



All Faculty Publications

2005-08-01

Sex-Related Acoustic Changes in Voiceless English Fricatives

Shawn L. Nissen
Shawn_Nissen@byu.edu

Robert Allen Fox

Follow this and additional works at: <https://scholarsarchive.byu.edu/facpub>

 Part of the [Communication Sciences and Disorders Commons](#)

Original Publication Citation

Fox, R. A., & Nissen, S. L. (25). Sex-Related Acoustic Changes in Voiceless English Fricatives. *Journal of Speech, Language, and Hearing Research*, 48, 753-765.

BYU ScholarsArchive Citation

Nissen, Shawn L. and Fox, Robert Allen, "Sex-Related Acoustic Changes in Voiceless English Fricatives" (2005). *All Faculty Publications*. 358.

<https://scholarsarchive.byu.edu/facpub/358>

This Peer-Reviewed Article is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Faculty Publications by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

Sex-Related Acoustic Changes in Voiceless English Fricatives

Robert Allen Fox

The Ohio State University, Columbus

Shawn L. Nissen

Brigham Young University, Provo, UT

This investigation is a comprehensive acoustic study of 4 voiceless fricatives (/f θ s ʃ/) in English produced by adults and pre- and postpubescent children aged 6–14 years. Vowel duration, amplitude, and several different spectral measures (including spectral tilt and spectral moments) were examined. Of specific interest was the pattern of normal development of the acoustic properties of fricatives and the nature of sex-specific patterns of fricative articulation in prepubescent children. Little evidence of amplitude or duration differences was found between speakers that was related to sex of the speaker. However, significant sex-specific differences in fricative articulation were found in all groups of speakers—even in the youngest children (ages 6–7 years)—although there was an indication that some of the acoustic differences between females and males is reduced or absent in the youngest children. Results from discriminant analysis demonstrated that a discriminant function based on the adult male tokens was generally better at classifying fricatives produced by male speakers than female speakers, regardless of age. This showed that sex-related differences (presumably a function of sex-linked vocal tract variation) were present even in the youngest speaker group. However, the classification accuracy of the female model showed a steady improvement with the increased age of the female speakers and may provide support for the claim that sex-related developmental differences may just be emerging in the youngest age group.

KEY WORDS: fricatives, fricative development, phonetic development, sex differences, speech production, spectral moments

It has been well documented that the acoustic characteristics of speech vary systematically as a function of the sex of the speaker. These differences are in part the result of the physical attributes typically exhibited by male and female speakers. Anatomical dimensions, such as vocal tract size and shape, as well as vocal fold length, have been found to be factors in determining the acoustic structure of an individual's speech. For example, it has been demonstrated that the lower overall fundamental frequency of typical adult male speakers is associated with their relatively larger larynx (Whiteside, 1996). In addition, studies have clearly shown that the dissimilarities in vowel formant patterns between male and female speakers are due principally to differences in vocal tract length (e.g., Fant, 1960).

Although sex-specific differences in vocal tract size are widely acknowledged in adults, the chronological age at which such distinctions develop in children remains somewhat controversial. Due to the health risks of radiographic studies and the relative expense of early magnetic resonance imaging, earlier studies investigating the morphology of the human vocal tract were often conducted on a relatively small number of adult participants (Baer, Gore, Gracco, & Nye, 1991; Dang, Honda, & Suzuki, 1994; Moore, 1992; Story, Titze, & Hoffman, 1996). However, in

recent years a much greater amount of morphometric data has been collected regarding the vocal tract morphology of children.

A study by Fitch and Giedd (1999), involving a relatively large number of children ($N = 129$), reported no significant anatomical sex differences in the vocal tract of younger speakers. In prepubescent children 10 years of age or younger, the authors found no significant differences in vocal tract morphology (i.e., vocal tract length, oral and pharyngeal cavity volume). According to their findings, sexual vocal tract dimorphism for most children occurs at the peripubertal and postpubertal stages of development.

Despite the strong association between an adult speaker's physical attributes and the acoustic realization of their speech, sex-specific variation cannot be entirely explained by anatomical differences alone. An early finding by Mattingly (1966), based on a reanalysis of the Peterson and Barney (1952) data on vowel formants, found that differences in formant frequencies that exist between adult male and female speakers is not accounted for solely by differences in vocal tract size. Evidence from additional studies has supported Mattingly's conclusion that a portion of the sex-linked acoustic variability of adults is associated with stylistic or behavioral factors (Sachs, Lieberman, & Erickson, 1973). This claim has been further advanced by research indicating that women may exhibit longer utterance durations, slower articulation rates, and less phonetic reduction than men (Byrd, 1994; Swartz, 1992; Whiteside, 1996).

Considering these findings, it is reasonable to postulate that men and women learn and acquire differing patterns of articulation. Some researchers have suggested that as a process of acculturation "men and women modify their articulation of the same phonetic elements to produce acoustic signals that correspond to the male-female archetype" (Sachs et al., 1973, p. 75). If acoustic variability is in part due to behavioral factors, it is interesting to determine at what age individuals begin to acquire such sex-specific patterns of articulation.

Several studies investigating the acoustic characteristics of vowel formant frequencies have found sex-specific differences in the speech of prepubescent children (Busby & Plant, 1995; Lee, Potamianos, & Narayanan, 1999; Perry, Ohde, & Ashmead, 2001; Sachs et al., 1973; Whiteside, 2001). In a large acoustic study involving 436 children between the ages of 5 and 17 years, Lee et al. (1999) found that the F1 values for vowels produced by young children varied systematically as a function of the child's sex. The authors concluded that these differences could not be solely attributed to physiological differences, but were probably the result of "social and psychological factors."

A reexamination of these data by Whiteside (2001) using formant scaling (Fant, 1966) and Bark transformations (Traunmüller, 1988, 1990) provided additional support for this conclusion, specifically, that additional factors besides vocal tract morphology were responsible for the sex-specific patterns in the children's acoustic data. Whiteside postulated that the emergence of "learned sex-specific speech behaviors" were partially responsible for the reported differences. In addition, an experiment examining the voice onset time (VOT) patterns in the stop articulations of pre-adolescent children found marked sex-linked differences by the age of 11 years (Whiteside & Marshall, 2001). These differences were characterized by young girls exhibiting longer VOT values than boys of a similar chronological age. Interestingly, the gender-specific VOT patterns varied as a function of the place of the stop articulation, with greater differences reported for bilabial stops as compared to alveolar stop productions. The authors concluded that such disparities in VOT between male and female children were linked to developmental or learned factors acquired with age.

Findings from these studies in association with the morphometric data presented earlier (Fitch & Giedd, 1999) provide support for the premise that prepubescent children are beginning to acquire learned patterns of articulation for vowel segments and temporal features of speech (i.e., VOT) at a relatively young age. However, there remain gaps in the literature regarding the possibility of sex-specific differences in the acoustic patterns of fricative productions. An early study by Martin Schwartz (1968) found that adult female productions of /s/ and /ʃ/ generally tended to have higher frequency spectra than productions by male speakers of similar age. Using relatively new analytical techniques (e.g., dynamic analysis of spectral slices and spectral moments), Jongman, Wayland, and Wong (2000) also found significant sex-specific differences in the acoustic properties of fricatives produced by adults. Specifically, the authors reported that female speakers' productions exhibited higher values for spectral peak location, spectral mean, variance, and kurtosis, yet significantly lower measures of skewness. However, since Jongman and his colleagues used only adult speakers, it is unclear if these results will hold true in younger speakers. Thus, it is of particular interest to investigate whether similar sex-specific differences will extend to prepubescent children.

