorder. The acoustic analyses indicated that three of the children produced significant, although frequently imperceptible, differences in VOT for a given stop when it represented different stops in adult speech. These small differences can be taken as evidence of productive phonological knowledge, and it was shown that such knowledge facilitated rapid generalization of correct production of the treated contrast. But when such knowledge was not evident in acoustic analysis, treatment over a longer period was needed to achieve production accuracy on the same treated contrast. But it should be noted that the voicing contrast can be based on several cues, not VOT alone. Forrest and Rockman (1988) suggested that a matrix of acoustic cues is needed to explain the perception of word-initial voicing in the speech of phonologically disordered children. In addition to VOT, these cues include fundamental frequency and F1 frequencies at the onset of voicing, and the amplitude of the burst and aspiration relative to the amplitude of the vowel onset.

Spectral Patterns

Formant descriptions (typically F1-F2 or F1-F2-F3, where Fn is a formant) are low-dimensional descriptions of vowel sounds. One advantage of a formant specification is that a fairly systematic relationship holds between formant pattern and vowel articulation (i.e., the acoustic-to-articulatory conversion). In the classic F1-F2 formant plot, the F1 and F2 frequencies are related principally to tongue height and advancement, respectively. Alternatively, the F2-F1 difference can be interpreted as tongue advancement/retraction. Formant patterns are readily observed in spectrograms or spectra and are among the most salient acoustic properties of speech.

The size of the vowel space, as typically displayed in an F1/F2 plot, is a potential index of the capacity for intelligible speech. Data on the acoustic vowel space in typically developing children have been summarized by Vorperian and Kent (2007). Data for children with speech disorders have been reported for several conditions including dysarthria (Higgins & Hodge, 2001; Liu, Tsao, & Kuhl, 2005), hearing loss (Kent, Osberger, Netsell, & Hustedde, 1987; Liker, Mildner, & Sindija, 2007; Rvachew, Slawinski, Williams, & Green, 1996; Schenk, Baumgartner, & Hamzavi, 2003), and various developmental disorders (Moura et al., 2008). Unusually small areas of the acoustic vowel space are correlated with reduced intelligibility, but it should be noted that some speakers maintain a fairly high level of intelligibility even with a compressed vowel space, so long as other acoustic cues are preserved. Furthermore, vowel-specific formant-frequency differences may have value in characterizing the vocal tract features of particular syndromes (Moura et al., 2008).

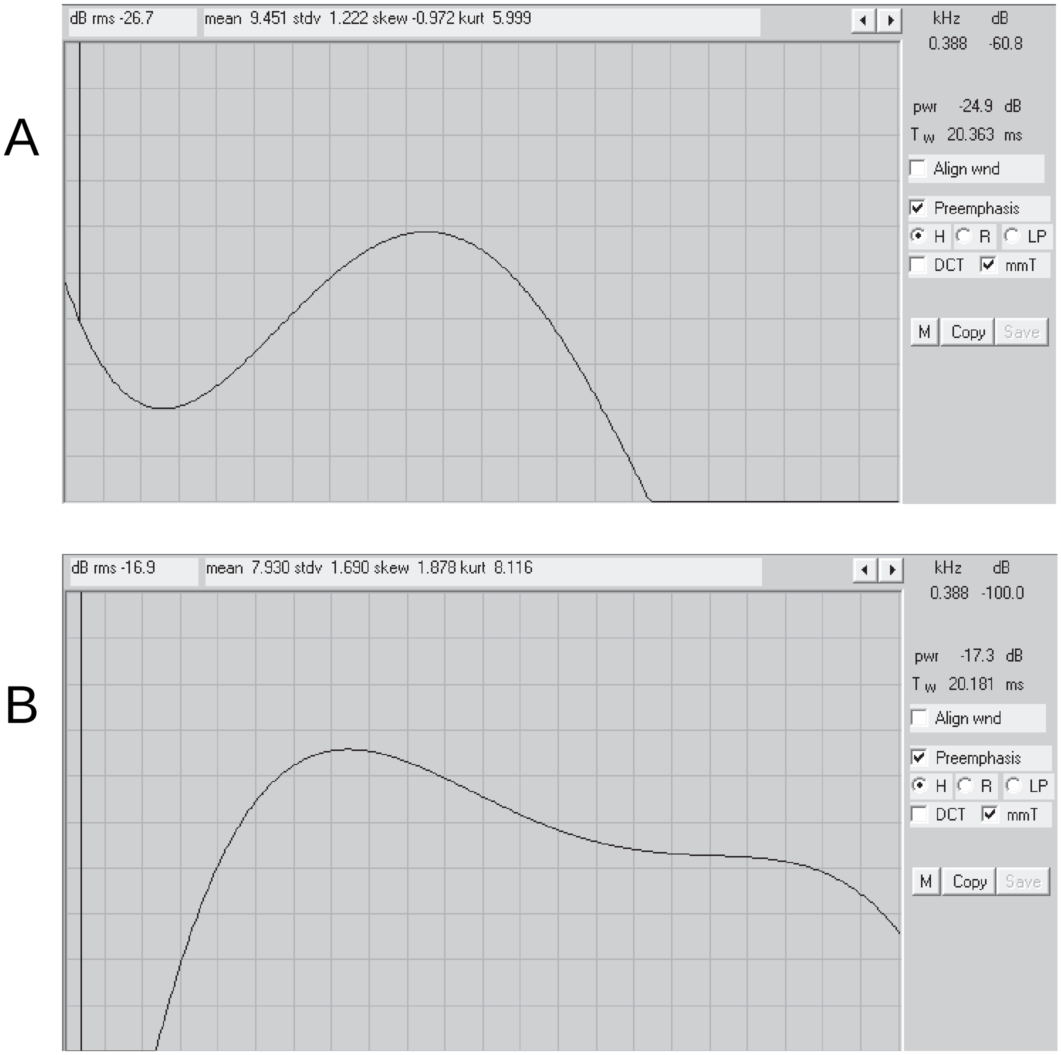
Among the most important noise events in speech are the bursts associated with stops and the frication intervals associated with fricatives and affricates. Generally, noise events in speech are characterized by diffuse spectra that possess varying degrees of resonant shaping. Without question, these events carry a great deal of phonetic information. What is less certain is how these acoustic intervals should be characterized. A valuable source of normative data for adults is the article by Jongman, Wayland, & Wong (2000). Some possibilities for the analysis of children's fricative sounds are considered next.

