SPECTRAL FEATURES OF VOICELESS FRICATIVES PRODUCED BY AUSTRALIAN ENGLISH-SPEAKING CHILDREN

Casey Ford & Marija Tabain

La Trobe University
ceford@students.latrobe.edu.au, m.tabain@latrobe.edu.au

ABSTRACT

This paper examines some spectral features of voiceless fricatives /s, ʃ, f, θ/ produced by Australian English-speaking children (5-13 years). It finds that sex differences are evident in fricative production, despite the unlikelihood of sex dimorphism in the vocal tract. These differences are especially evident in the sibilant fricatives /s/ and /ʃ/. Girls produce sibilants with higher spectral mean and lower spectral skewness than boys. Boys produce /f/ with a higher spectral mean and lower skewness than girls, while spectral mean and skewness for /θ/ are very similar between sexes. Spectral mean of /s/ and /ʃ/ declines significantly with age, while /f/ and /θ/ show no change. This work builds upon our current knowledge of sociophonetic variation in Australian English, as well as our knowledge of children's acquisition and use of socially-structured variation

Keywords: Sociophonetics, Australian English, fricatives, phonetic variation, children's speech.

1. INTRODUCTION

Sociophonetic work on voiceless fricatives, particularly voiceless sibilants, has shown that certain spectral features correlate with particular social factors. Speaker sex or gender is a particularly robust factor examined in existing literature. In part, the acoustic differences found between male and female sibilant realisation stem from the physiological differences that exist in the vocal tract. It is expected that females are more likely to have smaller vocal tracts than males, and is therefore assumed that resulting sibilant production will involve a smaller front cavity for female speakers and higher corresponding spectral frequencies than males [20]. These acoustic differences are least often found in non-sibilant fricatives such as labiodental /f/ and interdental /θ/ where cavity dimensions have less influence compared to sibilants. Indeed, listeners can identify the sex of a speaker from isolated productions of /s/ and /ʃ/, but not from productions of /f/ or /θ/ [19]. However, evidence suggests that speakers adjust sibilant production within their physiological limits to align with particular social groups, such as gender, sexuality, social class and geographic region [15, 17, 21, 23], in turn affecting the physiology-based acoustic features of the sibilant. Sociophonetic studies examining spectral characteristics of fricatives other than /s/ appear to be lacking, potentially due to its status a robust socio-indexical cue.

While such sex-dimorphism is not evident in children's vocal tracts prior to the onset of puberty [2, 22], sex-specific differences in the acoustic characteristics of children's voices have been found [1, 12, 16]. Examining children's speech highlights that sex-specific acoustic characteristics are not solely physiologically determined, but largely a result of socially-influenced and learned articulatory behaviours. Regarding children's fricatives, consistent evidence of sex-specific spectral characteristics in sibilants has been shown, correlating with those expected for their gender [3, 6]. Girls are reported to produce /s/ and /ʃ/ with higher spectral frequencies than boys of the same age, with differences increasing with an increase in age due to the onset of structural changes in the vocal tract for male speakers at puberty. Comparatively, non-sibilant fricatives are rarely reported on. Fox and Nissen [6] report spectral moment values of non-sibilant fricatives /θ/ and /f/ for American English-speaking children aged between 6-14 years. For these fricatives, they found that spectral means (M1) were similar between sexes, and spectral skewness (M3) to be significantly lower for girls than for boys. Individually, they found that girls have higher M3 than boys for /f/ and lower M3 for /θ/.

This paper examines spectral features of voiceless fricatives /s, ʃ, f, θ/ produced by Australian English-speaking (AusEng) children, looking particularly at sex- and age-related characteristics. This expands on previous work that focused on AusEng children's sibilants [4]. Sociophonetic variation of fricatives by AusEng-speaking adults is yet to be examined, thus direct comparison is not currently possible. Results will be compared with results reported in similar work in other English varieties [6].

