Acoustic Characteristics of /s/ in Adolescents

Peter Flipsen Jr.*
Lawrence Shriberg
Gary Weismer
Heather Karlsson
Jane McSweeny
Waisman Center
University of Wisconsin–Madison

The goal of the current study was to construct a reference database against which misarticulations of /s/ can be compared. Acoustic data for 26 typically speaking 9- to 15-year-olds were examined to resolve measurement issues in acoustic analyses, including alternative sampling points within the /s/ frication; the informativeness of linear versus Bark transformations of each of the 4 spectral moments of /s/ (Forrest, Weismer, Milenkovic, & Dougall, 1988); and measurement effects associated with linguistic context, age, and sex. Analysis of the reference data set indicates that acoustic characterization of /s/ is appropriately and optimally (a) obtained from the midpoint of /s/, (b) represented in linear scale, (c) reflected in summary statistics for the 1st and 3rd spectral moments, (d) referenced to individual linguistic-phonetic contexts, (e) collapsed across the age range studied, and (f) described individually by sex.

KEY WORDS: acoustic phonetics, articulation, fricatives, phonology, speech disorders

Continued reliance on phonetic transcription is a major measurement constraint in many areas of research in child speech disorders. Limitations of phonetic transcription have been described from several perspectives (e.g., Cucchiarini, 1996; Kent, 1996; McSweeny & Shriberg, 1995; Shriberg & Lof, 1991), as has the potential of acoustic analysis to capture significant information not available in a phonetic transcript (cf. Bond & Wilson, 1980; Forrest, Weismer, Elbert, & Dinnsen, 1994; Hewlett, 1988; Maxwell & Weismer, 1982; Weismer, 1984). Together with information on atypical speech acquisition, discussions in these and other citations suggest that, as an alternative to auditory-perceptual transcription, acoustic data and acoustic-aided phonetic transcription might have the sensitivity and reliability needed to identify and classify speech sound distortions throughout the lifespan (cf. Austin & Shriberg, 1996; Lewis & Shriberg, 1994; Shriberg, 1993; Shriberg & Kwiatkowski, 1994; Shriberg, Kwiatkowski, & Gruber, 1994). The data for /s/ in the present report (and for /r/ in a forthcoming report) were collected as a first step toward a reference database against which to compare theoretically and clinically significant speech sound distortions.

Acoustic Descriptions of /s/

The acoustic characterization of normally articulated /s/ has been a goal of at least 23 studies during the past approximately 4 decades (Avery & Liss, 1996; Bauer & Kent, 1987; Baum & McNutt, 1990; Behrens &

---

*Currently affiliated with Minnesota State University, Mankato

Blumstein, 1988; Boothroyd & Medwetsky, 1992; Daniloff, Wilcox, & Stephens, 1980; Harmes et al., 1984; Hoole, Ziegler, Hartmann, & Hardcastle, 1989; Hughes & Halle, 1956; Katz, Kripke, & Tallal, 1991; McGowan & Nittrouer, 1988; Menon, Jensen, & Dew, 1969; Miccio, Forrest, & Elbert, 1996; Nittrouer, 1995; Nittrouer, Studdert-Kennedy, & McGowan, 1989; Pentz, 1996; Pentz, Gilbert, & Zawadzki, 1979; Schwartz, 1968; Sereno, Baum, Marean, & Lieberman, 1987; Soli, 1981; Strevens, 1960; Tjaden & Turner, 1997; Yeni-Komshian & Soli, 1981). Table 1 comprises descriptive summaries of 21 of the 23 studies. Because the present concern was with the independent characterization of /s/, studies that focused primarily on the coarticulation of /s/ with surrounding segments (e.g., McGowan & Nittrouer, 1988; Soli, 1981) were not included in Table 1 unless they also reported independent data on /s/. Additionally, some studies reported data for both typical and disordered speakers (Baum & McNutt, 1990; Daniloff et al., 1980; Harmes et al., 1984; Miccio et al., 1996; Tjaden & Turner, 1997), in which case only data based on the typical (i.e., normal) speakers were included. Examination of Table 1 indicates a range of methodological differences across the studies. The following observations summarize the primary issues and considerations guiding the design of the study to be reported.

Sample Size

Until the technological advances of the current decade, it had been difficult to derive and manage the large quantities of data generated in acoustic studies. As indicated in Table 1, generalizations about the acoustics of /s/ have been based on as few as 2 speakers, with several studies using sample sizes of fewer than 5 speakers. As the current study is ultimately concerned with classification issues, sample sizes are needed that are sufficient to provide stable estimates of means and standard deviations. Moreover, especially for variables associated with small effect sizes, information is needed on the standard error of measurement (cf. Austin & Shriberg, 1996; Shriberg, Austin, Lewis, McSweeny, & Wilson, 1997).

Age

Although /s/ production has been studied across a wide range of ages, a gap exists in the available empirical data for older children and adolescents. Pentz and colleagues (Pentz, 1996; Pentz et al., 1979) provide data for older children, but there is no available information for 12- to 17-year-old speakers. More generally, because many of the studies in Table 1 do not report age of speakers, it is not known whether these data represent /s/ production across the entire life span. The goals of the present study were to focus on the 9- to 15-year period when presently unknown percentages of children beyond the developmental period (cf. Kent, 1976) have normalized either speech delay or distortions of /s/ and other sibilants.

Sex

The need to examine all speech data individually by sex is motivated by findings such as sex differences in oral cavity size (Daniloff et al., 1980), and proportionally greater individual variability and more aspiration noise in female voices (Klatt & Klatt, 1990). Of the 21 studies in Table 1, eight (38%) did not identify the sex of any of their speakers and another two (10%) identified sex for their adult speakers only. The available data indicated that both sexes were included in 10 (48%) of the studies, but only six (29%) reported data separately by sex, and none of the studies reported data blocked by age and sex. In the six studies that reported data by sex, each reported raw data trends for /s/ to be lower in frequency in males than females.

Linguistic Context

The linguistic contexts represented in the studies of /s/ production in Table 1 include /s/ in isolated prolongations (e.g., Schwartz, 1968), syllables (e.g., Yeni-Komshian & Soli, 1981), words (e.g., Hughes & Halle, 1956), sentences (e.g., Daniloff et al., 1980), spontaneous productions (Bauer & Kent, 1987), and reading passages (e.g., Tjaden & Turner, 1997). It is probable that atypical /s/ productions differ in decontextualized versus propositional speech (e.g., Morrison & Shriberg, 1992), but there is no consensus among investigators, using either auditory-perceptual or acoustic procedures, as to which linguistic context is most appropriate to sample typical speech (cf. Kent, 1996). The present study used short sentence frames and a variety of target contexts for /s/, which controls sentential stress and provides information on potential effects associated with word shape and phonetic context.

Sampling Point

The point in the frication at which /s/ tokens were sampled varied across the 21 studies. In 10 of the 16 studies in which sampling point was reported, there was a trend to fix the analysis to a specific time point (e.g., 100 ms) relative to the onset or offset of the frication for /s/. Such conventions, however, presume a relatively fixed coarticulatory field (defined temporally) both within and across speakers. Other investigators (Avery & Liss, 1996; Behrens & Blumstein, 1988; Boothroyd & Medwetsky, 1992; Harmes et al., 1984; Hoole et al., 1989; Schwartz, 1968) have preferred not to make this assumption and have identified relative temporal points in the frication (e.g., onset, midpoint, and offset) independently for each token. The latter position was taken in the current study.

Table 1 (part 1 of 2). Acoustic studies of typically produced /s/.