The earliest analyses used spectrograms and spectral analyses to characterize the noise energy in various /s/ distortions (Daniloff, Wilcox, & Stephens, 1986). One outcome of this work was recognition of the large inter- and intraspeaker variability for children who misarticulated the /s/ sound.

Daniloff et al. concluded that /s/ has a wide range of permissible acoustic allophonic variants, and that this sound accommodates a considerable variation in the upper and lower cutoff frequencies of the major noise energy, and the frequency and amplitude of major spectral peaks. An implication of this conclusion is that it may not be worthwhile to focus on fine spectral details for clinical purposes, but rather to emphasize major regions of noise energy. Taking together the results of the Daniloff et al. (1986), Weismer and Elbert (1980), and Munson (2004) studies reviewed earlier, it appears that both temporal and spectral variability are to be expected in children's misarticulated /s/. The variability is at once an interesting feature of misarticulated speech and a challenge to researchers and clinicians who would examine this sound.

Spectral moments were introduced as a speech analysis method by Forrest, Weismer, Milenkovic, and Dougall (1988) who treated FFTs as random probability distributions for which the first four moments (mean, variance, skewness, and kurtosis) were computed. The first spectral moment is the mean or center of gravity of the spectrum. The second moment is the distribution of energy around the mean, typically expressed as the variance or standard deviation. The third moment is skewness, which may appear as the degree of spectral tilt (although its exact meaning depends on the overall shape of the spectrum). The fourth moment is kurtosis, which is often defined as the degree of peakedness of the distribution or spectrum. Figure 5–4 illustrates the use of spectral moments for characterizing two fricatives produced by a child. It should be noted that these descriptions are most valid when the underlying distribution has the shape of the normal probability distribution. In fact, acoustic spectra rarely have that shape. The four moments are not uniform in their value in characterizing noise spectra, and a major goal of the ensuing discussion is to identify the moments that hold particular value in spectral description.

Spectral moments have been used to describe fricatives in typically and atypically developing speech. Normative data on /s/ production were reported for 26 children aged 9 to 15 years by Flipsen, Shriberg, Weismer, Karlsson, and McSweeny (1999). It was concluded that /s/ can be characterized satisfactorily by data for the midpoint of the /s/ frication presented in a linear scale (as opposed to the Bark scale), with preference for the 1st and 3rd spectral moments. In addition, the authors noted that the data should be referenced to individual linguistic-phonetic contexts. Rather different conclusions were reached by Nissen and Fox (2005), in a study of adults and children aged 3 to 6 years. Their results indicated that spectral slope and variance, usually neglected in earlier studies of child speech, contributed importantly to the differentiation and classification of the voiceless fricatives. The only measure that separated all for places of fricative articulation was spectral variance. Interestingly, it was also reported that /s/ and /ʃ/ were distinguished more sharply by adults than by children, with a remarkable change in several spectral parameters occurring at about 5 years of age. Munson (2004) compared spectral variability in /s/ production for adults and three groups of children (mean ages of 3;11; 5:04; and 8:04). Spectral variability was defined as changes in the spectral mean (first spectral moment) through the interval of frication noise. Adults produced the /s/ with less variability than the children's groups, who did not differ from one another. In view of the lack of effects of phonetic context on spectral variability, Munson concluded that that the differences between adults and children reflected a "subtle variability in place



A: dB rms -26.7 | mean 9.451 stdv 1.222 skew -0.972 kurt 5.999 | kHz 0.388 dB -60.8 | pwr -24.9 dB | T w 20.363 ms | ☐ Align wnd | ☑ Preemphasis | ● H ○ R ○ LP | ☐ DCT ☑ mmT

B: dB rms -16.9 | mean 7.930 stdv 1.690 skew 1.878 kurt 8.116 | kHz 0.388 dB -100.0 | pwr -17.3 dB | T w 20.181 ms | ☐ Align wnd | ☑ Preemphasis | ● H ○ R ○ LP | ☐ DCT ☑ mmT

Figure 5–4. Spectral moments display from a 4-year-old child, using a 20-msec window at the temporal midpoint of /s/ (Panel A) and /sh/ (Panel B). Note the difference in skewness between the two productions.