2. METHODS

2.1. Speech community, participants, and recording

Data were collected by the first author as part of a larger PhD project. The children attended a public

Age group	Sex	No. speakers	Age range (y;m)	Mean age (y;m)	/s/ tokens	/ʃ/ tokens	/f/ tokens	/θ/ tokens
Prep	Girls	6	5;7-6;2	5;11	441	71	132	24
	Boys	6	5;4-6;4	5;9	646	96	201	11
Year Three	Girls	6	8;3-9;3	8;8	852	94	293	69
	Boys	6	8;2-9;6	8;10	1127	99	370	75
Year Six	Girls	6	11;8-13;0	12;4	1351	169	426	138
	Boys	4	11;10-12;8	12;3	1070	127	324	155
				Totals	5488	656	1746	472

Table 1 Speaker group information.

primary school in the rural town of Yarrawonga, Victoria, 270km north-east of Melbourne, with a population of around 8000. Data were collected here as a starting point for examining sociophonetic variation between rural and urban speakers of AusEng. Speech recordings were taken of speakers in three primary school year levels, or age groups: Prep, Year Three, and Year Six. These groups were selected to provide an overview of the seven-year primary school period – a crucial time in the acquisition of sociolinguistic competence [11]. Information about each speaker group is reported in Table 1. Recordings were held in a quiet room on the school campus during school hours in sex- and age group-matched dyads. They participated in a range of child-friendly tasks designed to elicit a range of speech styles. These were a phase of spontaneous conversation, followed by a Map Task-style game, a reading task, and a picture identification task. All recordings were made using a Marantz Professional PMD661 solid-state recorder and two Shure SM94 microphones at a sampling rate of 44.1 kHz. Microphones were placed on a boom stand at around 20-30cm in front of each speaker's mouth. All speakers were in possession of their front teeth, and no speech delays were reported or observed.

2.2. Analysis

2.2.1. Data preparation

Recordings were segmented and force-aligned using the WebMAUS-multiple automated alignment service [9]. Segment boundaries were manually adjusted using the *Emu Speech Database System* [7]. Fricatives were segmented by placing the onset boundary at the onset of the frication noise, and then offset boundary at the cessation of the frication noise and the onset of the following segment. Tokens adjacent to another fricative sound, or contained instances of interference, such as interruption from another speaker or other background noises, were omitted. Across the speakers, 8452 tokens were extracted. 90 of these tokens were omitted, leaving 8362 tokens for analysis. Tokens were extracted using the *EmuR* interface [8]. Token numbers for each speaker group are shown in Table 1. Prosodic context, surrounding vowel contexts, and effect of speech task were not controlled for the current study, as the aim is to examine the spectral qualities of the fricatives in general. These elements will be examined in future work.

2.2.2. Spectra estimation and spectral moments

Fricative spectra were estimated using Fast Fourier Transform (FFT) with a 20ms Hamming window over the fricative midpoint. Spectral moments [5] were extracted at the temporal midpoint of each fricative within a spectral range of 1 – 15 kHz in order to capture the majority of the energy distribution while filtering out any potential background noise or adjacent voicing. The midpoint was measured to avoid the influence of co-articulatory gestures at the fricative onset and offset and any surrounding vowel formant transitions, making it the ideal place to examine the spectral qualities of the fricative itself.

The measurement of spectral moments is a common and robust approach in examining sociophonetic influence on fricative production. The measurement of spectral moments directly examines the spread of energy created in the production of a fricative. This paper focuses on the first (M1) and third (M3) spectral moments. The first spectral moment shows the mid-point frequency where all energy on either side of the point is even, while the third spectral moment shows the distribution of the energy, either being positively or negatively skewed. The relative fronting or retraction of the point of constriction in the production of /s/, for example, will affect these values, where a more fronted articulation will give rise to higher M1 values, and more negative M3 values. These spectral moments have been shown to be particularly sensitive to speaker sex [6, 8].