The work of Nitttrouer (1995) involved an examination of how the acoustic properties of speech segments change as a function of normal development/aging. One aim of Nitttrouer's study was to measure and compare the spectral moments of the voiceless sibilant fricative contrast (/s ʃ/) as produced by young children

and adults. For this particular contrast, significant differences were reported that might be related to articulatory changes that occur in developmental maturation. However, the study did not examine whether there were similar differences in other places of fricative articulation (i.e., /f θ/). Also, although Nittrouer found sex differences in the adult speakers, she did not examine sex differences in the younger speakers and how the sex differences develop as a function of maturation. Thus, the general purpose of this investigation is a comprehensive acoustic study of four places of English voiceless fricatives produced by adults and pre- and postpubescent children. Specifically, this research is intended to address the following questions:

1. What is the pattern of normal development in the acoustic properties of fricatives, as characterized by selected acoustic measures?
2. Are sex-specific patterns of fricative articulation present in the speech of prepubescent children and at what age might possible sex differences emerge? Would such differences be similar to those found by Jongman et al. (2000) and Nittrouer (1995) in adult speakers?
3. To what extent will a segment classification model based on the acoustic characteristics of fricatives produced by a sex-specific group of adult speakers successfully categorize the fricatives of younger speakers?

Method

Speakers

The primary study involved both male ($N = 50$) and female ($N = 50$) speakers, all of whom spoke a central Ohio version of Standard American English. The age range of this group was 6–52 years ($M = 14$ years, 4 months, $SD = 10$ years, 5 months). The speakers were organized into five different age groups with 10 male and 10 female speakers in each group (age means and standard deviations for younger groups are expressed in months): 6–7 years (females: $M = 84.0$, $SD = 6.8$; males: $M = 79.2$, $SD = 6.8$), 8–9 years (females: $M = 114.1$, $SD = 5.8$; males: $M = 111.8$, $SD = 3.4$), 10–12 years (females: $M = 139.7$, $SD = 10.3$; males: $M = 144.7$, $SD = 6.0$), 13–14 years (females: $M = 167.4$, $SD = 6.9$; males: $M = 169.2$ years, $SD = 6.9$), and adult (females: $M = 29.8$ years, $SD = 14.1$ years; males: $M = 32.4$ years, $SD = 16.3$ years). All speakers (and, for children, their parents) reported no speech, language, or hearing disorders. The absence of articulatory or phonological errors in the collected speech tokens was verified by a certified speech-language pathologist. All children had front incisors

(so there were no problems with physiologically induced lips) and the youngest group of children still had their baby teeth.

Recording Procedures

All speech samples were collected using laptop computers with high-quality microphones and specially designed computer programs.¹ All samples were recorded in a quiet room directly onto a hard disk drive using a head-mounted microphone positioned 1–2 in. from the speaker's lips. A computer program (written in C++) was used to collect the samples. In all cases, speakers saw a written version of the stimulus token on a computer screen and spoke the word. An experimenter monitored each production and determined whether the token was an acceptable version of the stimulus. The speaker was asked to repeat the token until an acceptable version was produced. Tokens were recorded using a 22.05 kHz sampling rate at 16-bit resolution and were low-pass filtered at 11.025 kHz.

The consonants of interest in this study included voiceless fricatives (/f θ s ʃ/) in a syllable initial /CV/ context. All speakers produced the following words (four different tokens for each consonant): *fall*, *fiber*, *follows*, *four*, *Thebes*, *thick*, *thieves*, *thing*, *sing*, *sick*, *sun*, *salt*, *ship*, *shape*, *she*, and *sheep*. Vowel quality in the stimulus set was not controlled; thus, spectra for a given place of articulation were pooled across tokens. As discussed later, interpreted results will be drawn from acoustic characteristics measured at the fricative midpoint (where acoustic parameters should be relatively immune to the effects of stimulus onset variations and the formant transitions into the following vowel).

Analysis Procedures

Location of acoustic landmarks in each of the stimulus tokens was determined using a waveform editing program (CoolEdit) and acoustic analyses (of segment durations, amplitudes, and spectra) were then completed using specialized MATLAB programs (written by one of the authors). Experimenters determined the location of the following landmarks (as appropriate) in each token containing a fricative: fricative onset, fricative offset, voicing onset, and vowel offset. Fricative onset was identified as the start of frication noise, and fricative offset was identified as the cessation of such

¹The stimulus tokens were part of a very large spoken database created by R. Fox in collaboration with Dr. Reiko Akahane-Yamada from the Advanced Telecommunication Research (ATR) Human Information Science Laboratories in Japan. This database contains speech tokens recorded from over 350 speakers. Speakers were selected on the basis of age, dialect, and overall clarity of the recordings.

noise. Voicing onset following the production of the voiceless fricative was identified as the start of quasi-periodic patterns in the waveform regardless of whether the fricative noise was absent. The onset of the vowel was set to the value of the fricative offset.

The experimenters recorded the values for each location (in milliseconds) into a file. All segmentation decisions were later checked and corrected (and then rechecked) using a MATLAB program that displayed the segmentation marks superimposed over a display of the token's waveform (in two different views: a view that included the entire token and an expanded view that concentrated primarily on the segment of interest). All subsequent acoustic analyses described below were completed at these locations.

Of particular interest in this study is a spectral moment analysis of the fricatives (see Forrest, Weismer, Milenkovic, & Dougall, 1988, and Jongman et al., 2000, for a general outline of this analysis procedure). Fast Fourier transforms (FFTs) were calculated using full Hamming windows with 98% pre-emphasis, and each FFT was treated as a probability distribution from which the first four distribution moments could be calculated. This article concentrates on the first (mean), third (skewness), and fourth (kurtosis) moments. In the current study, the second moment (variance) provided little insight into differences among the consonants or the speaker groups. The spectral moments discussed here were calculated from linear spectra only. However, the pattern of the results (in particular, consonant, age group, and sex group differences) was similar to the patterns found in moments calculated from bark-transformed spectra.

Following the procedure used by Jongman et al. (2000), spectral moments for each individual fricative were calculated with a 40 ms Hamming window at four different locations: (a) the first 40 ms, (b) the central 40 ms, (c) the final 40 ms, and (d) with the window centered over the fricative offset. Because some of the word-initial fricatives were extremely short, spectral moments for these tokens were calculated only for fricatives that were at least 60 ms in duration. This resulted in elimination of 31 tokens, broadly distributed across speakers. The spectral moment measures were extracted using 1,024-point FFTs with zero-padding. The spectral peak of the fricative was also determined; it represented the highest amplitude peak of the FFT spectrum. A final spectral measure calculated was spectral slope. Specifically, this parameter was the slope of the amplitude (FFT) spectrum between 1 and 11 kHz calculated on the basis of linear regression of the FFT values against their corresponding frequency values. In addition, root-mean-squared (RMS) amplitude of the entire fricative was calculated along with its normalized amplitude in decibels. Normalized ampli-

tude represented the difference between the amplitude of the fricative (in decibels) and the amplitude of the following vowel (in decibels) over the 40 ms of the vowel exhibiting the greatest amplitude. This 40 ms window was centered over the peak amplitude location in the vowel; if the start of the window located in this manner started before vowel onset, the window was relocated so that leading edge of the window was coterminous with the vowel onset. The trailing edge of the analysis window was never located past the vowel (token) offset. The duration of the fricative and following vowel were also calculated.