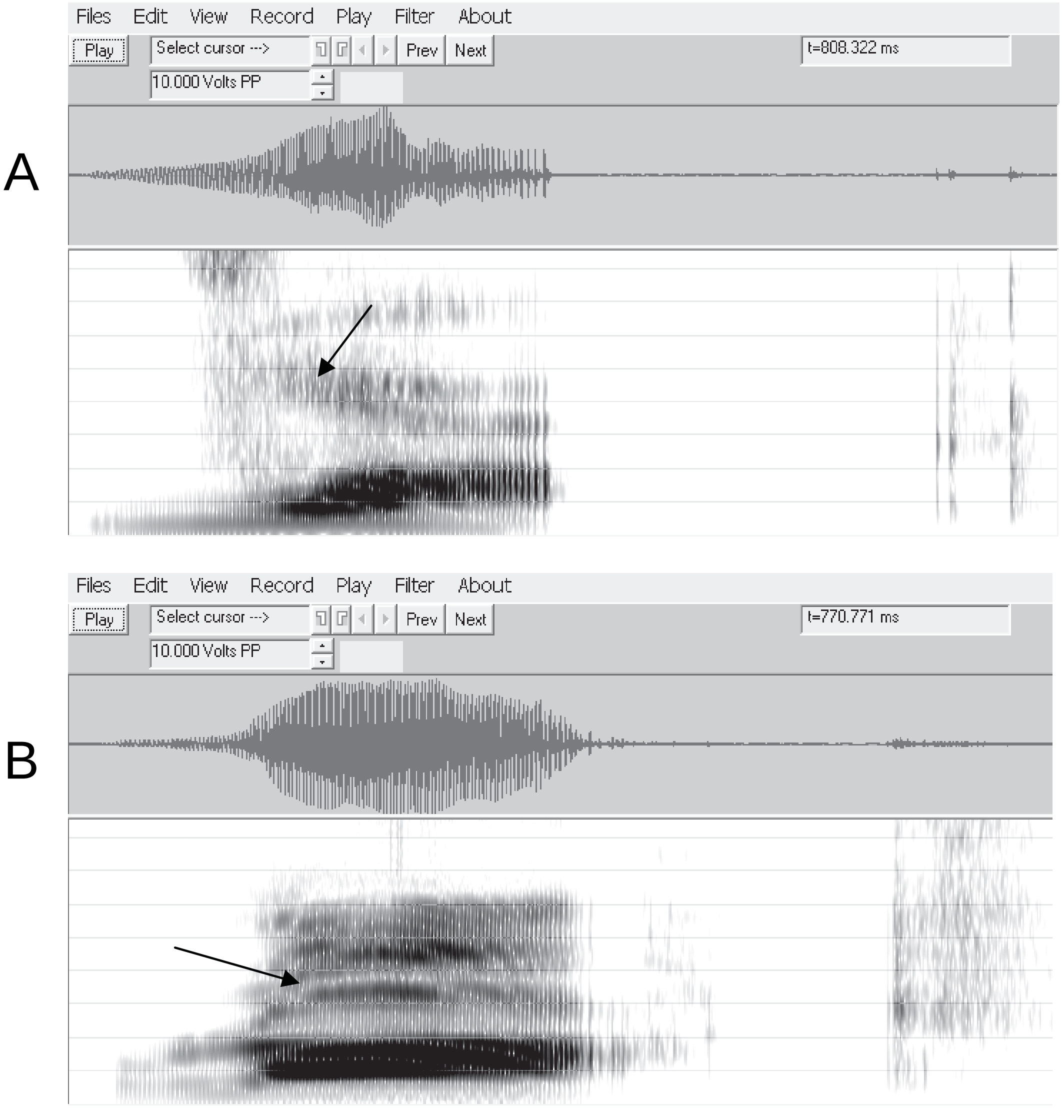
of articulation for /s/ in the children's productions (p. 58). It should also be noted that children's speech may differ from adults' speech in respect to the relative amplitude of its high-frequency components. Short-term spectra of children's speech sounds have been reported to have reduced amplitudes for /s/ and /ʃ/ and for vowel energy above kHz compared to the same speech sounds produced by adults (Pittman, Stelmachowicz, Lewis, & Hoover, 2003). These differences in relative amplitude are highly relevant to understanding the perception and transcription of children's speech.

Spectral analyses also have been reported for the burst of stop consonants, especially the voiceless stops /p t k/. In one of the earliest studies, spectral moments were calculated for word-initial /t/ and /k/ produced by both typically developing children and by children with phonological disorder (Forrest, Weismer, Hodge, Dinnsen, & Elbert, 1990). Using a discriminant function analysis, Forrest et al. achieved 82% correct classification of the two stops using the first, third and fourth moments. The discriminant function developed for the normally speaking children was applied to the phonologically disordered children, no distinction could be made between /t/ and /k/. Bunnell, Polikoff, and McNicholas (2004) compared spectral moments and Bark cepstral analyses for classification of children's word-initial voiceless stops. A better classification rate was achieved for the Bark cepstral analysis. For both types of analysis, four time frames that sampled the initial 40 msec of each burst was needed for the highest rates of correct classification. It is premature to recommend either spectral analysis or Bark cepstral analysis as the preferred method for inspecting stop bursts.

An example that demonstrates both clinical application and the sensitivity of acoustic analysis is for a common type of speech sound error in early speech development, omission or deletion of a segment. This is usually a conspicuous error, readily perceived by adult listeners. However, Weismer (1984) reported that in some cases of an ostensible deletion, acoustic analyses showed that the supposedly deleted consonant had formant transitions appropriate to its phonetic properties. The acoustic cue was not detected by listeners. In another study of apparent omission of word-final stops (Weismer, Dinnsen, & Elbert, 1981), it was shown that two of three children with the omission pattern produced vowel duration differences that were suited to the voicing characteristic of the omitted stop (i.e., longer vowels before voiced stops). Apparently, these two children preserved the stop-voicing feature in their speech, even though listeners judged the stop to be deleted. According to data reported by Krause (1982), the vowel duration cue for voicing appears at least by the age of 3 years. She described the early pattern of development as involving both exaggerated vowel lengthening (before voiced stops) and exaggerated vowel shortening (before voiceless stops).

Spectrotemporal Patterns for Liquids and Glides

The liquids in American English are the lateral /l/ and the rhotic /r/, both of which can be problematic for children acquiring speech. Acoustically, liquids are characterized especially by their formant pattern (/r/) or formant-antiformant pattern (/l/). Acoustic analyses for /r/ are illustrated in Figure 5–5. The glides in American English are the palatal /j/ and the labiovelar /w/ (and its voiceless allophone, which may not be used by all speakers. The glides are associated with a relatively gradual formant transition into the following vowel. Acoustic data on correctly produced /w, r, l/ in both children and adults were reported by Dalston (1975). Chaney (1988) studied three groups of children: a group that correctly produced /w, r, l, j/, a group with developmental w/r and w/l substitutions, and a group of articulation-impaired children who had w/r and w/l substitutions. The children with /w, r/ errors produced the glide /j/ with acoustic properties similar to those seen in the control group, but neither of the groups with errors differentiated among /w, r, l/ by either formant frequencies or transition rate. Interestingly, the /w/ produced for target /w/ and in substitution for /r/ and /l/ by some of the children with errors did not



10.000 Volts PP

t=808.322 ms

10.000 Volts PP

t=770.771 ms

Figure 5–5. Screen display of a waveform and spectrogram in TF32. Panel A shows a production of the word "rock" from a three-year-old typically developing child, where the /r/ phoneme is distorted. Panel B shows a production of the same word by a 6-year-old typically developing child, where the /r/ phoneme is produced appropriately. Note the difference in the 3rd formant frequency (as indicated by the arrows in each panel).

match the acoustic pattern of /w/ as produced by the children without errors. Shuster (1996) showed how speech resynthesis based on linear prediction coding (LPC) can be used to modify disordered productions of /r/ so that they approximate correct productions. Flipsen et al. (2001) suggested that speech-genetics research would be enhanced by the availability of acoustic phenotypes, such as residual distortions of rhotic sounds.