2.2.3. Statistical analysis

Linear mixed effects models were built to test the statistical effects of the social and linguistic factors on the spectral properties of the fricatives. Spectral parameters each served as dependent factors. Main effects of sex, age group, and fricative, and interactions of sex × age group × fricative were set as independent fixed factors. Speaker was included as a random intercept. Fricative place of articulation was set as a random slope. This was carried out using the *lmerTest* [10] package in R, with degrees of freedom estimated using the Satterthwaite approximation method. Post-hoc Tukey pairwise comparisons were carried out using *lsmeans* [13]. Significance was set at 0.05 and trend at 0.10

3. RESULTS

3.1. First Spectral Moment (M1)

Table 2 shows the significant main effects and interactions on M1. Neither sex nor age were significant main effects, but each interacted significantly with place of articulation. Figure 1 shows the spectral mean of the four fricatives by speaker sex. It shows that the girls produce both of the sibilant fricatives, /s/ and /ʃ/, with higher spectral mean than boys. These results are consistent with those previously found for American English-speaking children [6]. For the non-sibilant fricatives, spectral mean of /f/ is lower for the girls compared to the boys, while there is little difference between sexes for /θ/. This result is contrary to those reported by Fox and Nissen, particularly for /f/, where they found M1 to be similar between sexes. Statistically, speaker sex interacted significantly with place of articulation, as shown in Table 2. Post-hoc analyses (Table 3) showed that /s/ and /ʃ/ were significantly higher for girls, while /f/ was significantly higher for boys. The spectral means for each fricative measured were significantly different to one another (all comparisons *p*<.0001).

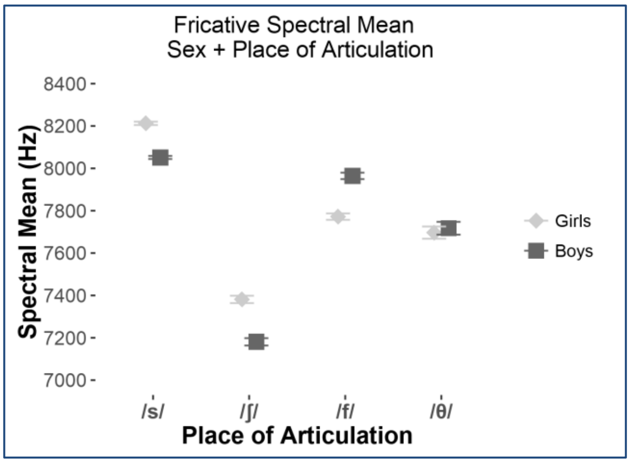

Figure 1 Spectral mean values (Hz) for each fricative by girls and boys overall.

Effects	Model output	*p*-value
place	F[3,89.8]=169.71	<.0001
sex:place	F[3,83.12]=12.72	<.0001
age group:place	F[6,87.7]=2.77	.02

Table 2 Statistically significant effects and interactions on M1.

Figure 2 shows M1 values for each fricative across the three age groups. M1 for the sibilants /s/ and /ʃ/ decreases with an increase in age. This decrease was statistically significant for both sibilants, as shown in Table 3. /f/ does not appear to be sensitive to age, while older speakers had higher M1 for /θ/ than younger speakers, although this was not statistically significant. These patterns are consistent with those reported by Fox and Nissen [6].

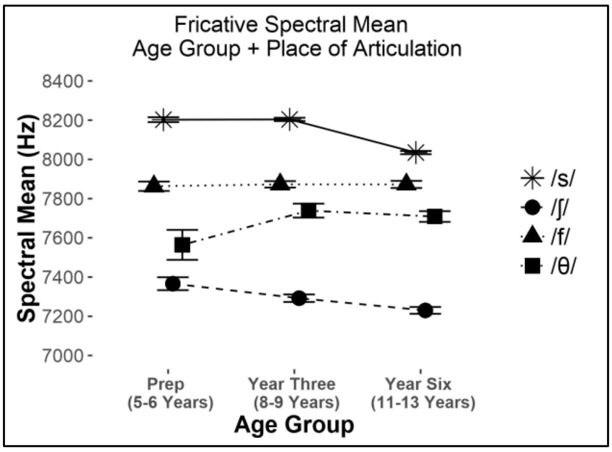

Figure 2 Spectral mean values (Hz) for each fricative in each age group.