Study	Sample size[a]	Speaker ages[b]	Speaking context	Analysis tools (+ algorithm[c])	Sampling point	Sampling rate	Analysis range	Frequency of /s/
Hughes & Halle (1956)	3+ adults[d]	NR	isolated words (initial and final)	Hewlett-Packard Wave Analyzer	NR	–	0–10 kHz	Peak energy above 4 kHz
Strevens (1960)	13 adults (NR)	NR	prolonged in isolation	Kay Sonograph	NR	–	0–8 kHz	Energy range: 3.5–8+ kHz
Schwartz (1968)	9 adults (M) 9 adults (F)	NR	isolation	Kay Sonograph	Middle 500 ms	–	NR	Peaks[e]: M = 5.5 kHz F = 6.5 kHz
Menon et al. (1969)	10 adults (M)	NR	V(s)C syllables (phrase final)	Kay Sonograph	NR	–	NR	Low cutoffs: 2.9–3.7 kHz
Pentz et al. (1979)	10 children (M) 11 children (F)	8;5–11;3	isolated words (initial and final)	Kay 7029A spectrograph	Smaller of midpoint or 30 ms after onset	–	0–16 kHz	Peaks: 8.4 kHz
Daniloff et al. (1980)	2 children (NR)	6.5–8.5	words in sentences	Voiceprint 700 spectrograph	NR	–	0–8 kHz	Peaks: 6.2–7.2 kHz
Yeni-Komshian & Soli (1981)	1 adult (M) 1 adult (F)	NR	(s)VCV syllables in isolation	PDP-12 (LPC)	Last 150 ms	20 kHz	0–10 kHz	Peaks: M = 5.5 kHz F = 6.7 kHz
Harmes et al. (1984)	2 adults (M) 2 adults (F)	49–73 yrs[f]	isolated words (varied contexts)	PDP 11/34 (LPC)	Midpoint	20 kHz	0–8.5 kHz	Centroids: 5.6–7.0 kHz
Bauer & Kent (1987)	24 infants (NR)	3–12 months	Spontaneous fricatives and trills	Kay 7029A spectrograph (LPC & FFT)	NR	40 kHz	0–16 kHz	Peaks: 4.8–13.1 kHz
Sereno et al. (1987)	4 adults (NR) 8 children (NR)	NR 3–7 yrs	(s)V syllables in isolation	PDP-11/34 (LPC)	70 ms before offset	20 kHz	0–9 kHz	Adult peaks: 4.2–4.7 kHz Child peaks: 5.6–6.0 kHz
Behrens & Blumstein (1988)	3 adults (M)	NR	(s)V syllables in isolation	PDP-11/34 (LPC)	first 15 ms mid 15 ms last 15 ms	20 kHz	0–9 kHz	Peak range: 3.8–8.5 kHz
Hoole et al. (1989)	2 adults	NR	(s)V(s)V in sentence frame	LSI-11/73 (FFT)	middle 50 ms	20 kHz	0–9 kHz	Peaks: 5 kHz
Nittrouer et al. (1989)	4 adults (M) 4 adults (F) 8 children (NR) 8 children (NR) 8 children (NR) 8 children (NR)	20–21 yrs 20–21 yrs 3 yrs 4 yrs 5 yrs 7 yrs	(s)VCV syllables in isolation	Kay Digital Spectrograph (DFT)	100 ms before offset	20 kHz	0–9.6 kHz	Centroids: M = 6.8–6.9 kHz F = 7.4–7.8 kHz 3 = 6.9–7.2 kHz 4 = 6.8–7.1 kHz 5 = 6.8–7.2 kHz 7 = 6.7–6.8 kHz
Baum & McNutt (1990)	5 children (NR) 5 children (NR)	6 yrs 8 yrs	(s)V(t) syllables in isolation	BLISS speech analysis system (DFT)	100 ms after onset	20 kHz	0–8.5 kHz	Centroids: 6 = 5.9–6.4 kHz 8 = 6.0–6.7 kHz
Katz et al. (1991)	10 adults (5/5)[g] 10 children (5/5) 10 children (5/5) 10 children (5/5)	26–45 yrs 3 yrs 5 yrs 8 yrs	(s)V syllables (phrase final)	CSpeech (DFT, moments)	100 ms before offset	20 kHz	0–10 kHz	Centroids: A = 6.1–6.8 kHz 3 = 6.2–6.6 kHz 5 = 6.3–6.6 kHz 8 = 6.1–6.4 kHz

Table 1 (part 2 of 2). Acoustic studies of typically produced /s/.

Study	Sample size[a]	Speaker ages[b]	Speaking context	Analysis tools (+ algorithm[c])	Sampling point	Sampling rate	Analysis range	Frequency of /s/
Boothroyd & Medwetsky (1992)	5 adults (M) 5 adults (F)	NR	V(s) syllables from nonsense strings	NR (FFT)	midpoint	20 kHz	0–10 kHz	Low prominent peak: M = 2.9–5.7 kHz F = 5.4–8.9 kHz
Nittrouer (1995)	10 adults (5/5)[g] 10 children (4/6) 10 children (5/5) 10 children (4/6)	20–40 yrs 3 yrs 5 yrs 7 yrs	(s)V syllables (phrase medial)	CSpeech (FFT, moments)	100 ms before offset	25 kHz	0–12.5 kHz	Centroids[e]: M = 6.2 kHz F = 8.0 kHz 3 = 8.0 kHz 5 = 8.1 kHz 7 = 7.9 kHz
Miccio et al. (1996)	3 children (NR) 3 children (NR)	4 yrs 5 yrs	(s)V syllables (phrase medial)	CSpeech (FFT, moments)	first 40 ms	22 kHz	0–9 kHz	Peaks[e]: 4 = 5.8 kHz 5 = 4.2 kHz
Pentz (1996)	NR adults (F)[h] NR children (NR) NR children (NR) NR children (NR)	NR 7 yrs 9 yrs 11 yrs	isolated words (initial and final)	CSL 4300 (FFT, moments)	100 ms before offset	40 kHz	0–20 kHz[i]	Means: A = 8.7 kHz 7 = 8.1 kHz 9 = 8.4 kHz 11 = 8.3 kHz
Avery & Liss (1996)	8 adults (M)	19–48 yrs[j]	words from reading passage (varied contexts)	CSpeech (FFT, moments)	Midpoint	22 kHz	NR	Centroids[k]: LMS: 5.2–6.5 kHz MMS: 5.0–5.3 kHz
Tjaden & Turner (1997)	4 adults (M) 3 adults (F)	34–68 yrs[f]	word initial position from reading passage	CSpeech (FFT, moments)	first 4 time slices (each 10 ms)	22 kHz	0–9 kHz	Centroids: M = 4–7 kHz F = 6.5–8.1 kHz

[a]Sex is given in parentheses (M = male, F = female, NR = not reported). [b]NR = Specific ages not reported. [c]Analysis method (LPC = linear predictive coding; FFT = fast Fourier transforms; DFT = discrete Fourier transforms; moments = see text). [d]Report "... a number of English speakers ..." (p. 303); present data for 2 males and 1 female. [e]Estimated from graphs. [f]For disordered speakers matched to normal speakers; exact ages not given for normal speakers. [g]Figures in parentheses are males/females. [h]No exact numbers were given (reports approximately 12 per group). [i]Not specified; inferred from sampling rate. [j]Age range for initial group of 35 speakers. [k]LMS = less-masculine-sounding speakers, MMS = more-masculine-sounding speakers.

Acoustic Analyses and Algorithms

Over the 5 decades of research represented in Table 1, platforms for acoustic analysis have evolved from dedicated spectrographic instruments to personal computer-based software. There have been relatively fewer changes in analysis algorithms, with studies since 1980 using either a version of Fourier analysis, such as a discrete Fourier transform (DFT) or a fast Fourier transform (FFT), or linear predictive coding (LPC) to profile the frequency and intensity characteristics of /s/. Comparative data suggest equivalences among these approaches (Bauer & Kent, 1987, p. 506; see also Kent & Read, 1992, Figures 4–8, p. 73). The goals of each of the studies in Table 1 have dictated the selection of an algorithm. Where a central tendency value is viewed as the appropriate statistic to characterize /s/ production (e.g., amplitude peak, low prominent peak), LPC appears to be the appropriate technique. Alternatively, where one or more measures of distribution of the frication noise are viewed as appropriate (e.g., spectral moments), Fourier analysis has provided the needed information. The current needs to characterize /s/ in typical speakers for comparison to /s/ in atypical speakers mandated the greater information available in FFT analyses.

Sampling Rate

For the older analog procedures, such as the Kay sonograph, sampling rate is not applicable, but with studies using digital approaches, rates typically ranged from 20–25 kHz. There were at least two exceptions to this, both involving higher rates. Bauer and Kent (1987) sampled at 40 kHz in an attempt to look for very high frequency components in fricatives and trills produced by infants. These authors postulated the presence of such high frequencies as a direct function of the very small oral cavities in infants, a hypothesis supported by the acoustic data. Pentz (1996) also used a 40 kHz sampling rate. It is noteworthy that, in both of these studies, the obtained frequency values were higher than values reported in all other studies using lower sampling rates. One significant disadvantage of a higher sampling rate is the creation of very large data files, which, until very recently, posed data storage and retrieval problems. Higher sampling rates may also not have been feasible in some digital studies due to the limited frequency response ranges of some hardware components. In the current study, for example, sampling rate was limited to 22 kHz by a combination of the frequency response characteristic of the tape recorder used for both recording and digitizing, which had an upper limit of 10 kHz, and the use of an antialiasing filter (see below).