Pagan and Wertzner (2007b) studied the acoustic patterns of two of the three Brazilian Portuguese liquids (/r/, a voiced alveolar tap, and /l/, a voiced alveolar lateral) as produced by typically developing children and by children with a phonological disorder who had r/l substitutions. It was found that the /l/ produced as a substitution for /r/ was different from both the /l/ correctly produced by the phonologically disordered children and the /l/ produced by the control group. The /l/ substituting for /r/ had a longer duration, different steady-state values, and a smaller formant slope. This result is another example of an acoustic differentiation for sounds that are judged to be the same by listeners

Coarticulation

Coarticulation, or the simultaneous adjustment of the articulators to two or more phones, is a basic characteristic of competent adult speech. In forward or anticipatory coarticulation, a phonetic property of a given phone is assumed earlier in the phonetic string. For example, lip rounding for the vowel in the word appears during the initial consonant /s/. In backward or retentive coarticulation, a phonetic property of a given phone is retained to a later position in the phonetic string. An example is the nasalization of the vowel in the word *no*. The development of coarticulation is not well understood, and rather different conclusions have been reached from research on children's speech. In several studies, young children were observed to show more extensive coarticulation than adults (Nittrouer, Studdert-Kennedy, & McGowan, 1989; Nittrouer, Studdert-Kennedy, & Neely, 1996; Siren & Wilcox, 1995), whereas other studies showed no developmental difference, variable patterns across sounds, or greater coarticulation in adults than children (Flege, 1988; Katz, Kripke, & Tallal, 1991; Kent, 1983; Kuijpers, 1993; Repp, 1986; Sereno, Baum, Marean, & Lieberman, 1987; Sussman, et al., 1996; Turnbaugh, Hoffman, Daniloff, & Absher, 1985). The different results probably can be explained by reference to the different methods of analysis, and the phonetic properties of the speech material interacting with the maturational status of the child. It appears that there is no single maturational pattern of coarticulation for various sounds, and that a child seeks to balance coarticulatory adjustments against contrastive distinctiveness (Gibson & Ohde, 2007; Nittrouer, 1993; Sussman, Duder, Dalston, & Cacciatore, 1999; Sussman, Hoemeke, & McCaffrey, 1992).

Acoustic Correlates of Speaker Intelligibility

A long-term goal in the application of acoustics is to determine the acoustic correlates of intelligibility (see also Chapter 10). Research on this topic is hindered by the potentially large number of acoustic features that can be considered, and also by the fact that speakers can deploy acoustic cues in various combinations to achieve a satisfactory degree of intelligibility. It seems safe to conclude on the basis of available evidence that the same general acoustic properties are relevant to both adult and child speech (Hazen & Markham, 2004).

Research on "clear" versus "conversational" speech holds value in understanding the acoustic bases of speech intelligibility (Picheny et al., 1985, 1986, 1989). Acoustic analyses of the two forms of speech have shown consistent differences, thereby laying a foundation for a general understanding of the acoustic correlates of intelligibility (Picheny et al., 1985, 1986, 1989). As compared to "clear" speech, "conversational" speech tends to have modified or reduced vowels, nonreleased word-final stops, and reduced intensities for obstruents. Although "clear" speech typically is slower than "conversational" speech, it is important to note that enhancements of intelligibility can be achieved even at rapid speaking rates (Krause & Braida, 1995). The acoustic differences between "clear" and "conversational" speech may explain intrinsic intelligibility differences among individual speakers. Bond and Moore (1994) studied the acoustic-phonetic differences between a talker with relatively high intelligibility and two talkers with relatively low intelligibility. The high-intelligibility talker had many acoustic-phonetic properties similar to those described for "clear" speech. In a similar study of individual differences in intelligibility, Bradlow, Torretta, & Pisoni (1996) concluded that global characteristics (e.g., speaking rate and mean F_0 level) did not correlate strongly with intelligibility, but the fine-grained characteristics (F_0 and F_1 variation, formant frequency range for vowels, intersegmental timing) did correlate. The profile of a highly intelligible speaker was one who produced sentences with a relatively wide range of F_0, a relatively expanded vowel space that includes a substantial F_1 variation, precise articulation of the point vowels, and a high precision of intersegmental timing. Therefore, there is an important linkage between two general approaches to the study of intelligibility differences in normal speakers.

Similar results can be seen in studies of dysarthric and deaf speakers that have established fine-grained acoustic characteristics relating to differences in speaker intelligibility (Kent et al., 1989; Metz et al., 1985; Monsen, 1976; Weismer & Martin, 1992). In the main, the results from dysarthric and deaf speech agree with the results reviewed above for normal speech. That is, the differences in intelligibility appear to be rooted in a common set of fine-grained acoustic measures including vowel formant frequencies and intersegmental timing.

If these results can be generalized to speech development and to developmental speech disorders, then the implication is that fine-grained acoustic properties are the key to understanding differences in speech intelligibility.

Variability as an Index of Precision and Maturation

As earlier sections of this chapter make clear, variability has been a particular focus of research on both typical and atypical speech development. Across motor skills, it is generally presumed that increasing accuracy is a characteristic of skill maturation. One way of gauging accuracy is to determine the variability in a motor response, or, in the case of speech, the acoustic consequences of that motor response. In one of the earliest studies to address this issue, Eguchi and Hirsh (1969) reported that there were nearly continuous decreases in the variability of both F1 and F2 frequencies from 3 to 11 years of age in typically developing children. One interpretation is that motor skill for speech improves with age, and acoustic measures of formant structure reflect this improvement up until the age of puberty. But other acoustic data point to a different conclusion. Nittrouer (1993) reported that the variability in F1 frequency was minimal by the age of 3 years whereas variability in F2 frequency continued to decrease beyond that age. The early accuracy in F1 frequency was related to an early maturation of jaw movement control, given that jaw movement has a strong effect on F1 frequency. In fact, the maturation of motor control over different oral structures is open to discussion. Although it has been reported that children's jaw movements are less variable than lip movements (Green, Moore, & Reilly, 2002; Walsh & Smith, 2002), it also has been shown that there are parallel decreases in the variability of jaw and lip movements with maturation (Walsh & Smith, 2002).