Discriminant Analysis

Calculation of the acoustic measurements described above will allow us to determine how different acoustic variables may vary as a function of speaker age, speaker sex, and fricative place of articulation. However, none of these analyses will allow a determination of how well combinations of acoustic parameters (such as fricative duration, vowel duration, normalized fricative amplitude, spectral slope, spectral peak, spectral mean, spectral skewness, and spectral kurtosis) will allow for correct place of articulation categorization of the fricatives; for this purpose we use discriminant analysis (DA). To allow relatively direct comparisons to the discriminant results of Jongman et al. (2000), classification models were first calculated using only the adult data (three different models were created using all adults, female adults, and male adults). The female adult and male adult models were then used to classify the place of articulation of the fricatives for all of the children's data.

Results

Acoustic Measures

Below is a description of the pattern of results for each of the acoustic measures examined. Unless otherwise indicated, mixed design analyses of variance (ANOVAs) were used, with age and sex as between-subjects factors and place of articulation as the within-subjects factor. Reported probability levels for F tests reflect Greenhouse-Geisser adjustments for potential violations of the sphericity assumption. Results of significant F tests will include a measure of effect size: partial eta squared, or η^2 (in general, partial eta squared can be considered a measure of the proportion of variance explained by a dependent variable when controlling for other factors).

Fricative duration. The first acoustic measure examined was the duration of the fricative. An ANOVA

revealed a significant difference in fricative duration as a function of place of articulation, $F(3, 270) = 135.64, p < .001, \eta^2 = 0.60$. The mean fricative length was longer for the sibilants (mean duration of /s ʃ/ was 199 and 207 ms, respectively) than for the nonsibilants (mean duration was 142 and 139 ms for /f θ/, respectively). Paired-sample t tests with Bonferroni adjustments² showed that all pairwise comparisons for duration between sibilant (/s ʃ/) and nonsibilant (/f θ/) fricatives were significant at the .001 level. However, the duration comparisons between /s/ and /ʃ/, as well as /f/ and /θ/, failed to reach significance. The range of mean duration measures (averaged across speakers and vowel context) was 68 ms, which is considerably greater than the range of values (15 ms) reported by Jongman et al. (2000). A possible explanation for such differences between studies may be the fact that the nature and context of the elicited speech samples were different. There was also a significant Place \times Sex of Speaker effect, $F(3, 270) = 3.33, p = .021, \eta^2 = 0.04$. Although the mean durations of /f s ʃ/ were similar for male and female speakers (136, 195, and 207 ms, respectively, for females; 147, 204, and 208 ms, respectively, for males), the female speakers had significantly shorter /θ/ fricatives than did male speakers (125 ms for females and 153 ms for males), one-tailed $t(98) = 2.771, p = .003$. No other main or interaction effects were significant, although the age effect approached significance ($p = .063$). Younger speakers had longer fricative durations than did older speakers (e.g., mean duration for the 6–7 year old speakers was 183 ms while the mean for adults was 153 ms).

Amplitude measures. As one would expect, the normalized and RMS amplitude of the sibilant fricatives were significantly greater than those of the nonsibilant fricatives. An ANOVA revealed that the only significant result was a main effect of place of articulation for both normalized amplitude, $F(3, 270) = 584.77, p < .001, \eta^2 = 0.87$, and RMS amplitude, $F(3, 270) = 495.91, p < .001, \eta^2 = 0.85$. Paired-sample Bonferroni-adjusted t tests showed that all pairwise comparisons for RMS amplitude (48.40, 50.89, 62.21, and 62.36 dB for /f θ s ʃ/, respectively) and normalized amplitude (–24.31, –20.13, –9.36, and –7.37 dB for /f θ s ʃ/, respectively) between nonsibilant and sibilant fricative types were significant at the .001 level; only the comparisons between /s/ and /ʃ/ failed to reach significance.

Spectral Measures

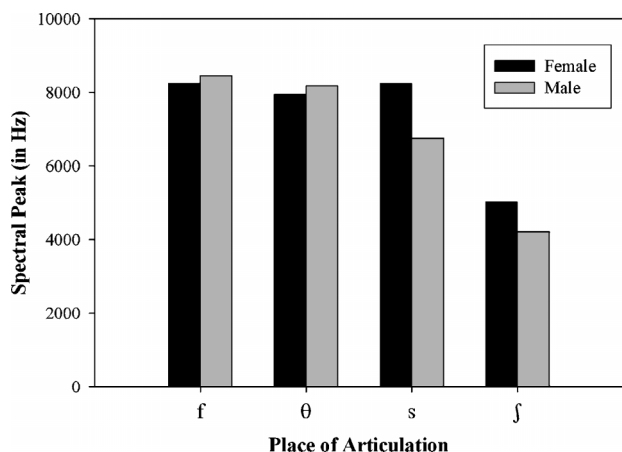
Although there were four different analysis window locations to choose from, here only the spectral moments

²Tukey pairwise comparisons are used as post hoc tests for between-subject factors, and Bonferroni-adjusted paired-sample t tests were used as post hoc tests for within-subject factors.

of the mid-fricative location were analyzed, as we are interested in the nature of the fricative noise itself. This choice reduced and/or eliminated the influence of vowel quality differences between different CV tokens. It is important to note that the patterns of differences among all the spectral measures as a function of analysis window closely matched the patterns found by Jongman et al. (2000).

Spectral peak. The first spectral measure examined was spectral peak. An ANOVA showed a significant main effect of the sex of the speaker—the mean spectral peak of the female speakers was higher than that of the male speakers (7365 and 6895 Hz, respectively), $F(1, 90) = 16.12, p < .001, \eta^2 = 0.15$. There was also a significant main effect of place of articulation, $F(3, 270) = 316.19, p < .001, \eta^2 = 0.78$, with mean spectral peak decreasing as the place of articulation moved back in the oral cavity (mean spectral peaks for /f θ s ʃ/ were 8348, 8056, 7497, and 4518 Hz, respectively). Paired-comparison Bonferroni-adjusted t tests showed that only the comparison between /f/ and /θ/ failed to reach significance at the .05 level. The main effect of age group was significant, $F(4, 90) = 4.41, p = .003, \eta^2 = 0.16$. Tukey pairwise comparisons indicated that the spectral peaks of the 6–7-year-old and 8–9-year-old groups were significantly higher (at the .05 level) than were the peaks of the other groups (the mean peaks for the 6–7-year-old, 8–9-year-old, 10–12-year-old, 13–14-year-old, and adult groups were 7367, 7384, 6851, 7229, and 6822 Hz, respectively). These significant main effects should be interpreted in light of the two significant interactions. As shown in Figure 1, there was a significant Place \times Speaker Sex interaction, $F(3, 270) = 18.56, p < .001, \eta^2 = 0.17$, reflecting the fact that spectral peak means for female and male speakers were similar for /f θ/ but significantly different for /s/, $t(98) = 7.88, p < .001$, and /ʃ/, $t(98) = 3.45, p < .001$. There was also a significant Place \times Age interaction, $F(12, 270) = 2.40, p = .006, \eta^2 = 0.10$. As shown in Figure 2, the difference in spectral peaks as a function of age became progressively greater as the place of articulation moved posterior with the greatest age differences found for /ʃ/. Separate univariate analysis of the spectral peaks for each place of articulation revealed a minimally significant effect of age for /s/, $F(4, 90) = 2.54, p = 0.045, \eta^2 = 0.10$, but a very significant effect of age for /ʃ/, $F(4, 90) = 8.09, p < .001, \eta^2 = 0.26$. Tukey pairwise comparisons of the /ʃ/ data indicated that the spectral peaks of the 6–7-year-old and 8–10-year-old groups were significantly higher (at the .05 level) than were the peaks of the other groups (the mean peaks for /ʃ/ the 6–7-year-old, 8–9-year-old, 10–12-year-old, 13–14-year-old, and adult groups were 5493, 5159, 4270, 4163, and 4002 Hz, respectively). There were no significant age effects for spectral peaks for /f/ or /θ/. No other effects were significant.