Analysis Range

Most of the studies in Table 1 (14 of 18, 78%) were limited to examination of the spectral energy of 10 kHz or lower, despite the fact that, nearly 40 years ago, Strevens (1960) noted that spectral energy in fricatives produced by adults may be present up to at least 12 kHz. Strevens did note that his failure to analyze above 8 kHz was a function of instrumental limitations. In addition to limiting the analysis range, at least nine of the studies made use of low-pass filtering, probably in response to concerns about aliasing (Bauer & Kent, 1987; Baum & McNutt, 1990; Behrens & Blumstein, 1988; Harmes et al., 1984; Hoole et al., 1989; Miccio et al., 1996; Nittrouer et al., 1989; Sereno et al., 1987; Tjaden & Turner, 1997). Such filtering is evident in those studies in Table 1 in which the upper limit to the analysis range is less than one half of the specified sampling rate. The current study used an antialiasing filter of 89% of the analysis range.

Spectra of /s/

The diversity of ways to characterize /s/ makes it difficult to compare findings across the studies summarized in Table 1. For example, the two earliest studies reported the range of frequencies at which most of the spectral energy for /s/ is concentrated (Hughes & Halle, 1956; Strevens, 1960), whereas the later studies characterized the spectra of /s/ with a single value (i.e., highest amplitude peak, centroid, low cutoff) selected in a variety of ways. Several investigators identified the highest amplitude peak present (Bauer & Kent, 1987; Behrens & Blumstein, 1988; Daniloff et al., 1980; Hoole et al., 1989; Pentz et al., 1979; Schwartz, 1968; Sereno et al., 1987). Pentz (1996) reported unweighted mean frequencies. Boothroyd and Medwetsky (1992) reported what they termed the *low prominent peak*, defined as the lowest frequency peak present that was within 10 dB of the highest amplitude peak. The choice of measure appeared appropriate in the latter study, however, because the goal had been to identify the minimum bandwidth needed for perception of fricatives by hearing aid users. The other studies in Table 1 reported centroids, or amplitude-weighted means, but those centroids themselves were calculated at least three different ways. Harmes et al. (1984) derived their centroid values from FFT data weighted to include 40% of the spectral energy surrounding the amplitude peak. Most centroid calculations used FFT data and formulae presented by Forrest et al. (1988; Avery & Liss, 1996; Baum & McNutt, 1990; Nittrouer, 1995; Tjaden & Turner, 1997). Nittrouer et al. (1989) and Katz et al. (1991) also used the Forrest et al. formulae to derive their centroids, but based their analyses on DFT data.

For the present study, the *moments analysis* approach described by Forrest et al. (1988) was selected to characterize the /s/ spectra of typical speakers. Moments analysis was adapted to the case of speech analysis to provide acoustic indices of the whole spectrum, as opposed to indices that capture just a single aspect of the spectrum such as peak frequency. The approach accomplishes its goal by characterizing the entire distribution of fricative energy statistically, using FFTs calculated from the digitized signal, with each FFT treated as a random probability distribution. The first four spectral moments of that distribution (mean, standard deviation, skew, and kurtosis) are derived from the FFTs on a linear frequency scale. The moments approach, as implemented in CSpeech (Milenkovic, 1996), also includes transformation of the FFTs to a Bark scale (Syrdal & Gopal, 1986). Bark-transformed moments have been shown to be useful in discriminating among fricative types (Forrest et al., 1988), but it is not known whether linear or Bark-transformed moments provide a better characterization of typical /s/ production.

Summary and Goal of Study

The many methodological differences reviewed above preclude summary consensus on the acoustic findings listed in Table 1. Two generalizations appear to be well supported. First, typical /s/ production can be characterized as a band of energy above approximately 4 kHz, a conclusion reached by the very first study completed over 4 decades ago (Hughes & Halle, 1956). However, as indicated by the wide range of values listed in Table 1, all attempts to characterize /s/ in more detail are limited by subject sampling and measurement constraints. A second generalization from these studies is that frequency values for /s/ production by females tend to be higher than values produced by males. As indicated previously, this sex difference is consistent with sex differences in the size of the resonating cavity in front of the point of constriction for /s/.

The goal of the current study was to provide a set of measurement procedures and reference data on the typical production of /s/ by 9- to 15-year-old speakers for speech disorders research. In view of the gaps and equivocal findings in the precedent literature, five independent variables under inspection included age, sex, sampling point for /s/, linguistic context for /s/, and linear versus Bark-transformed data. The candidate-dependent variables for the reference data were the four spectral moments obtained using CSpeech.

Method

Participants

Children were recruited in the context of a larger study of children with histories of several types of speech-sound disorders. For the goals of the present study, 2 children of each sex for each grade level from grades 4 to 9 were recruited. Classroom teachers in the Madison Metropolitan School District randomly selected 4 typically developing children from each class in which children with speech disorders had been obtained for the larger study. In addition to grade and sex criteria, typically speaking participants met the following criteria: (a) no history of a speech problem, as confirmed by the school speech-language pathologist; (b) no history of special education services, as confirmed by school records; (c) standard scores above 90 on the Peabody Picture Vocabulary Test–Revised, Form M (Dunn & Dunn, 1981), as later obtained by the speech examiner who administered the speech protocol; (d) no dialectal differences from general American speech, as later confirmed by the speech examiner; and (e) no dental braces or other orthodontic appliances, as later confirmed by the examiner. Parents were invited by mail to participate in the study; those who responded and whose children met study criteria were scheduled for assessment. This recruitment procedure yielded 24 typically speaking children who were tested over a 3-month period. To fill a gap in the age distribution for males, data for 2 typically speaking boys identified through acquaintances and siblings of children enrolled in other studies were added several months later. Table 2 is a summary of the ages of the typically speaking children. Ages ranged from 9;7 to 15;2, with a mean age of 12;4 (*SD* = 1;9). There was no significant age difference between female and male speakers (*t* = 0.01, *df* = 22, *p* = .99).

Table 2. Age and fundamental frequency data for the 26 typically speaking adolescents.

		Age (years;months)			Fundamental frequency (Hz)		
	n	*M*	*SD*	Range	*M*	*SD*	Range
Females	12	12;4	1;11	9;7–15;2	209	12.0	191–228
Males	14	12;4	1;9	9;7–14;10	200	48.1	118–279
Overall	26	12;4	1;9	9;7–15;2	204	35.9	118–279

Administration of the /s/-Sample Task

An /s/-sample task was administered to each of the 26 children as part of a 90-minute protocol conducted by one speech-language examiner in a comfortable test suite at the Waisman Center on Mental Retardation and Human Development on the University of Wisconsin–Madison campus. Parents had the option of observing the assessment through a one-way mirror. The /s/-sample task was embedded in a larger 120-item task intended to sample production of words containing both /s/ and /r/. The 120-item task was the fourth task in the assessment protocol and required approximately 10 minutes to complete. A Shure Model SM10-A low-impedance, unidirectional headset microphone was placed on each subject's head, and the headband was adjusted for head size and comfort. The microphone was positioned approximately 1.5 inches from the lips and no more than 2 inches from the subject's nose, and the microphone head was tilted so that it pointed toward the nose. Prior experience indicated that this configuration provided optimal signal capture while minimizing negative effects from plosive bursts. Productions were recorded on a Sony TCM-5000EV analog cassette recorder. Recording volume was manually adjusted while the child produced pretest words. The examiner later transcribed all the samples following a system of narrow-phonetic transcription and conventions developed for research in child phonology (Shriberg, 1986; Shriberg & Kent, 1995).

Stimuli for the 120-item task consisted of five randomized lists of 24 words (120 total tokens) produced in the carrier phrase "Say _____ again." Words were presented live by the examiner, who read from a typed list; the child could not see the list or the examiner's face during this task. Speakers were asked to repeat the target word in the carrier phrase while maintaining loudness within a preset range as indicated by the VU meter on the tape recorder. The examiner monitored the participants' alertness and performance and asked children to repeat a phrase if the target word appeared not to be understood, was produced incorrectly, or contained obvious interword pauses or dysfluencies. All 120 tokens were recorded within a single session. Speakers were permitted a brief break after 60 tokens. If the examiner sensed any fatigue, the tape recorder was stopped, and the child was given a short rest. Immediately following completion of the acoustics task, the examiner played back a portion of the tape to confirm that the quality of the recording was acceptable. At completion of the entire assessment protocol, each participant received $25.00.

The current study involved analysis of the 50 tokens for /s/, which included five repetitions each of the following 10 words: *sin*, *spin*, *skin*, *soon*, *spoon*, *kiss*, *kits*, *kicks*, *cosine*, and *assign*. These 10 words sample /s/ in five canonical forms, five word positions, and several consonant and vowel contexts.