Variability in the temporal patterns of speech also has been examined. The general conclusion is that variability declines with age until late childhood, puberty, or adolescence (Kent & Forner, 1978; Lee, Potamianos, & Narayanan, 1999; Lehman & Sharf, 1989; Munson, 2004; Smith & Kenney, 1999). However, changes in variability are not necessarily uniform across different segments (Kent & Forner, 1978; Smith & Kenney, 1999).

Variability is actually relevant to several developmental issues, including the following:

(a) Estimates of the variability of either spectral or temporal features have been proposed as an index of the maturation of speech motor control, as noted above. Variability, commonly expressed as a standard deviation, is considered as an estimate of precision of articulation. This approach requires analysis of multiple tokens of a given speech target. It is assumed that the speaker is able to create a stable representation of the target behavior from which motor commands to the articulators can be formulated.

(b) In general, the variability in temporal segments is related to speaking rate, such that a slow rate is associated with greater variability. Because children typically have a slower rate of speech than adults, speaking rate is confounded with the maturational factor mentioned in (a) above (Kent & Forner, 1978). Children with a speech-language disorder may have even slower speaking rates than typically developing children. This slow rate may be related to the combined effects of development and disorder.

(c) Variability may be a gauge of category breadth or coarticulatory range. For example, as children add elements to their vowel systems, the allowable range for any one vowel may be adjusted to accommodate the insertion of new vowel sounds. Similarly, variability in producing a particular word may be related to the lexical density for that word. Presumably, a word with a high neighborhood density would be produced more accurately than a word with a low neighborhood density.

(d) Variability in a spectral or temporal feature (or a spectrotemporal property) may be an indication of destabilizing forces. In a dynamic systems perspective, periods of destabilization may be optimum times for intervention.

(e) Measures of temporal pattern are not as sensitive to age and gender variables as are measures of formant or general spectral pattern. Reliability estimates of temporal measures are reviewed in Kent and Read (2002).

Obviously, a simple interpretation of variability is not likely to be correct unless this list can be pruned to one or two applicable alternatives. Unfortunately, many developmental studies were not designed to address each of these factors in an empirical fashion that allows their separation.

The clinical implication is not necessarily that a clinician will record ten or more tokens of a sound pattern and then calculate standard deviations for a selected measurement. Such a procedure may be forbiddingly tedious for both the child and the clinician. Rather, the object is more likely to be to ascertain the stability of production in relation to a clinical objective. Say, for example, that the objective is treatment of a speech sound disorder. Frequently, clinicians want to establish a degree of stability in production of a certain sound pattern before introducing a change of some kind, such as working on another target sound, changing the phonetic context of sound production, or varying prosodic features such as speaking rate or stress.

Sensitivity

Acoustic analysis is capable of resolving fine differences in the timing and spectra of speech sounds. Differences that are not detected by the ear can be detected by suitable acoustic analyses that are performed in the time domain (waveform), frequency domain (FFT or LPC spectrum, cepstrum, or other analysis), or the time-frequency domain (spectrogram or other running spectral display). The issue here is not necessarily quantification, as important as that may be, but identifying the shear presence or absence of an acoustic property. Examples from the literature are discussed below to illustrate the concept for both segmental and suprasegmental aspects of speech.

The sensitivity of acoustic analysis does not necessarily depend on quantification. Sometimes, simply observing the presence or absence of an acoustic phenomenon is sufficient. In some examples given earlier in this chapter, measurements were not always needed. Rather, the person performing the analysis used acoustics as a kind of alternative visual display—a highly sensitive one—to the analysis performed by the ear. This approach made it possible to detect (a) acoustic differences between stressed and unstressed syllables that were not consistently perceived by adults (Munson, Bjorum, & Windsor, 2003), (b) small differences in VOT even for stimuli that were not distinguished by adult listeners (Macken & Barton, 1980; Tyler, Edwards, & Saxman (1990), and (c) acoustic evidence of a phonetic feature of a speech sound that was supposedly omitted (Weismer, 1984; Weismer, Dinnsen, & Elbert, 1981).

This "look and listen" strategy can be quite powerful, as it enables the observer (e.g., clinician, researcher) to observe the visual display of speech and to reconcile it with what is heard. Qualitative analysis has much to recommend it. As noted by Liss and Weismer (1992), "traditional acoustic measures of temporal and spectral characteristics of normal speech may not necessarily reveal the inherently 'important' aspects of disordered speech production" (p. 2984). This is not to assert that quantitative analyses are irrelevant to the study of disordered speech, but rather to say that qualitative analyses are a valuable complement to quantitative methods. For additional discussion of this issue, see Weismer and Liss (1991).

Each individual clinician must ask herself or himself whether acoustic tools will make for better clinical services. Technology is only as useful as the use to which it is put. The dramatic progress in speech technology (automatic speech recognition, speech synthesis, no-cost or low-cost speech analysis software) presents a powerful set of tools for the future practice of speech-language pathology.

Conclusion

The first author was a long-term faculty colleague of Larry Shriberg at the University of Wisconsin-Madison. We co-authored a text, *Clinical Phonetics*, now in its third edition.

In preparing the text and the accompanying audiotapes, we accomplished a kind of mutual calibration of our "phonetic" ears as we listened (repeatedly) to samples of children's speech disorders. Our labors began in the pre-DSP days, which meant that speech samples existed physically as pieces of audiotape (analog recordings). I recall seeing strips of tape hanging around the room where we worked. These were eventually assembled by tape-splicing methods into tapes for auditory exercises for phonetic transcription. If we undertook that effort today, it would be very different. We would use digital signal processing to record, store, and analyze the samples. Rather than compare notes strictly on our respective auditory impressions (which differed now and then) of each sample, we would examine visual displays of acoustic information. Would this information be helpful? I have no doubt that it would.