Figure 1. Mean spectral peaks of fricatives by place of articulation for female and male speakers.



Spectral slope. The spectral slope (from 1 to 11 kHz) is not a measure that was used in either the Jongman et al. (2000) or Nittrouer (1995) studies; however, as we shall see, it provides a very useful and effective parameter in the DAs to be discussed later. An ANOVA showed a significant main effect of sex of the speaker, $F(1, 90) = 4.71, p = .033, \eta^2 = 0.05$, with female speakers having slightly higher slope values than male speakers (1.2 and 1.1, respectively). There were also significant spectral slope differences as a function of place of articulation, $F(3, 270) = 282.95, p < .001, \eta^2 = 0.76$, but there was no significant effect of age. However, the Place \times Age interaction was significant, $F(12, 270) = 3.54, p < .001, \eta^2 = 0.14$. The spectral slopes for the youngest group of speakers (6–7-year-olds) were lower than those for all other speaker groups for /s/, but the

Figure 2. Plot of mean spectral peaks of fricatives by place of articulation and age group.

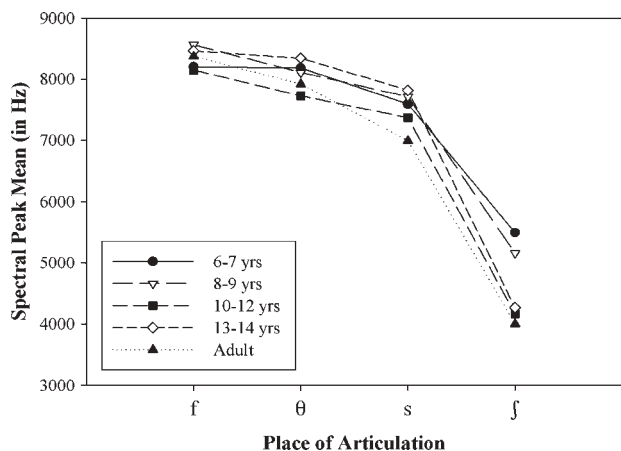
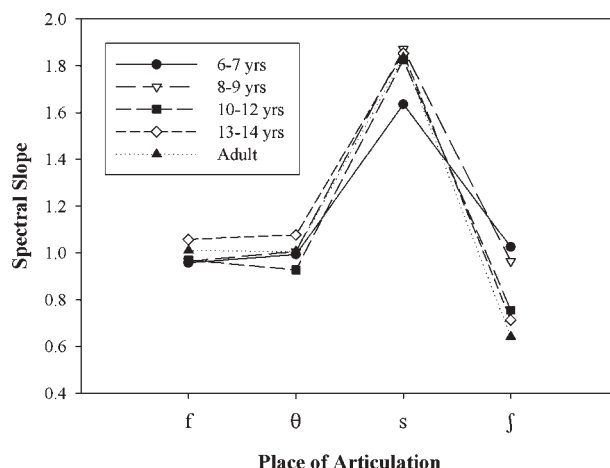


Figure 3. Mean spectral slopes as a function of fricative place of articulation and speaker age group.



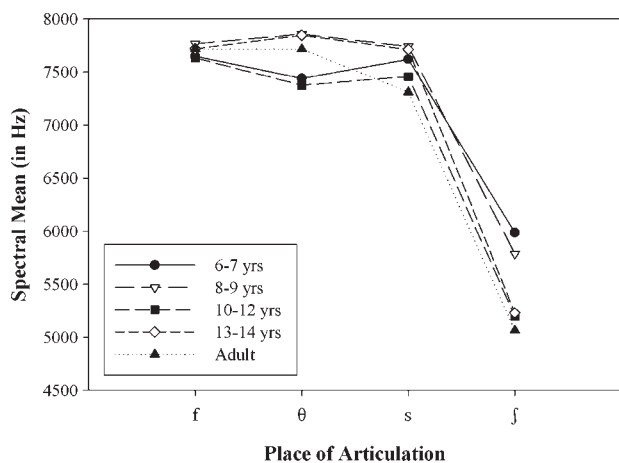
slopes for the two youngest groups of speakers (6–7-year-olds and 8–9-year-olds) were significantly higher than the older speaker groups for /ʃ/ (Tukey pairwise comparisons at the .05 level), as shown in Figure 3.

The Place of Articulation \times Sex of the Speaker, $F(3, 270) = 14.66, p < .001, \eta^2 = 0.14$, and the Place \times Age \times Sex of the Speaker, $F(12, 270) = 2.48, p = .004, \eta^2 = 0.10$, interactions were also significant. Overall, the female speakers had significantly higher spectral slopes than male speakers for /s/, $t(98) = 3.94, p < .001$, and /ʃ/, $t(98) = 3.84, p < .001$, but nonsignificantly lower slopes for /f θ/. However, this pattern did not hold for the youngest speakers (6–7-year-olds), with male speakers having nonsignificantly higher slopes than the female speakers for all but /ʃ/.

Spectral mean. Turning now to analysis of spectral means, we find, as expected, a significant main effect of place of articulation, $F(3, 270) = 288.45, p < .001, \eta^2 = 0.76$. The spectral mean of /ʃ/ was significantly lower (using Bonferroni-adjusted t tests at the .001 level) than the spectral mean of each of the other three fricatives (/f θ s/), but unlike Jongman et al. (2000), we found no significant differences among these three fricatives. There was also a main effect of the sex of the speaker, $F(1, 90) = 7.99, p = .006, \eta^2 = .08$, with the overall spectral mean higher for female speakers than for male speakers (7140 and 6910 Hz, respectively).

There was a significant main effect of age group, $F(4, 90) = 3.22, p = .016, \eta^2 = 0.12$. In general, the spectral means are lower for older speakers than for younger speakers, but this difference was most evident (in the expected direction) within the sibilants. Tukey pairwise comparisons showed a significant difference between the youngest speakers (6–7- and 8–9-year-olds) and the adult speakers at the .05 level. There was also a

Figure 4. Spectral means by fricative place of articulation and speaker age group.

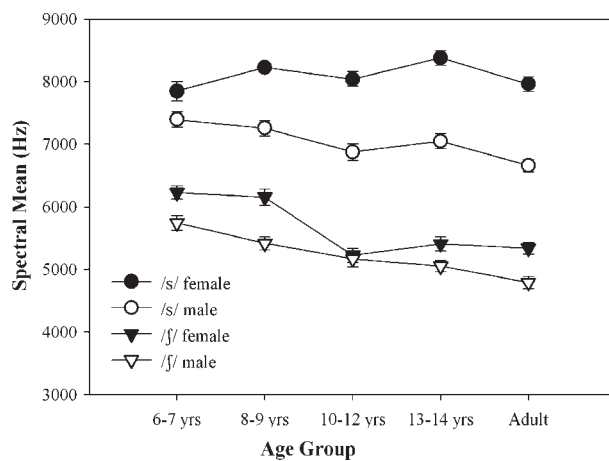


significant Place of Articulation \times Age Group interaction, $F(12, 270) = 2.32, p = .008, \eta^2 = 0.09$. This interaction stems from the fact that the adult speakers had high spectral means for /f θ/ but had the lowest spectral means for /s ʃ/. The basis for these effects is evident in Figure 4. Following the observed age trend for higher spectral means in younger children, we should expect that the spectral means will be highest in very young children. This was precisely the result described by Nissen (2003), who found the highest spectral means in the youngest children (aged 3–4 years).