Speaker Validation

To evaluate the representativeness of the speakers for this age range, speaking fundamental frequency (f_0) was obtained using an average of up to 10 measures per speaker from productions of "Say assign again," including five measurements from /e͞ɪ/ in *say* and five from the /ɑ͞ɪ/ in *assign*. The interval from the 5th glottal pulse of the vowel to a point 50 ms later in time was isolated. The pitch period was identified using the two most similar but contiguous glottal pulses, and CSpeech then provided f_0 values derived from the inverse of the period. Group f_0 values by sex are shown in Table 2. Female values were relatively stable across this age range, whereas male values predictably dropped significantly across this developmental period. The individual speaker values were consistent with age and sex expectations (Wilson, 1987, pp. 119–124).

Speech Analysis

Acoustic analyses were accomplished by two trained research assistants, each of whom was randomly assigned half of the subjects. The assistants, who had each completed a course in speech acoustics, followed a well-developed protocol for the analyses (Flipsen, Tjaden, Weismer, & Karlsson, 1996). Using a second Sony 5000EV tape recorder as the input source, tokens were digitized using a Sound Blaster AWE32 PNP A/D sound card connected to a Pentium-based PC. The signal was sampled at 22 kHz with 15 bits of quantization and low-pass filtered at 9.8 kHz using the record utility of the software program CSpeech (Milenkovic, 1996). Tokens were again evaluated during digitizing for incorrect productions, obvious interword pauses, or dysfluencies. Pauses were measured from the spectrogram, with a pause defined as any period of silence of 250 ms or longer (Miller, Grosjean, & Lomanto, 1984). Of 1,300 possible tokens, only 13 (1%) were rejected because of the presence of pauses, dysfluencies, or production of an incorrect target. The target word in the interval from the start of /e͞ɪ/ in *say* to the closure for /g/ in *again* was isolated and stored.

The following segmentation criteria were used to identify the /s/ segment for analysis: (a) for *sin* and *soon*, from the last glottal pulse of /e͞ɪ/ in *say* or beginning of frication (glottal pulse criterion preferred unless frication is not continuous with the vowel) to the first glottal pulse of vowel following /s/; (b) for *skin*, *spin*, and *spoon*, from the last glottal pulse of /e͞ɪ/ in *say* or beginning of frication to the onset of closure for the stop consonant following /s/; (c) for *assign*, from the last glottal pulse of the /ə/ to the first glottal pulse of /ɑ͞ɪ/; (d) for *cosine*, from the last glottal pulse of /o͞ʊ/ to the first glottal pulse of /ɑ͞ɪ/; (e) for *kicks* and *kits*, from the burst release of the second /k/ (or /t/) to either the first glottal pulse of the /ə/ in *again* or the end of frication; specific evidence for a burst release of the stop consonant was not always evident, but, when present, could not be reliably separated from the frication for /s/; where it was not present, /s/ was deemed to begin at the onset of frication; conversely, if multiple stop-bursts occurred, the onset of frication was based on the burst that was contiguous with the fricative energy (i.e., the last one); and (f) for *kiss*, from the last glottal pulse of /ɪ/ to either the first glottal pulse of /ə/ in *again* or the end of frication.

The /s/ segments for each speaker were analyzed using the *moments* batch command function in CSpeech. Analyses were completed using a 20-ms Hamming window with a 10-ms step. For each token produced, three temporal points in the frication noise were then identified from the moments output (onset, midpoint, and offset). The onset was defined as the first analysis window in the output (i.e., the first 20 ms of the frication). The midpoint was defined as the middle analysis window (i.e., the middle 20 ms). In cases where a token contained an even number of analysis windows, one of the two middle windows was selected randomly. Offset was defined as the last analysis window (i.e., the last 20 ms).

Reliability Estimates

To estimate the intrajudge reliability of the acoustic measures, each assistant resegmented 20% of the tokens from a randomly determined 2 (15%) of the approximately 13 speakers each had analyzed. An interjudge reliability estimate was obtained by having each assistant segment 20% of the tokens from 2 (15%) speakers originally measured by the other assistant. The resegmented tokens were then analyzed using the moments batch function (as above). Reliability estimates obtained for each of the three points in the frication (onset, midpoint, and offset) are shown in Table 3. The entries in Table 3 indicate the mean difference between values obtained (along with standard deviations).

The estimates in Table 3 indicate adequate intrajudge and interjudge agreement for the present purposes. The tendency for smaller mean differences at the midpoint of the frication likely reflects the fact that cursor placement (i.e., the basis for resegmentation) would likely affect onsets and offsets more so than the

Table 3. Intrajudge and interjudge agreement data for the two research assistants who accomplished the acoustic measurements.

	Intrajudge agreement for Research Assistant 1			Intrajudge agreement for Research Assistant 2			Interjudge agreement		
Acoustic variable	**Onset**	**Midpoint**	**Offset**	**Onset**	**Midpoint**	**Offset**	**Onset**	**Midpoint**	**Offset**
Moment 1									
Frequency Mean (kHz)	0.57	0.35	0.73	0.56	0.20	0.33	0.67	0.10	0.42
	(0.65)	(0.45)	(0.81)	(0.46)	(0.18)	(0.28)	(0.64)	(0.13)	(0.66)
Moment 2									
Frequency SD (kHz)	0.14	0.15	0.27	0.26	0.13	0.21	0.27	0.07	0.17
	(0.22)	(0.30)	(0.34)	(0.26)	(0.15)	(0.24)	(0.27)	(0.07)	(0.21)
Moment 3									
Frequency Skew	0.43	0.28	0.56	0.50	0.44	0.33	0.45	0.19	0.36
	(0.55)	(0.27)	(0.57)	(0.50)	(0.52)	(0.31)	(0.43)	(0.22)	(0.41)
Moment 4									
Frequency Kurtosis	0.86	1.08	1.42	2.03	2.78	1.89	1.54	1.60	0.89
	(1.37)	(1.59)	(1.83)	(2.42)	(3.40)	(1.97)	(1.60)	(3.23)	(1.04)

Note. Cell entries are mean differences between measurements (and standard deviations).

Table 4. Standard error of measurement estimates for each of the spectral moments measures at /s/ onset, midpoint, and offset.

	Onset				Midpoint				Offset			
	M	***SD***	**ρ**	**SEM**	***M***	***SD***	**ρ**	**SEM**	***M***	***SD***	**ρ**	**SEM**
Moment 1 linear frequency mean (kHz)	5.07	1.11	.76	0.55	6.48	1.19	.98	0.16	5.23	1.32	.79	0.60
Moment 2 linear frequency *SD* (kHz)	1.59	0.51	.68	0.29	1.20	0.43	.93	0.11	1.41	0.50	.90	0.16
Moment 3 linear frequency skew	–0.13	0.85	.74	0.45	–0.02	1.10	.96	0.23	0.15	1.08	.84	0.43
Moment 4 linear frequency kurtosis	1.55	2.33	.57	1.53	4.31	9.02	.92	2.55	3.31	3.97	.90	1.24

midpoint. The notable lack of information on reliability in prior studies makes it difficult to compare present with precedent findings. Of the 21 studies in Table 1, only three included reliability data (Avery & Liss, 1996; Menon et al., 1969; Yeni-Komshian & Soli, 1981). Menon et al. were the only group to report reliability data on measurement of /s/ itself. They reported that at least 92% of their frequency measures on /s/ were within 60 Hz of each other, higher agreement than found in the present study. However, they measured the lowest frequency present in the frication noise, whereas the current study included the entire distribution of frequencies at each of three separate points in the frication. Moreover, Menon et al. appear to have sampled a much narrower range of frequencies (approximately 7 kHz) compared to the 9.8 kHz used in the present study.