References

Boersma, P., & Weenink, D. (2008). PRAAT (computer program). University of Amsterdam; available at http://www.fon.hum.uva.nl/praat .

Bond, Z. S., & Moore, T. J. (1994). A note on the acoustic-phonetic characteristics of inadvertently clear speech. *Speech Communication*, *14*, 325–337.

Borden, S. J., Harris, K. S., & Raphael, L. J. (1994). *Speech science primer: Physiology, acoustics and perception of speech.* Baltimore: Williams & Wilkins.

Bradlow, A. R., Torretta, G. M., & Pisoni, D. B. (1996). Intelligibility of normal speech. I. Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication*, *20*, 255–272.

Bunnell, H. T., Polikoff, J., & McNicholas, J. (2004). *Spectral moment vs. Bark cepstral analysis of children's word-initial voiceless stops.* Available at http://www.asel.udel.edu/speech/reports/icslp04/ICSLP04_paper.pdf .

Catts, H. W., & Jensen, P. J. (1983). Speech timing of phonologically disordered children: Voicing contrast of initial and final stop consonants. *Journal of Speech and Hearing Research*, *26*, 501–510.

Chaney, C. (1988). Acoustic analysis of correct and misarticulated semivowels. *Journal of Speech and Hearing Research*, *31*, 275–287.

Dalston, R. M. (1975). Acoustic characteristics of English /w, r, l/ spoken correctly by young children and adults. *Journal of the Acoustical Society of America*, *57*, 462–469.

Daniloff, R. G., Wilcox, K., & Stephens, M. I. (1980). An acoustic-articulatory description of children's defective /s/ productions. *Journal of Communication Disorders*, *13*, 347–363.

Eguchi, S., & Hirsh, I. J. (1969). Development of speech sounds in children. *Acta Otolaryngologica, Supplementum*, *257*, 1–51.

Flege, J. E. (1988). Anticipatory and carry-over nasal coarticulation in the speech of children and adults. *Journal of Speech and Hearing Research*, *31*, 525–536.

Flipsen, P., Jr., Shriberg, L., Weismer, G., Karlsson, H., & McSweeny, J. (1999). Acoustic characteristics of /s/ in adolescents. *Journal of Speech, Language, and Hearing Research*, *42*, 663–677.

Flipsen, P., Jr., Shriberg, L., Weismer, G., Karlsson, H., & McSweeny, J. (2001). Acoustic phenotypes for speech-genetics studies: Reference data for residual /Er/ distortion. *Clinical Linguistics and Phonetics*, *15*, 603–630.

Forrest, K., & Rockman, B. K. (1988). Acoustic and perceptual analysis of word-initial stop consonants in phonologically disordered children. *Journal of Speech, Language, 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.

Forrest, K., Weismer, G., Milenkovic, P., & Dougal, R. N. (1988). Statistical analysis of word-initial voiceless obstruents: Preliminary data.

Journal of the Acoustical Society of America, *84*, 115–123.

Gerken, L., & McGregor, K. (1998). An overview of prosody and its role in normal and disordered child language. *American Journal of Speech-Language Pathology*, *7*, 38–48.

Gibson, T., & Ohde, R. N. (2007). F2 locus equations: Phonetic descriptors of coarticulation in 17- to 22-month-old children. *Journal of Speech, Language, and Hearing Research*, *50*, 97–108.

Green, J. R., Moore, C. A., & Reilly, K. U. (2002). The sequential development of jaw and lip control in speech. *Journal of Speech, Language, and Hearing Research*, *45*, 66–79.

Gurgueira, A. L. (2006). *Estudo acústico dos fonemas surdos e sonoros do português do Brasil, em crianças com distúrbio fonológico apresentando o processo fonológico de ensurdecimento*. [Acoustic study of the voice onset time (vot) and the vowel duration for the distinctiion of voicing of the stop sounds of the brazilian portuguese in children with typical development and with phonological disorders]. Doctoral dissertation. Department of Linguistics of the Faculty of Philosophy, Sciences and Literature and Languages from the University of São Paulo, Brazil.

Hazen, V., & Markham, D. (2004). Acoustic-phonetic correlates of talker intelligibility for adults and children. *Journal of the Acoustical Society of America*, *116*, 3108–3118.

Henrich, J., Lowit, A., Schalling, E., & Mennen, I. (2006). Rhythmic disturbances in ataxic dysarthria: A comparison of different measures and speech tasks. *Journal of Medical Speech-Language Pathology*, *14*, 291–296.

Higgins, C. M., & Hodge, M. M. (2001). F2/F1 vowel quadrilateral area in young children with and without dysarthria. *Canadian Acoustics*, *29*, 66–68.

Huggins, A. W. F. (1980). Better spectrograms from children's speech: A research note. *Journal of Speech and Hearing Research*, *23*, 19–27.

Ingram, K., Bunta, F., & Ingram, D. (2004). Digital data collection and analysis. *Language, Speech, and Hearing Services in Schools*, *35*, 112–121.

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

Katz, W. F., Kripke, C., & Tallal, P. (1991). Anticipatory coarticulation in the speech of adults and young children. *Journal of Speech and Hearing Research*, *34*, 1222–1232.

Kehoe, M., Stoel-Gammon, C., & Buder, E. H. (1995). Acoustic correlates of stress in young children's speech. *Journal of Speech and Hearing Research*, *38*, 338–350.

Kent, R. D. (1976). Anatomical and neuromuscular maturation of the speech mechanism: Evidence from acoustic studies. *Journal of Speech and Hearing Research*, *19*, 421–447.

Kent, R. D. (1983). Segmental organization of speech. In P. F. MacNeilage (Ed.), *The production of speech*. New York: Springer Verlag.

Kent, R. D. (1996). Hearing and believing: Some limits to the auditory-perceptual assessment of speech and voice disorders. *American Journal of Speech-Language Pathology*, *7*, 7–23.