There was also a significant Place \times Sex of Speaker interaction, $F(3, 270) = 22.1, p < .001, \eta^2 = 0.20$. Although the spectral means for the female and male speakers were similar in the nonsibilants /f/ (7592 and 7801 Hz, respectively) and /θ/ (7648 and 7647 Hz, respectively), Bonferroni-adjusted t tests showed that the female speakers had significantly higher spectral means than did the male speakers for both /s/ (8091 and 7047 Hz, respectively) and /ʃ/ (5671 and 5235 Hz, respectively) at the .05 level.

To make appropriate comparison to Nittrouer (1995), data from the /s ʃ/ fricatives were analyzed separately. The only notable difference was that in these data (but not in Nittrouer's) there was a significant Place \times Age \times Speaker Sex effect, $F(4, 90) = 2.37, p = .050, \eta^2 = 0.10$, as shown in Figure 5. As a general rule, older speakers had lower spectral means for /ʃ/ than did younger speakers. Tukey pairwise comparisons showed that spectral means for /ʃ/ were significantly higher for the 6–7- and 8–10-year-olds than for the other three groups of speakers at the .05 level, no doubt reflecting expected differences in vocal tract size. However, although there was a similar age-related decline for /s/ in male speakers (younger speaker groups

Figure 5. Spectral means for /s ʃ/ as a function of speaker age group and speaker sex.

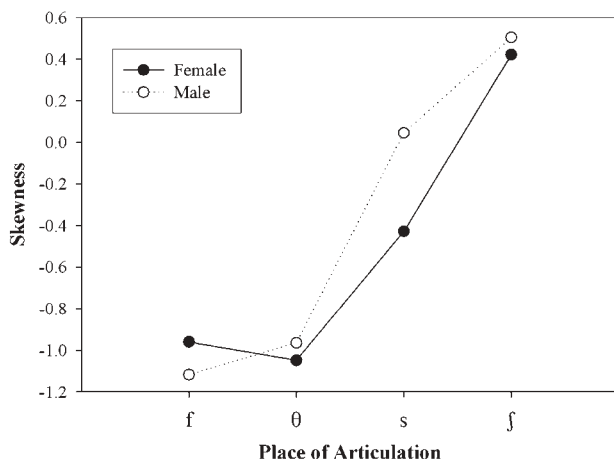


had consistently higher spectral means than older speaker groups, with Tukey pairwise comparisons showing that the youngest group of speakers had significantly higher spectral means than the adults, at the .05 level), the spectral mean for female speakers remained high for all age groups (with no significant differences obtained between age groups).

Spectral skewness. There was a main effect of age group, $F(4, 90) = 3.85, p = .006, \eta^2 = 0.15$. Tukey pairwise comparisons indicated that the skewness values for the two youngest groups (6–7-year-olds = -0.835 and 8–9-year-olds = -0.547) were significantly lower than those of two of the older groups (10–12-year-olds = -0.273 and adults = -0.3630) at the .05 level (skewness for the 13–14-year-olds was -0.468). In Nissen (2003), the 3–4-year-olds had the lowest skewness values of all age groups. The Place \times Age interaction was not significant and indicated that the skewness differences among these four fricatives remained the same through the development period addressed here.

As is illustrated in Figure 6, there was a significant main effect of sex of speaker, $F(1, 90) = 4.56, p = .035, \eta^2 = 0.05$ —overall, male speakers had higher skewness ($M = -.375$) than did female speakers ($M = -.441$). However, there was also a significant Place \times Sex of Speaker interaction, $F(3, 270) = 7.91, p < .001, \eta^2 = 0.08$. Spectral skewness for female speakers was significantly lower than for male speakers in the production of /f/, $t(98) = 2.14, p = 0.035$, but significantly higher than for male speakers in the production of /s/, $t(98) = -4.493, p < .001$. Bonferroni-adjusted t tests also showed that although there was no significant difference between /f θ/ for either male or female speakers, each of these nonsibilants was significantly different from both /s/ and /ʃ/ (/s/ and /ʃ/ were also significantly different from each other) at the .05 level for both speaker sexes.

Figure 6. Spectral skewness as a function of fricative place of articulation and speaker sex.



In this study, both spectral skewness and spectral mean were entered into the DA as independent measures. However, one should recognize that these two measures are not necessarily independent. In particular, when using a fixed analysis bandwidth, as the concentration of acoustic energy moves up in frequency, the asymmetry of the spectrum will likely become more negatively skewed—this is the pattern found for spectral skewness in the youngest versus oldest speaker. Although this pattern is not always found in data reported here—for example, spectral skewness was significantly higher for female speakers than for male speakers in the production of /s/—the potential intercorrelations between spectral skewness and spectral mean should be considered when interpreting our results.

Spectral kurtosis. The final spectral measure examined was the fourth moment of the distribution (kurtosis). An ANOVA of the kurtosis values revealed a significant place of articulation effect, $F(3, 270) = 27.65, p < .001, \eta^2 = 0.23$. Mean kurtosis values for /f θ s ʃ/ were 1.83, 1.45, 0.69, and -0.21 , respectively. Bonferroni-adjusted t tests showed that both sibilant fricatives were significantly different from both nonsibilants and from one another at the .05 level. As was found in Jongman et al. (2000) and Nittrouer (1995), /s/ had a higher kurtosis than did /ʃ/. However, the kurtosis value obtained for /s/ was more similar to those obtained by Nittrouer (who obtained a mean value of 0.95 across all speaker groups) than Jongman et al. (whose kurtosis values for /s/ ranged from a little below 2 to a little above 3.5). The values obtained in this study for /f θ/ were somewhat higher than those obtained by Jongman et al., although the pattern of the nonsibilant fricatives as compared to /ʃ/ were similar. There were no significant effects due to either age or sex of the speaker. Of the interaction effects, only the Place \times Sex of the Speaker

interaction was significant, $F(3, 270) = 8.24, p < .001, \eta^2 = 0.08$. Bonferroni-adjusted t tests show that the kurtosis values for /s/ were significantly higher for female speakers than for male speakers at the .05 level (mean kurtosis was 1.51 and -0.14 , respectively) but not for the other fricatives.

Discriminant Analysis (DA)

To allow relatively direct comparison to the study by Jongman et al. (2000), the acoustic parameters described above were analyzed using a series of linear DAs using SPSS. DA is an application that derives a set of linear equations (discriminant functions) that are designed to classify individual cases (fricatives) into one of several mutually exclusive groups (place of articulation) based on their values for a set of predictor variables (acoustic measures). In all cases, the analysis method used here selects predictor variables on a stepwise basis in terms of which of the available acoustic variables will significantly minimize Wilks's Λ . When all such variables have been selected, the stepwise procedure stops—this leaves out of the discriminant functions those variables that do not significantly contribute (at the .01 level as determined by an F test) to the discrimination among fricatives. A second use of Wilks's Λ is in an overall chi-square test of significance of the analysis. All the DA results discussed below were significant at the .01 or better according to this test of significance.