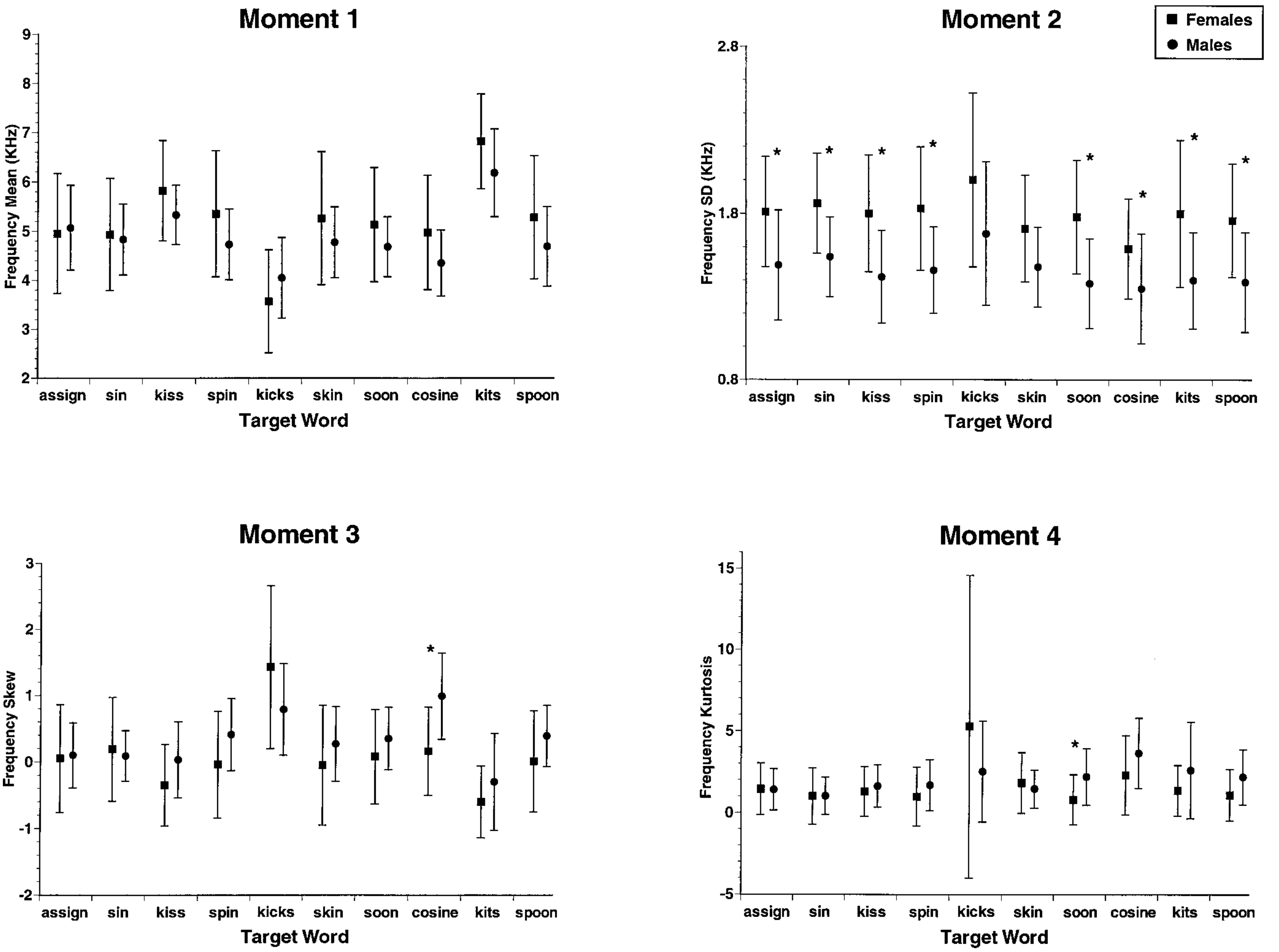
Standard Error of Measurement Estimates

Table 4 provides estimates of the standard error of measurement (SEM) for the reference data in Figures 1–3. The columns in Table 4 represent the average mean, average standard deviation, Spearman rank-order correlation, and SEM for each pair-wise comparison in the interjudge reliability data for each point in the frication. Spearman rank-order correlations were used in the SEM calculations because 9 of the 24 distributions (4 moments by 3 points in the frication by 2 examiners) were nonnormal as determined using the Anderson-Darling test of normality. These estimates, indicating the error boundaries around speakers' obtained values, will ultimately be useful when inspecting effect sizes in speech disorders research.

Results and Discussion

Figures 1–3 provide reference data for /s/ production in 9- to 15-year-old typically speaking children. Means and standard deviations for each of the spectral moments for the onset (Figure 1), midpoint (Figure 2), and offset (Figure 3) of the frication of /s/ for each of the 10 words are presented for the 12 girls and 14 boys. Decisions and findings motivating the format and content of the data in Figures 1–3 relevant to candidate versus selected independent and dependent variables are posed as questions and discussed in the following sections. Note that the order of presentation for the words

Figure 1. Reference data for /s/ at frication onset for 26 typically speaking adolescents. Significant ($p < .05$) female-male differences on the Wilcoxon-Mann-Whitney test are indicated by an asterisk.



in Figures 1–3 was motivated empirically using highest–lowest frequency mean (Moment 1, Figure 2) at the midpoint in the frication. Although the goal of the current study was to generate a reference data set, figures were felt to be more appropriate in the present context because they more clearly convey the relative relationships both within and between variables. For the interested reader, tables containing all obtained data values (including linear and Bark-transformed values) are archived in a technical report that can be downloaded from the University of Wisconsin–Madison Phonology Project Web site (http://www.waisman.wisc.edu/phonology/).

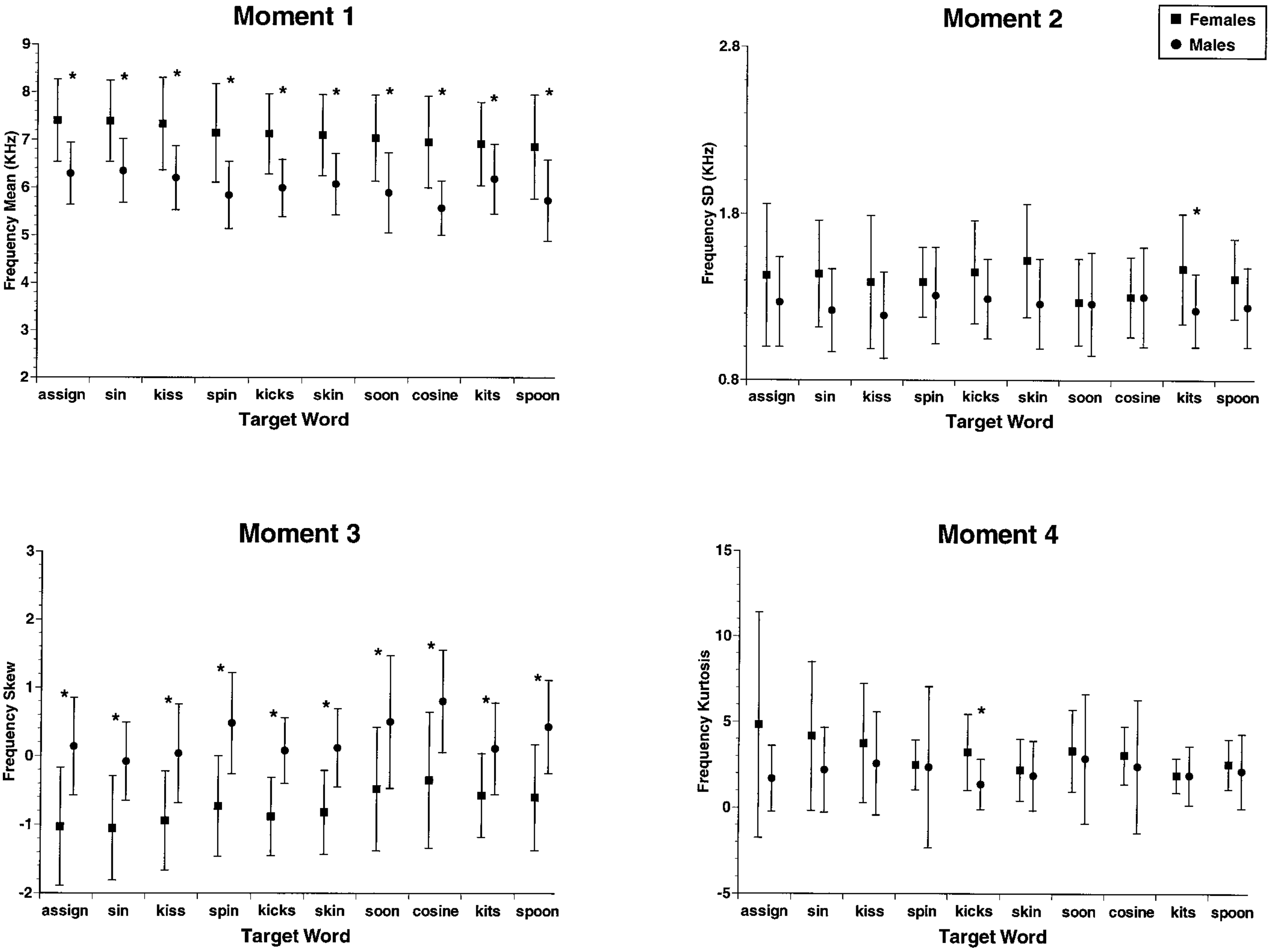
What Is the Appropriate Sampling Point Within the /s/ Frication for the Reference Data?