Kent, R. D., & Forner, L. L. (1980). Speech segment durations in sentence recitations by children and adults. *Journal of Phonetics*, *8*, 157–168.

Kent, R. D., & Hustad, K. C. (2009). Speech production, development. In L. Squire, T. Albright, F. Bloom, F. Gage, & N. Spitzer (Eds.), *New encyclopedia of neuroscience.* Oxford, UK: Elsevier.

Kent, R. D., & Murray, A. D. (1982). Acoustic features of infant vocalic utterances. *Journal of the Acoustical Society of America*, *72*, 353–365.

Kent, R. D., Osberger, M. J., Netsell, R., & Hustedde, C. G. (1987). Phonetic development in twins who differ in auditory function. *Journal of Speech and Hearing Disorders*, *52*, 64–75.

Kent, R. D., & Read, C. (2002). *The acoustic analysis of speech* (2nd ed.). Albany, NY: Singular/Thomson Learning.

Kent, R. D., & Tilkens, C. (2007). Oral motor foundations of speech. To appear in S. McLeod

(Ed.), *The international guide to speech acquisition*. Albany, NY: Delmar.

Kent, R. D., & Vorperian, H. K. (1995). Anatomic development of the craniofacial-oral-laryngeal systems: A review. *Journal of Medical Speech-Language Pathology*, *3*, 145–190. (Also published as a monograph (1995) San Diego, CA: Singular.).

Kent, R. D., Weismer, G., Kent, J. F., & Rosenbek, J. C. (1989). Toward phonetic intelligibility testing in dysarthria. *Journal of Speech and Hearing Disorders*, *54*, 482–499.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, *59*, 1208–1221.

Krause, J. C., & Braida, L. D. (1995). The effects of speaking rate on the intelligibility of speech for various speaking modes. *Journal of the Acoustical Society of America*, *98*, 2982.

Krause, S. E. (1982). Developmental use of vowel duration as a cue to postvocalic stop consonant voicing. *Journal of Speech and Hearing Research*, *25*, 388–393.

Kuijpers, C. (1993). Temporal aspects of the voiced-voiceless distinction in speech development of young Dutch children. *Journal of Phonetics*, *21*, 313–327.

Lee, S., Potamianos, A., & Narayanan, S. (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *Journal of the Acoustical Society of America*, *105*, 1455–1468.

Lehman, M. E., & Sharf, D. J. (1989). Perception/production relationships in the development of the vowel duration cue to final consonant voicing. *Journal of Speech and Hearing Research*, *32*, 803–815.

Liker, M., Mildner, V., & Sindija, B. (2007). Acoustic analysis of the speech of children with cochlear implants: A longitudinal study. *Clinical Linguistics and Phonetics*, *21*, 1–11.

Liss, J. M., & Weismer, G. (1992). Qualitative acoustic analysis in the study of motor speech disorders. *Journal of the Acoustical Society of America*, *92*, 2984–2987.

Liu, H. M., Tsao, F. M., & Kuhl, P. K. (2005). The effect of reduced vowel working space on speech intelligibility in Mandarin-speaking young adults with cerebral palsy. *Journal of the Acoustical Society of America*, *117*, 3879–3889.

Low, E. L., Grabe, E., & Nolan, F. (2000). Quantitative characterizations of speech rhythm: Syllable-timing in Singapore English. *Language and Speech*, *43*, 377–401.

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

Metz, D., Samar, V., Schiavetti, N., Sitler, R., & Whitehead, R. (1985). Acoustic dimensions of hearing-impaired speakers' intelligibility. *Journal of Speech and Hearing Research*, *28*, 345–355.

Monsen, R. B. (1976). Normal and reduced phonological space: The productions of English vowels by deaf adolescents. *Journal of Phonetics*, *4*, 189–198.

Moura, C. P., Cunha, L. M., Vilarinho, H., Cunha, M. J., Freitas, D., Palha, M., et al. (2008). Voice parameters in children with Down syndrome. *Journal of Voice*, *22*, 34–42.

Munson, B. (2004). Variability in /s/ production in children and adults: Evidence from dynamic measures of spectral mean. *Journal of Speech, Language, and Hearing Research*, *47*, 58–69.

Munson, B., Bjorum, E. M., & Windsor, J. (2003). Acoustic and perceptual correlates of stress in nonwords produced by children with suspected developmental apraxia of speech and children with phonological disorder. *Journal of Speech, Language, and Hearing Research*, *46*, 189–202.

Nissen, S. L., & Fox, R. A. (2005). Acoustic and spectral characteristics of young children's fricative productions: A developmental perspective. *Journal of the Acoustical Society of America*, *118*, 2570–2578.

Nittrouer, S. (1993). The emergence of mature gestural patterns is not uniform: Evidence from an acoustic study. *Journal of Speech and Hearing Research*, *36*, 959–972.

Nittrouer, S., Studdert-Kennedy, M., & McGowan, R. S. (1989). The emergence of phonetic segments: Evidence from the spectral structure of fricative-vowel syllables spoken by children and adults. *Journal of Speech and Hearing Research*, *32*, 120–132.

Nittrouer, S., Studdert-Kennedy, M., & Neely, S. (1996). Howe children learn to organize their speech gestures: Further evidence from fricative-vowel syllables. *Journal of Speech and Hearing Research*, *39*, 379–389.

Pagan, L. O., & Wertzner, H. F. (2007a). *Perceptual analysis of the three Brazilian Portuguese liquid sounds.* Poster to be presented at the 27th World Congress of the International Association of Logopedics and Phoniatrics, Copenhagen, Denmark.

Pagan, L. O., & Wertzner, H. F. (2007b). *Description of the acoustic characteristics of the liquid sounds /l/ and /r/ in phonologically disordered children.* Poster presented at the 2nd International Association of Logopedics and Phoniatrics Composium, Sao Paulo, Brazil, March 24–25.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech. *Journal of Speech and Hearing Research*, *28*, 96–103.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard of hearing II: Acoustic characteristics of clear and conversational speech. *Journal of Speech and Hearing Research*, *29*, 434–446.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1989). Speaking clearly for the hard of hearing III: An attempt to determine the contribution of speaking rate to difference in intelligibility between clear and conversational speech. *Journal of Speech and Hearing Research*, *32*, 600–603.