After all the predictor variables were selected, the salience or strength of each predictor variable entered into the discriminant functions was determined through an analysis of the standardized canonical discriminant function coefficients (SCDFs; see Klecka, 1980). These coefficients indicate the partial contribution of each predictor variable selected to the discriminant functions, controlling for the other independent variables entered into the equations. In general, the larger the coefficient, the greater the contribution of that variable to the discrimination among fricatives. Also calculated were the structure coefficients (SCs), which represent a correlation between a predictor variable and the discriminant functions (providing an indication of how closely the predictor variable and the discriminant function are related). Finally, we calculated the parallel discriminant ratio coefficients (DRCs), the product of corresponding SCDFs and SCs, which provide a indication of the relative importance of the variable to the discriminant functions (Thomas, 1992).

DA for each separate age group and gender group as well as the combined dataset was done using eight acoustic variables available for entry as predictor

Table 1. Results of stepwise linear discriminant analysis for self-classified sets of data (using the jackknife procedure). All values represent percentage correct classification.

Groups Self-Classified		% Fricatives Correctly Classified				
		Overall	/f/	/θ/	/s/	/ʃ/
Group 1 6–7 years	All speakers	65.4	63.9	42.4	67.5	83.8
	Female speakers	62.3	68.6	16.1	75.0	80.0
	Male speakers	67.1	70.3	48.6	62.5	85.0
Group 2 8–9 years	All speakers	78.7	77.6	54.7	86.3	91.3
	Female speakers	77.3	78.9	50.0	90.0	95.0
	Male speakers	79.3	73.7	43.8	95.0	97.5
Group 3 10–12 years	All speakers	76.6	65.3	53.7	90.0	92.5
	Female speakers	77.2	67.6	50.0	90.0	95.0
	Male speakers	76.7	71.4	42.9	97.5	100.0
Group 4 13–14 years	All speakers	73.8	55.6	61.4	82.5	92.5
	Female speakers	76.0	51.5	54.5	90.0	100.0
	Male speakers	79.5	74.4	59.5	82.5	100.0
Group 5 adults	All speakers	80.3	71.4	54.8	93.8	95.0
	Female speakers	80.0	75.7	50.0	90.0	95.0
	Male speakers	83.1	75.0	52.9	100.0	100.0

variables: fricative duration,³ vowel duration, normalized fricative amplitude, spectral slope, spectral peak, spectral mean, spectral skewness, and spectral kurtosis (values of the spectral variables were all calculated at the fricative midpoint). In each analysis, the classification results were based on the cross-validation (“jackknife”) procedure.⁴ Since fricatives shorter than 60 ms were not included in the analyses, the prior probabilities were computed from the number of fricatives of each type that were included.

The percentage of correct classifications for these “self-classified” results for each separate age group and gender are shown in Table 1 in the first data column. In general, the classifications of the fricatives were best for the adults (80.3% correctly classified)—a classification rate somewhat better than that obtained by Jongman et al. (2000; 77%) even though they used 21 predictor variables. The classification rates for the young age groups are somewhat worse and are poorest for the youngest children, age 6–7 years (65.4% correctly classified). Also shown in Table 1 are the classification rates for each separate fricative /f θ s ʃ/. For all age groups, the sibilant fricatives /s ʃ/ were correctly classified more often than were the nonsibilant fricatives /f θ/, a result also compatible with Jongman et al. It should be noted that the classification of even the sibilant fricatives is markedly poorer in the youngest

age group (only 62.5% of the /s/ fricatives were correctly classified for the youngest male speakers).

An indication of which predictor variables were “important” in the discriminant functions for each age group is shown in Table 2. For each of these DAs, there were three statistically significant discriminant functions (as measured by Wilks’s Λ). Since the third discriminant function contributed no more than 2% of the variance accounted for in any set of data, it is not reported here. Parallel DRCs can be considered a measure of the contribution of individual response variables to the discriminant function and, therefore, represent a measure of the importance of that variable to group discrimination (see Thomas, 1992, for an extended discussion). There is no single best decision metric to determine what value of DRC would indicate that it is important to the function. However, following Thomas, we will specify that a variable with a DRC value of 0.0625 ($1/2p$, where p = number of response variables) or less indicates that it is not an important predictor variable. Shown in Table 2 are the SCDFs, SCs, and DRCs for the most important predictor variables contributing to the first two discriminant functions that match this criterion. Although there are some differences among the different age groups, it is clear that the most important predictors, overall, are normalized fricative amplitude, spectral peak, spectral mean, spectral skewness, and spectral slope. This is compatible with Jongman et al. (2000), who found that spectral peak, two amplitude measures (normalized fricative amplitude and relative amplitude), and spectral mean classified the fricatives with reasonable accuracy. We

³All fricatives analyzed using DA were at least 60 ms in duration.

⁴Specifically, the SPSS cross-validation procedure was used in which each case is classified on the basis of the discriminant functions derived from all cases other than itself.

found that the variables vowel duration, fricative duration, and spectral kurtosis made little or no significant contribution to the place-of-articulation classifications.

Of special interest in this study are developmental differences in fricative production and the development of speaker sex differences as children mature. To address this question, a series of separate DAs was done in which the discriminant functions (the classification models) were based on the all adult model, the adult male model, and the adult female model. In this use of DA, there are two separate phases of analysis. First is the analysis phase, in which a “classification rule” (a set of discrimination functions) is developed using cases for which the category membership is known. These are the functions developed in the “self-classification” analyses described above. Second is the testing/classification phase that uses this classification rule to classify cases for which category membership is not known. We used two different classification rules to classify the children’s data: one derived using the adult male data and one derived using the adult female data.

Results of these DAs are shown in Table 3. One general pattern of interest that can be seen is that a

classification model based on data from adult female speakers is more successful at identifying the fricatives of the younger speakers than is a classification model based on adult male speakers. However, other patterns are also of interest. Note that the male adult model successfully categorizes a higher percentage of fricatives produced by male children than does the female adult model in all but the youngest age group. In addition, the adult male model is relatively poor at categorizing fricatives produced by male and female speakers in this youngest age group. The female adult model, on the other hand, is better than the adult male model at classifying the fricatives from both male and female children in the group (ages 6–7 years). However, this pattern changes for the next three age groups (which include children aged 8–14 years); here the adult female model is more successful at classifying the fricatives produced by female children than is the adult male model, but the reverse is true for fricatives produced by male children.

In the analysis of the spectral mean data, we found a significant Place \times Age \times Speaker Sex effect for the sibilant fricatives /s/ and /ʃ/. Likewise, the ability of the discriminant functions of both the female and male

Table 2. Summary of standard discriminant function coefficients (SDFCs), structure coefficients (SCs), and parallel discriminant ratio coefficients (DRCs) for each separate age group.