To identify the appropriate sampling point at which to characterize /s/, values for each available token (maximally five) of each of the 10 word types were averaged by word for each of the 26 speakers and then averaged by gender for each position in the frication. These gender-average values are presented in Figures 1–3.

Beginning with the Moment 1 data (the top-left panels in each of the three figures), the highest values were obtained at the midpoint of the frication for every word with one exception. In the case of the word *kits*, the frequencies at onset and midpoint were identical for the male speakers and nearly identical for the female speakers. This similarity is not surprising given that the segmentation criteria for /s/ in *kits* included the burst release for /t/. Because /t/ and /s/ are homorganic, the burst for /t/ and the frication for /s/ would be expected to be similar in frequency.

Comparing the Moment 2 data in each of the three figures (the top-right panels), the average standard deviation values (representing the dispersion of frequency values within each token) were lowest at the midpoint for 7 of the 10 words for the female speakers, and 9 of the 10 words for the male speakers. The exceptions were

Figure 2. Reference data for /s/ at frication midpoint for 26 typically speaking adolescents. Significant ($p < .05$) female-male differences on the Wilcoxon-Mann-Whitney test are indicated by an asterisk.

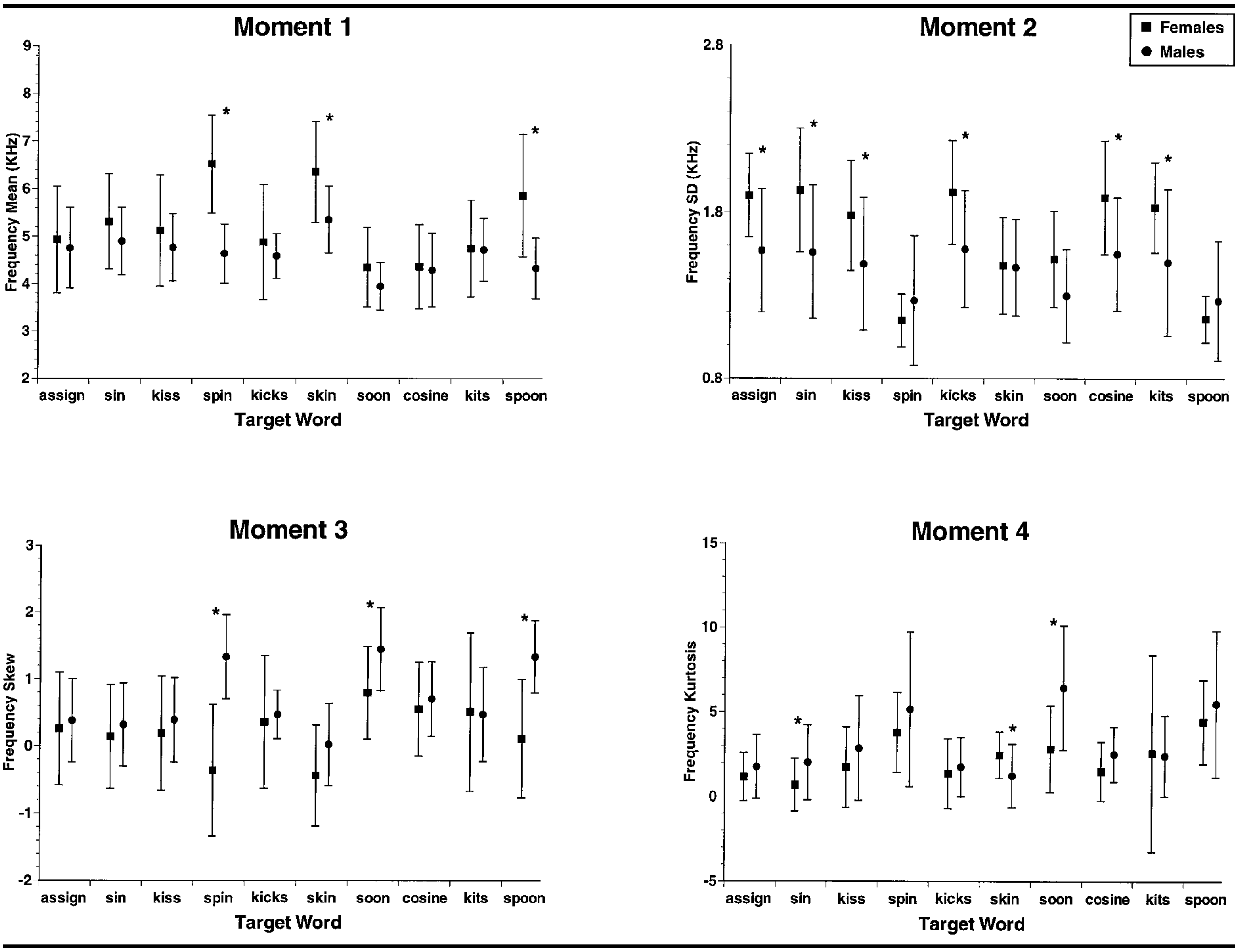


all words beginning with /s/ clusters (*spin* for the males; *spin*, *skin*, and *spoon* for the females) for which average Moment 2 values were lowest at the offset of the frication. Less within-token dispersion at the offset of the frication for these words is expected because the offset of the frication coincided with the onset of closure for the following stop consonant. It may be that the presence of the stop consonant following /s/ constrains the articulatory behavior of /s/ more than would be seen with a following vowel, where the demands on tongue movements are less stringent. This notion of greater articulatory demands reducing variability is supported by findings from Klatt (1974) who reported that, although /s/ may be longer in duration before a stressed vowel, the standard deviations of those durations were smaller when compared to /s/ before an unstressed vowel.

For the Moment 3 data in the bottom-left panels in Figures 1–3 (the frequency skew), the only consistent pattern was that the values were all lowest and consistently negative for the female speakers at the midpoint. This pattern will be discussed in a section to follow.

Finally, as shown in the bottom-right panels in each figure, the only observable pattern for the Moment 4 data (frequency kurtosis) was that, at the onset of the frication for the word *kicks* (Figure 1), a much larger standard deviation was obtained for the female speakers than for the males. Review of the individual speaker data indicated notably high mean kurtosis values for two female speakers, ages 9;7 and 13;2 years. Each speaker's data had been reduced by a different assistant. Spectrographic review indicated that multiple burst events may have occurred on several word tokens; recall that the burst for /k/ is included in the segmentation procedures for *kicks*. But in cases of multiple bursts, only the final burst is to be included. Because these multiple bursts were not always sharply defined, the assistants needed to make a judgment about whether or not to include a burst in the frication. Anecdotal reports by the

Figure 3. Reference data for /s/ at frication offset for 26 typically speaking adolescents. Significant ($p < .05$) female-male differences on the Wilcoxon-Mann-Whitney test are indicated by an asterisk.



assistants suggested that segmentation of tokens was often problematic for this word (as it was in *kits*), potentially lowering the obtained frequency (Moment 1) values for /s/. Given a strong negative Spearman correlation (–.84) between Moment 1 and Moment 4 for this data set, a decrease in Moment 1 values would likely be accompanied by an increase in Moment 4 values.

Overall, these findings support the expectation that, particularly as reflected in frequency mean and standard deviation values, /s/ differs at the midpoint compared to onset and offset. The lower frequency mean (Moment 1) and greater within-token dispersion of frequencies (Moment 2) at the perimeter of the frication suggest generalized context effects, making the midpoint more attractive for the independent characterization of /s/. A possible negative consequence of choosing this point, however, may be the loss of information sensitive to developmental effects. Context effects have been implicated as potentially informative for understanding articulatory development (cf. Nittrouer, 1995; Nittrouer et al., 1989; Pentz, 1996; Sereno et al., 1987; Soli, 1981). However, significant age trends (see below) were not observed at any of the three points in the frication, suggesting that, if present in this age range, developmental trends are not accessible with the measures used in this descriptive study. Studies suggesting developmental changes have used measures intended to identify potential contextual interactions (e.g., an energy peak in the frequency range of that expected for an adjoining vowel).

Should the Reference Data Be Collapsed by Word Type (Linguistic Complexity, Phonetic Context)?