Pittman, A. L., Stelmachowicz, P. G., Lewis, D. E., & Hoover, B. M. (2003). Spectral characteristics of speech at the ear. *Journal of Speech, Language, and Hearing Research*, *46*, 649–657.

Read, C., Buder, E., & Kent, R.D. (1990). Speech analysis systems: A survey. *Journal of Speech and Hearing Research*, *33*, 363–374.

Repp, B. (1986). Some observations on the development of anticipatory coarticulation. *Journal of the Acoustical Society of America*, *79*, 1616–1619.

Rvachew, S., Slawinski, E. G., Williams, M., & Green, C. L. (1996). Formant Frequencies of vowels produced by infants with and without early onset ititis media. *Canadian Acoustics/Acoustique Canadienne*, *24*, 19–28.

Schenk, B. S., Baumgartner, W. D., & Hamzavi, J. S. (2003). Changes in vowel quality after cochlear implantation. *ORL Journal of Otorhinolaryngology Related Specialties*, *65*, 184–188.

Schwartz, G. R. (1995). Effect of familiarity on word duration in children's speech: A preliminary investigation. *Journal of Speech and Hearing Research*, *38*, 76–84.

Sereno, J. A., Baum, S. R., Marean, G. C., & Lieberman, P. (1987). Acoustic analyses and perceptual data on anticipatory coarticulation in adults and children. *Journal of the Acoustical Society of America*, *81*, 512–519.

Shuster, L. I. (1996). Linear predictive coding parameter manipulation/synthesis of incorrectly produced /r/. *Journal of Speech and Hearing Research*, *39*, 827–832.

Siren, K. A., & Wilcox, K. A. (1995). Effects of lexical meaning and practiced productions on coarticulation in children's and adults' speech. *Journal of Speech and Hearing Research*, *38*, 351–359.

Smith, B. L., & Kenney, M. K. (1999). A longitudinal study of the development of temporal properties of speech production: Data from 4 children. *Phonetica*, *56*, 73–102.

Snow, D. (1994). Phrase-final syllable lengthening and intonation in early child speech. *Journal of Speech, Language, and Hearing Research*, *37*, 831–840.

Snow, D. (1998). Prosodic markers of syntactic boundaries in the speech of 4-year-old children with normal and disordered language development. *Journal of Speech, Language, and Hearing Research*, *41*, 1158–1170.

Sussman, H. M., Duder, C., Dalston, E., & Cacciatore, A. (1999). An acoustic analysis of the development of CV coarticulation. *Journal of Speech, Language, and Hearing Research*, *42*, 1080–1096.

Sussman, H. M., Hoemeke, K. A., & McCaffrey, H. A. (1992). Locus equations as index of coarticulation for place of articulation distinctions in children. *Journal of Speech and Hearing Research*, *35*, 769–781.

Sussman, H. M., Minifie, F. D., Buder, E. H., Stoel-Gammon, C., & Smith, J. (1996). Consonant-vowel interdependencies in babbling and early words. *Journal of Speech and Hearing Research*, *39*, 424–433.

Thom, S., Hoit, J., Hixon, T., & Smith, A. (2005). Velopharyngeal function during vocalization in infants. *The Cleft Palate-Craniofacial Journal* [published online 15 November; doi: 10.1597/05-113].

Turnbaugh, K., Hoffman, P., Daniloff, R. G., & Absher, R. (1985). Stop-vowel coarticulation in 3-year-olds, 5-year-olds, and adults. *Journal of the Acoustical Society of America*, *77*, 1256–1258.

Tyler, A. A., Edwards, M. L., & Saxman, J. H. (1990). Acoustic validation of phonological knowledge and its relationship to treatment. *Journal of Speech and Hearing Disorders*, *55*, 251–261

Vorperian, H. K., & Kent, R. D. (2007). Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *Journal of Speech, Language, and Hearing Research*, *50*, 1510–1545.

Walsh, B., & Smith, A. (2002). Articulatory movements in adolescents: Evidence for protracted development of speech motor control processes. *Journal of Speech, Language, and Hearing Research*, *45*, 1119–1133.

Wang, Y.-T., Kent, R. D., Duffy, J. R., & Thomas, J. E. (2005). Dysarthria in traumatic brain injury: A breath group and intonational analysis. *Folia Phoniatrica et Logopaedica*, *57*, 59–89.

Wang, Y.-T., Kent, R. D., Duffy, J. R., Thomas, J. E., & Fredericks, G. V. (2006). Dysarthria following cerebellar mutism secondary to resection of a fourth ventricle medulloblastoma: A case study. *Journal of Medical Speech-Language Pathol*ogy, *14*, 109–122.

Weismer, G. 1984). Acoustic analysis strategies for the refinement of phonological analyses. In M. Elbert, D. A. Dinnsen, & G. Weismer (Eds.), *Phonological theory and the misarticulation child* (pp. 30–52). *ASHA Monographs* (No. 22). Rockville, MD: American Speech-Language-Hearing Association.

Weismer, G., Dinnsen, D., & Elbert, M. (1981). A study of the voicing distinction associated with omitted, word-final stops. *Journal of Speech and Hearing Disorders*, *46*, 320–328.

Weismer, G., & Elbert, M. (1982). Temporal characteristics of "functionally" misarticulated /s/ in 4- to 6-years-old children. *Journal of Speech and Hearing Research*, *25*, 275–287.

Weismer, G., & Liss, J. M. (1991). Acoustic/perceptual taxonomies of disordered speech. In C. Moore, K. Yorkston, & D. Beukelman (Eds.), *Dysarthria and apraxia of speech: Perspectives on management* (pp. 245–270). Baltimore: Brookes.

Weismer, G., & Martin, R. (1992). Acoustic and perceptual approaches to the study of intelligibility. In R. Kent (Ed.), *Intelligibility in speech disorders* (pp. 67–118). Philadelphia: John Benjamins.