Training Set	First Discriminant Function			Second Discriminant Function				
	Input Variable ^a	SDFCs	SCs	DRCs ^b	Input Variable ^a	SDFCs	SCs	DRCs ^b
Group 1 6–7 years	FricAmp	0.700	0.718	0.503	SpSkew	0.845	0.920	0.777
	SpPeak	-0.414	-0.587	0.243	SpVar	-0.375	-0.536	0.201
	SpSlope	-0.458	-0.506	0.232				
	SpMean	0.598	0.312	0.187				
Group 2 8–9 years	SpMean	-1.010	-0.360	0.364	SpSlope	0.694	0.862	0.598
	FricAmp	0.471	0.665	0.313	SpMean	0.731	0.682	0.499
	SpSlope	0.821	0.262	0.215	FricAmp	-0.317	-0.061	0.193
	VowDur	-0.313	-0.225	0.070				
Group 3 10–12 years	FricAmp	0.583	0.705	0.411	SpSlope	0.869	0.910	0.791
	SpSkew	0.484	0.416	0.201	SpPeak	0.339	0.700	0.237
	SpSlope	0.583	0.303	0.177				
	SpPeak	-0.270	-0.338	0.091				
	VowDur	-0.309	-0.264	0.082				
Group 4 13–14 years	SpMean	0.744	0.591	0.440	SpSlope	0.871	0.895	0.780
	FricAmp	-0.548	-0.640	0.351	SpMn	0.291	0.610	0.178
	SpPeak	0.268	0.616	0.165	SpPeak	0.283	0.459	0.130
Group 5 adults	FricAmp	0.739	0.690	0.510	SpSlope	1.121	0.869	0.974
	SpPeak	-0.435	-0.548	0.238	SpPeak	0.188	0.390	0.073
	SpSkew	0.267	0.520	0.139				

Note. FricAmp = fricative amplitude; SpPeak = spectral peak; SpSlope = spectral slope; SpMean = spectral mean; VowDur = vowel duration; SpSkew = spectral skewness; SpVar = spectral variance.

^aVariables are ordered according to their relative contribution to each discriminant function based on the DRCs (≥ 0.625). ^bDRC = SDFC \times SC.

Table 3. Summary of fricative classification results for each age and gender group based on the female adult model or the male adult model. All values represent percentage correct classification.

Group Classified		Discriminant Functions Based on:	
		Female Adult Model	Male Adult Model
Group 1 6–7 years	All speakers	62.2	50.9
	Female speakers	57.5	46.3
	Male speakers	66.9	55.6
Group 2 8–9 years	All speakers	70.9	63.8
	Female speakers	68.8	53.8
	Male speakers	73.1	73.8
Group 3 10–12 years	All speakers	75.6	68.8
	Female speakers	78.5	65.0
	Male speakers	72.5	72.5
Group 4 13–14 years	All speakers	71.3	70.6
	Female speakers	72.5	66.3
	Male speakers	70.0	75.0
Group 5 adults	All speakers	—	—
	Female speakers	—	66.2
	Male speakers	69.5	—
Entire data set	All speakers	—	—
	Female speakers	—	58.8 ^b
	Male speakers	70.6 ^a	—

^aRepresents the categorization of all male speakers' fricatives based on a classification model derived from all female speakers' fricatives. ^bRepresents the categorization of all female speakers' fricatives based on a classification model derived from all male speakers' fricatives.

adult speakers to categorize these two fricatives was also found to contain sex-related differences in young children. Shown in Figure 7 are the percentage correct classifications of /s/ in male and female speakers by both adult models broken down by age group. As can be seen, the adult male model best categorizes the male fricatives in all age groups. A similar pattern is present for the adult female model and the female speakers, except for the two youngest age groups.

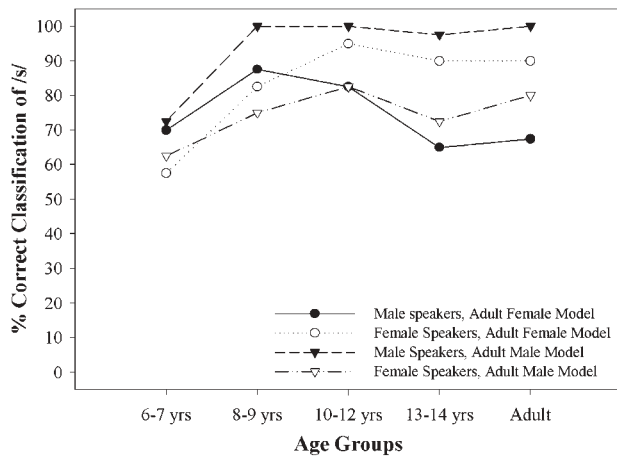
Shown in Figure 8 are the same data for categorization of /ʃ/ where a different pattern from the categorization of /s/ emerges. In particular, the adult male model does very poorly at categorizing /ʃ/ for female speakers at the two youngest ages and does consistently worse than the female model for all age groups. The adult male model also does relatively poorly in categorizing /ʃ/ for the male speakers at the youngest age. On the other hand, the female model does a reasonably good job of categorizing the fricatives of both sexes in all age groups, with steady improvement in the categorization of the female speakers. These data can be interpreted as providing additional support for the claim that acoustic variations as a function of speaker sex have not been completed in the earliest age group.

Discussion

Of primary interest in this study was the examination of systematic variation in a set of acoustic parameters as a function of fricative place of articulation, speaker age, and speaker sex. Analysis of both amplitude and spectral parameters for fricatives revealed significant variation due to place of articulation, sex, and age on selected acoustic parameters. What are the answers to the three questions raised at the beginning of this paper?

In terms of the duration and amplitude measurements of the fricative, we found little evidence of differences between speakers related to sex of the speaker. The only significant main effect related to sex of the speaker was the discovery that female speakers, overall, had significantly shorter nonsibilant fricatives than did male speakers; but this difference was relatively small (on the order of 10 ms). There was no significant Sex of the Speaker \times Age effects. However, this is not surprising as differential increases in vocal tract length and vocal fold size for females and males (expected in normal development through puberty) might not be expected to affect directly either fricative amplitudes or durations. The primary differences due to

Figure 7. Discriminant classification of /s/ for all age groups using adult discriminant functions.

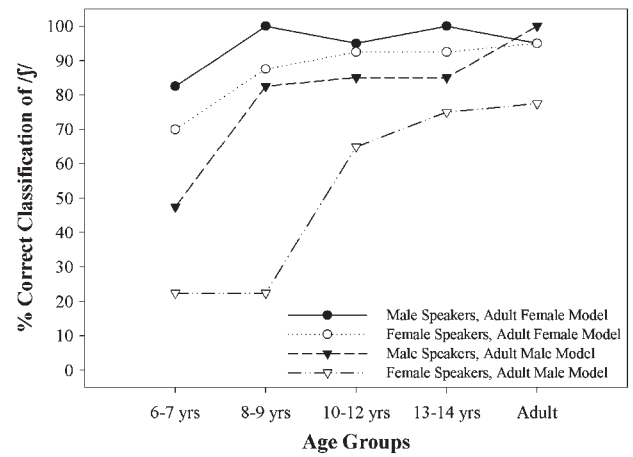


speaker age were that younger children generally produced longer fricatives and vowels than did older children and adults.

In general, the pattern of spectral differences among the fricatives across all age groups used here is similar to that found in the Jongman et al. (2000) and Nittrouer studies (1995)—in particular, variation in the values of the spectral peak and the first, third, and fourth spectral moments reflected place distinctions. However, unlike Jongman et al. we did not find that the spectral mean values, for any age group, separated /s/ from /f θ/. Like Jongman et al., we found that spectral peak values did separate /s/ from /f θ/ in male speakers but not in female speakers. A parameter not measured by Jongman et al. or Nittrouer—spectral slope—was found to separate /s/ from /f θ/ (and from /ʃ/ in all age groups). In terms of age effects, the spectral differences obtained match those expected from the normal development of the vocal tract (which increases in length through puberty).