At least three independent findings justify a decision to present data at the level of word types, indicating that phonetic context has significant effects on /s/

production. First, there is the failure to separate Moment 1 data at onset and midpoint as discussed above for the word *kits*. Second, data for the word *kicks* (especially for the female speakers) at the onset of the frication differ from those of the other words: the Moment 1 (frequency mean) values for both sexes are lowest of all the words, and the Moment 2 (standard deviations) values are the largest. Additionally, the means and standard deviations of Moment 4 (kurtosis) are the most variable for both sexes. As with the word *kits*, these findings likely reflect the inclusion of lower frequency energy due to segmentation criteria that included the burst release for /k/, expected to be lower in frequency than /s/. Finally, there is the lack of sex differences at the onset of the frication for Moment 2 for words involving /s/ in cluster contexts with /k/, and at the offset of the frication for Moment 2 in words beginning with /s/ clusters and in the word *soon* (see below).

Should the Reference Data Be Collapsed by Sex?

The Moment 1 data in Figures 1–3 suggest the presence of reliable sex differences in the characteristics of the frication for /s/. A series of Wilcoxon-Mann-Whitney tests to assess the strength of sex differences was completed for each of the 10 target words for each of the four moments at each of the three points in the frication. For the present purposes, a Bonferroni correction was not applied to a .05 alpha significance level for each comparison. Results of the inferential statistical analyses are indicated in Figures 1–3.

Frication Onset

At the onset of the frication (Figure 1), the only significant pattern of sex differences was obtained for Moment 2 (within-token dispersion). Statistically significant sex differences were obtained on 8 of the 10 comparisons, with female values being higher than male values. The two exceptions (involving /sk/ clusters) were both in the same direction as the other eight, but the alpha values (both ps < .10) were not significant. Thus, compared to males within this age range, females have a wider distribution of fricative energy at the onset of the frication of /s/. Possible explanations for this finding are considered in the following section.

Frication Midpoint

The most notable findings for the midpoint of the frication (Figure 2) were significant sex differences for all 10 words for both Moment 1 (frequency mean) and Moment 3 (frequency skew). Female values for Moment 1 were significantly higher than male values by about 1.1 kHz. Moment 3 values for the female speakers were negative (male values tended to be near zero or slightly positive) with a difference between females and males of approximately 1 skew unit. Sex differences were much smaller for both Moments 1 and 3 on *kits* (0.7 kHz and 0.68 skew units, respectively).

The observed sex differences for Moment 1 may have been a function of differences in the size of the resonating cavity anterior to the point of constriction (Daniloff et al., 1980), with the higher values for females presumably reflecting a smaller space. Sex differences may be expected, even in younger children. Nellhaus (1968) observed that “at all ages [birth to 18 years] the head circumference of boys is approximately 0.9 cm larger than that of girls” (p. 107). Such structural differences may underlie consistent sex differences in the size of the resonating cavity for /s/.

An alternative explanation for the observed sex difference in frequency mean may be that the female speakers used a tongue placement for /s/ that was relatively farther forward than that of the male speakers, thereby producing a temporarily smaller cavity. This possibility is supported by the fact that perceptual transcription of the /s/-sample task tokens revealed significantly more of the female speakers producing at least one dentalized token of /s/ (6/12 vs. 2/14 males; $\chi^2 = 3.869$, $df = 1$, $p = .049$). Also, significantly fewer female speakers produced at least one palatalized token of /s/ (1/12 vs. 6/14 males; $\chi^2 = 3.914$, $df = 1$, $p = .048$). It is, of course, possible that both a structural difference and a placement difference were operating concurrently to generate the observed sex differences, with lowered accuracy resulting from a more forward tongue position attempting to hit a smaller target.

The negative skew (Moment 3) values observed at the midpoint of the frication for the female speakers may represent either some natural upper limit to the distribution of frequencies produced by the female speakers or a statistical artifact of the upper limit of the analysis band (9.8 kHz). Both of these possibilities were supported by a series of Pearson-product-moment correlations between Moment 1 and Moment 3, completed using a single randomly selected token for each word per speaker. Correlation values ranged from 0.73–0.87 (all ps < .001) across the 10 target words. If the negative skew represents a statistical artifact, it would suggest that truncation of higher frequency components in the female productions had occurred. Such a bias in the data would indicate that the frequency means for females reported in Figure 2 may underestimate the true values that would be obtained with an analysis band extending to frequencies higher than 9.8 kHz (cf. Bauer & Kent, 1987). Data for male speakers would be less suspect for bias, given males’ lower frequency (M = 5.6–6.3 kHz) and less negative skew, as would all the data at the onset and offset of frication (with the possible exception of onset data for *kits*). This would then suggest that the sex difference observed for frequency mean may be larger than indicated. Determining whether the negative skew represents the actual distribution of frequencies produced or the effects of truncation was, however, not possible with the current data set.

Frication Offset

Findings for the offset of the frication (Figure 3) illustrate a pattern of sex differences somewhat similar to that seen in the onset data. Moment 2 again provided the most consistent sex differences for 6 of the 10 words, with female values again being higher (more within-token dispersion). It could be argued that if tongue placement were more fronted in female speakers, as proposed above, the speaker may need to begin movement to the following sound sooner, resulting in a less stable production at the end of the frication.

Exceptions to sex difference findings for Moment 2 included all three words beginning with /s/ clusters. The potential constraints of a stop consonant compared to a following vowel were discussed above (cf. sampling point). The one other exception to the sex differences observed for Moment 2 was for the word *soon*; there is no obvious articulatory-phonetic motivation for this exception to the trends for the other words with singleton /s/ onsets. Overall, these findings provide strong support for the need to provide individual reference data for each sex.

Should the Reference Data Be Collapsed Across Age?

All speaker data were inspected for trends across age. A series of linear regressions were conducted (10 words, 8 moments, 2 sexes) at each point in the frication (onset, midpoint, offset). At /s/ onset, only 3 of 160 analyses were significant ($p < .05$); at /s/ midpoint, only 4 of 160 analyses were significant; and at /s/ offset, none of the 160 analyses was significant. All statistically significant analyses were attributed to chance. Thus, because there was no support for age trends by word, by moment, or by sex at any of the three points in the frication, it is appropriate to collapse the reference data across age.

Lack of significant age trends may have been associated with power constraints due to the relatively small sample size. However, adult data from Weismer and Bunton (1998) support the probability that /s/ production may not differ within the current age range. Using identical spectral analysis procedures, Weismer and Bunton found frequency means for /s/ at the midpoint in the frication in *suit* of 7.0 kHz and 5.7 kHz for 10 female and 11 male speakers, respectively (ages 21–25 years). These values are consistent with those obtained in the present study for *soon* (see Figure 2). Thus, at least as indexed by frequency means, /s/ production may stabilize at these values after approximately age 9.5 years.

Which Scale Is Appropriate for the Reference Data: Linear or Bark-Transformed Spectra?

CSpeech calculates moments from both linear (unweighted) and Bark-transformed spectra (cf. Forrest et al., 1988; Syrdal & Gopal, 1986), allowing inspection of potential differences in assembling a reference data set for /s/ production. A primary finding in these comparisons was that the sex differences described above for linear moments were not observed for any of the moments at any of the three positions in the frication using data from the Bark-transformed scale. Moreover, age trends, which were not apparent in the linear data, were also not observed with the Bark data. Thus, although Bark-transformed data have previously been shown to be useful in developing discriminant functions to distinguish among fricative types (e.g., /s/ vs. /ʃ/; Forrest et al., 1988), in the present study they lacked the sensitivity to the sex differences identified with linear data and were not associated with any additional findings. Therefore, the linear data were deemed appropriate for the reference database.

Which One or More of the Four Moments Best Characterizes /s/?

The primary goal of the analyses was to motivate the use of one or more of the four spectral moments to characterize /s/. Considering that the midpoint of the frication was deemed to be the most appropriate sampling point, and that the first and third spectral moments were most informative on sex differences at the midpoint and that linear scale data highlighted these differences, linear scale data for the first and third moments at the midpoint are considered essential to characterize /s/ independent of phonetic context. As discussed, however, where phonetic context is of potential interest, the second moment may include relevant data when sampled in the onset and offset of /s/ frication. Finally, the analyses suggested that information on the fourth moment does not add significant insights into the characteristics of /s/ in typically speaking children from 9–15 years of age.

Acknowledgments

Thanks to the following colleagues for their assistance: Chad Allen, Diane Austin, Kate Bunton, Frederic Gruber, Robert Hagiwara, Doris Kistler, Rachel Riely, Kristin Tjaden, and David Wilson. Preparation of this paper was supported by the National Institute on Deafness and Other Communication Disorders, National Institutes of Health, Grant DC00496.