According to these data, sex-specific differences in fricative articulation were found in the speech of prepubescent children. The results that are most notable involve the interactions between age, sex of the speaker, and place of articulation. Although the sex differences found in Jongman et al. (2000) and not examined by Nittrouer (1995) are evident even in the younger children (ages 6–7 years), there is evidence that some acoustic differences between females and males are reduced or absent in the youngest children. For example, the spectral slopes in the youngest age group did not differ. In all other groups, the spectral slope of the female speakers was higher than that of the male speakers. Interestingly, preliminary analysis with even younger speakers (Nissen, 2003) indicates that similar sex differences are not evident in 3–4-year-old children, but may start to emerge at approximately 5 years of age.

Figure 8. Discriminant classification of /ʃ/ for all age groups using adult discriminant functions.



Sex-specific differences also varied as a function of the fricative place of articulation, suggesting that the course of normal development for at least one of the fricatives may be different for females and males. In particular, we found that the spectral means declined for /ʃ/ as age increased, regardless of speaker sex. However, across all age groups a similar age-related decline for /s/ was only exhibited by the male speakers.

Thus far we have discussed age and sex related differences in terms of individual acoustic variables. In addition, findings from DA showed developmental differences in fricative production when considering a combination of acoustic variables. As expected, discriminant functions based on adult tokens (using only the oldest group) were less accurate in classifying the segments produced by the younger speakers. Results demonstrated that a discriminant function based on the adult male tokens was generally better at classifying fricatives produced by male speakers than female speakers—regardless of age. This indicated that sex-related differences (presumably a function of sex-linked vocal tract variation) were present even in the youngest speaker group. However, the fact that the classification accuracy of the female model showed a steady improvement with the increased age of the female speakers may provide support for the claim that sex-related developmental differences may just be emerging in the youngest age group. It is also interesting to note that sex-related differences in classification varied as a function of the place of fricative articulation.

Considering recent magnetic resonance imaging data suggesting that sexual vocal tract dimorphism for most children occurs at the peripubertal and postpubertal stages of development, it is likely that the sex-related acoustic differences found in this study cannot be entirely explained by anatomical differences alone. It may be concluded that a portion of these sex-linked differences may be associated

with learned or behavioral factors. These data support the contention that articulatory development follows different patterns in girls as opposed to boys, differences which may start to emerge as early as 6–7 years of age.

Acknowledgments

This research was supported by research support funds from the Advanced Telecommunication Research (ATR) Laboratory in Nara, Japan, by the College of Social and Behavioral Sciences at The Ohio State University, and by an Institutional National Research Service Award from the National Institutes of Health.

References

- Baer, T., Gore, J. C., Gracco, L. C., & Nye, P. W.** (1991). Analysis of vocal tract shape and dimensions using magnetic resonance imaging: Vowels. *Journal of the Acoustical Society of America*, *90*, 799–828.
- Busby, P. A., & Plant, G. L.** (1995). Formant frequency values of vowels produced by preadolescent boys and girls. *Journal of the Acoustical Society of America*, *97*, 2603–2606.
- Byrd, D.** (1994). Relations of sex and dialect to reduction. *Speech Communication*, *15*, 39–54.
- Dang, J., Honda, K., & Suzuki, H.** (1994). Morphological and acoustical analysis of the nasal and paranasal cavities. *Journal of the Acoustical Society of America*, *96*, 2088–2100.
- Fant, G.** (1960). *The acoustic theory of speech production*. The Hague, the Netherlands: Mouton.
- Fant, G.** (1966). A note on vocal tract size factors and nonuniform *F*-pattern scalings. *Speech Transactions Laboratory Quarterly Progress and Status Report*, *4*, 22–30.
- Fitch, W. T., & Giedd, J.** (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *Journal of the Acoustical Society of America*, *106*, 1511–1522.
- Forrest, K., Weismer, G., Milenkovic, P., & Dougall, R.** (1988). Statistical analysis of word-initial voiceless obstruents: Preliminary data. *Journal of the Acoustical Society of America*, *84*, 115–124.
- Jongman, A., Wayland, R., & Wong, S.** (2000). Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America*, *108*, 1252–1263.
- Klecka, W. R.** (1980). *Discriminant analysis*. Beverly Hills, CA: Sage.
- 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.
- Mattingly, I.** (1966). Speaker variation and vocal tract size. *Journal of the Acoustical Society of America*, *39*, 1219(A).
- Moore, C. A.** (1992). The correspondence of vocal tract resonance with volumes obtained from magnetic resonance images. *Journal of the Acoustical Society of America*, *35*, 1009–1023.
- Nissen, S.** (2003). *An acoustic analysis of voiceless obstruents produced by adults and typically developing children*. Unpublished doctoral dissertation, Department of Speech and Hearing Science, The Ohio State University.
- 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.
- Perry, T. L., Ohde, R. N., & Ashmead, D. H.** (2001). The acoustic bases for gender identification from children's voices. *Journal of the Acoustical Society of America*, *109*, 2988–2998.
- Peterson, G. E., & Barney, H. L.** (1952). Control methods used in the study of vowels. *Journal of the Acoustical Society of America*, *24*, 175–184.
- Sachs, J., Lieberman, P., & Erickson, D.** (1973). Anatomical and cultural determinants of male and female speech. In R. W. Shuy & R. W. Fasold (Eds.), *Language attitudes* (pp. 74–83). Washington, DC: Georgetown University Press.
- Schwartz, M. F.** (1968). Identification of speaker sex from isolated, voiceless fricatives. *Journal of the Acoustical Society of America*, *43*, 1178–1179.
- Story, B. H., Titze, I. R., & Hoffman, E. A.** (1996). Vocal tract area functions from magnetic resonance imaging. *Journal of the Acoustical Society of America*, *100*, 537–554.
- Swartz, B. L.** (1992). Gender differences in voice onset time. *Perceptual Motor Skills*, *75*, 983–992.
- Syntrillium Corporation.** (2000). CoolEdit (Version 1.1) [Computer software]. Phoenix, AZ: Author.
- Thomas, D.** (1992). Interpreting discriminant functions: A data analytic approach. *Multivariate Behavioral Research*, *27*, 335–362.
- Trautmüller, H.** (1988). Paralinguistic variation and invariance in the characteristic frequencies of vowels. *Phonetica*, *45*, 1–29.
- Trautmüller, H.** (1990). Analytical expressions for the Tonotopic Sensory Scale. *Journal of the Acoustical Society of America*, *88*, 97–100.
- Whiteside, S. P.** (1996). Temporal-based acoustic-phonetic patterns in read speech: Some evidence for speaker sex differences. *Journal of the International Phonetic Association*, *26*, 23–40.
- Whiteside, S. P.** (2001). Speaker's sex-specific fundamental and formant frequency patterns in a cross-sectional study. *Journal of the Acoustical Society of America*, *110*, 464–478.
- Whiteside, S. P., & Marshall, J.** (2001). Development trends in voice onset time: Some evidence for sex differences. *Phonetica*, *58*, 196–210.

Received November 20, 2003

Revision received November 1, 2004

Accepted January 14, 2005

DOI: 10.1044/1092-4388(2005/052)

Contact author: Robert Allen Fox, Speech Perception and Acoustic Laboratory, Department of Speech and Hearing Science, 1070 Carmack Road, The Ohio State University, Columbus, OH, 43210-1002. E-mail: fox.2@osu.edu