References

Austin, D., & Shriberg, L. D. (1996). *Lifespan reference data for ten measures of articulation competence using the Speech Disorders Classification System (SDCS)* (Tech. Rep. No. 3). Phonology Project, Waisman Center on Mental Retardation and Human Development, University of Wisconsin–Madison.

Avery, J. D., & Liss, J. M. (1996). Acoustic characteristics of less-masculine-sounding male speech. *Journal of the Acoustical Society of America, 99*, 3738–3748.

Bauer, H. R., & Kent, R. D. (1987). Acoustic analyses of infant fricative and trill vocalizations. *Journal of the Acoustical Society of America, 81*, 505–511.

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

Behrens, S. J., & Blumstein, S. E. (1988). Acoustic characteristics of English voiceless fricatives: A descriptive analysis. *Journal of Phonetics, 16*, 295–298.

Bond, Z. S., & Wilson, H. F. (1980). /s/ plus stop clusters in children's speech. *Phonetica, 37*, 149–158.

Boothroyd, A., & Medwetsky, L. (1992). Spectral distribution of /s/ and frequency response of hearing aids. *Ear and Hearing, 13*, 150–157.

Cucchiarini, C. (1996). Assessing transcription agreement: Methodological aspects. *Clinical Linguistics & Phonetics, 10*, 131–155.

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.

Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test–Revised*. Circle Pines, MN: American Guidance Service.

Flipsen, P., Jr, Tjaden, K., Weismer, G., & Karlsson, H. (1996). *Acoustic analysis protocol* (Tech. Rep. No. 4). Phonology Project, Waisman Center on Mental Retardation and Human Development, University of Wisconsin–Madison.

Forrest, K., Weismer, G., Elbert, M., & Dinnsen, D. A. (1994). Spectral analysis of target-appropriate /t/ and /k/ produced by phonologically disordered and normally articulating children. *Clinical Linguistics & Phonetics, 8*(4), 267–281.

Forrest, K., Weismer, G., Milenkovic, P., & Dougall, R. N. (1988). Statistical analysis of word-initial voiceless obstruents: Preliminary data. *Journal of the Acoustical Society of America, 84*, 115–123.

Harmes, S., Daniloff, R. G., Hoffman, P. R., Lewis, J., Kramer, M. B., & Absher, R. (1984). Temporal and articulatory control of fricative articulation by speakers with Broca's aphasia. *Journal of Phonetics, 12*, 367–385.

Hewlett, N. (1988). Acoustic properties of /k/ and /t/ in normal and phonologically disordered speech. *Clinical Linguistics and Phonetics, 2*, 29–48.

Hoole, P., Ziegler, W., Hartmann, E., & Hardcastle, W. (1989). Parallel electropalatographic and acoustic measures of fricatives. *Clinical Linguistics and Phonetics, 3*, 59–69.

Hughes, G. W., & Halle, M. (1956). Spectral properties of fricative consonants. *Journal of the Acoustical Society of America, 28*, 303–310.

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

Kent, R. D. (1976). Anatomical and neuromuscular maturation of the speech mechanism: Evidence from acoustic studies. *Journal of Speech and Hearing Research, 19*(3), 421–447.

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, 5*(3), 7–23.

Kent, R. D., & Read, C. (1992). *The acoustic analysis of speech*. San Diego, CA: Singular.

Klatt, D. H. (1974). The duration of /s/ in English words. *Journal of Speech and Hearing Research, 17*, 51–63.

Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America, 87*, 820–857.

Lewis, B. A., & Shriberg, L. D. (1994, November). *Life span interrelationships among speech, prosody-voice, and nontraditional phonological measures*. Paper presented at the annual meeting of the American Speech-Language-Hearing Association, New Orleans, LA.

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

McGowan, R. S., & Nittrouer, S. (1988). Differences in fricative production between children and adults: Evidence from an acoustic analysis of /ʃ/ and /s/. *Journal of the Acoustical Society of America, 83*, 229–236.

McSweeny, J. L., & Shriberg, L. D. (1995). *Segmental and suprasegmental transcription reliability* (Tech. Rep. No. 2). Phonology Project, Waisman Center on Mental Retardation and Human Development, University of Wisconsin–Madison.

Menon, K. M. N., Jensen, P. J., & Dew, D. (1969). Acoustic properties of certain VCC utterances. *Journal of the Acoustical Society of America, 46*, 449–457.

Miccio, A. W., Forrest, K., & Elbert, M. (1996). Spectra of voiceless fricatives produced by children with normal and disordered phonologies. In T. Powell (Ed.), *Pathologies of speech and language: Contributions of clinical phonetics and linguistics* (pp. 223–236). New Orleans, LA: International Clinical Phonetics and Linguistics Association.

Milenkovic, P. (1996). CSpeech (Version 4) [Computer Program]. Madison, WI: University of Wisconsin–Madison, Department of Electrical Engineering.

Miller, J. L., Grosjean, F., & Lomanto, C. (1984). Articulation rate and its variability in spontaneous speech: A reanalysis and some implications. *Phonetica, 41*, 215–225.

Morrison, J. A., & Shriberg, L. D. (1992). Articulation testing versus conversational speech sampling. *Journal of Speech and Hearing Research, 35*, 259–273.

Nellhaus, G. (1968). Head circumference from birth to eighteen years. *Pediatrics, 41*, 106–114.

Nittrouer, S. (1995). Children learn separate aspects of speech production at different rates: Evidence from spectral moments. *Journal of the Acoustical Society of America, 97*, 520–530.

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.

Pentz, A. (1996, November). *Developmental changes in school-age children's fricative spectra: Preliminary findings*. Poster session presented at the Annual Convention of the American Speech-Language-Hearing Association, Seattle, WA.

Pentz, A., Gilbert, H. R., & Zawadzki, P. (1979). Spectral properties of fricative consonants in children. *Journal of the Acoustical Society of America, 66*, 1891–1893.

Schwartz, M. F. (1968). Identification of speaker sex from isolated, voiceless fricatives. *Journal of the Acoustical Society of America, 43*, 1178–1179.

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

Shriberg, L. D. (1986). *PEPPER: Programs to examine phonetic and phonologic evaluation records*. Hillsdale, NJ: Lawrence Erlbaum.

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

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

Shriberg, L. D., & Kent, R. D. (1995). *Clinical phonetics* (2nd ed.). Boston, MA: Allyn & Bacon.

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

Shriberg, L. D., Kwiatkowski, J., & Gruber, F. A. (1994). Developmental phonological disorders II: Short-term speech-sound normalization. *Journal of Speech and Hearing Research, 37*, 1127–1150.

Shriberg, L. D., & Lof, G. L. (1991). Reliability studies in broad and narrow phonetic transcription. *Clinical Linguistics and Phonetics, 5*, 225–279.

Soli, S. D. (1981). Second formants in fricatives: Acoustic consequences of fricative-vowel coarticulation. *Journal of the Acoustical Society of America, 70*, 976–984.

Strevens, P. (1960). Spectra of fricative noise in human speech. *Language and Speech, 3*, 32–49.

Syrdal, A. K., & Gopal, H. S. (1986). A perceptual model of vowel recognition based on the auditory representation of American English vowels. *Journal of the Acoustical Society of America, 79*, 1086–1100.

Tjaden, K., & Turner, G. S. (1997). Spectral properties of fricatives in amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research, 40*, 1358–1372.

Weismer, G. (1984). Acoustic analysis strategies for the refinement of phonological analysis [Monograph]. *Asha, 22*, 30–52.

Weismer, G., & Bunton, K. (1998). *Influence of pellet markers on speech production behavior: Acoustical and perceptual measures*. Manuscript submitted for publication.

Wilson, D. K. (1987). *Voice problems of children* (3rd ed.). Baltimore, MD: Williams & Wilkins.

Yeni-Komshian, G., & Soli, S. D. (1981). Recognition of vowels from information in fricatives: Evidence of fricative-vowel coarticulation. *Journal of the Acoustical Society of America, 70*, 966–975.

Received May 26, 1998

Accepted October 1, 1998

Contact author: Lawrence D. Shriberg, the Phonology Project, Waisman Center on Mental Retardation and Human Development, University of Wisconsin–Madison, 1500 Highland Avenue, Madison, WI 53705. Email: shriberg@waisman.wisc.edu

Acoustic Characteristics of /s/ in Adolescents

Peter Flipsen, Jr., Lawrence Shriberg, Gary Weismer, Heather Karlsson, and Jane McSweeny

J Speech Lang Hear Res 1999;42;663-677