Comparison	Difference (Hz)	Std. Error	Df	t-ratio	*p*
Girls – Boys /s/	187.0	61.9	67.3	3.02	.004
Girls – Boys /ʃ/	201.5	68.8	101.0	2.93	.004
Girls – Boys /f/	-185.4	64.3	78.1	-2.88	.005
Prep – Year Six /s/	197.1	77.1	67.6	2.56	.01
Year Three – Year Six /s/	193.1	76.6	65.8	2.52	.01
Prep – Year Six /ʃ/	198.8	85.3	99.2	2.33	.02

Table 3 Statistically significant pairwise comparisons for M1.

3.2. Third Spectral Moment (M3)

Table 4 shows the statistically significant main effects and interactions on fricative M3. The only significant main effect in the model was place of articulation, and pairwise comparisons showed all fricatives were significantly different from one another in M3 (*p*<.0001). Sex and place of articulation emerged as a highly significant interaction on M3, but the age group and place of articulation interaction did not quite reach significance as it did for M1.

Effects	Model output	*p*-value
place	F[3,91.5]=176.69	<.0001
sex:place	F[3,91.5]=12.29	<.0001
age group:place	F[6,89.3]=2.15	.06

Table 4 Statistically significant effects and interactions (and trends) on M3.

Figure 3 shows the M3 values by speaker sex at each place of articulation. Both sibilants /s/ and /ʃ/ are produced by girls with lower M3 values compared to boys. Pairwise comparisons (Table 5) showed the sex difference for /s/ to be highly statistically significant, but just under significance for /ʃ/. For the non-sibilant fricatives, /f/ was significantly higher in M3 for girls compared to boys, while there was no difference in /θ/ between the sexes. Again, these results are consistent with those reported by Fox and Nissen for M3 [6].

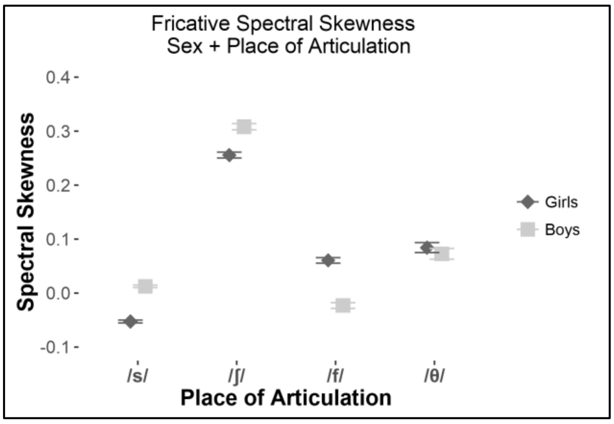

Figure 3 Spectral skewness values for each fricative by girls and boys overall.

Comparison	Difference	Std. Error	Df	t-ratio	*p*
Girls – Boys /s/	-0.1	0.02	82.6	-3.04	.003
Girls – Boys /ʃ/	0.04	0.02	120.9	-1.92	.06
Girls – Boys /f/	0.1	0.02	94.9	3.64	<.0001

Table 5 Statistically significant pairwise comparisons (and trends) for M3.

4. DISCUSSION

The main interest of this paper was to examine the spectral features of voiceless fricatives /s, ʃ, f, θ/ produced by AusEng-speaking children. Spectral differences in these fricatives as a function of speaker sex and age were of particular interest. Further, non-sibilant fricatives are not commonly included in work examining spectral characteristics, and thus the paper aimed to be a starting point of reference for these fricatives produced by AusEng-speaking children.

From the spectral moments analysis of the four voiceless fricatives, several points emerge. First, sex-specific differences in the fricatives produced by these AusEng-speaking children are evident. Sex-specific differences in the first and third spectral moments of the sibilant fricatives /s/ and /ʃ/ are consistent with what has been found in previous work on both adult and child production [4, 6, 8, 14]. The girls produce sibilant fricatives with higher spectral mean and lower spectral skewness in comparison to boys. This suggests that these children produce /s/ and /ʃ/ with articulatory behaviours that are learned from their social environment. As reported by Fox and Nissen [6] no sex differences were evident for the non-sibilant fricative /θ/. However, the labiodental fricative /f/ showed evidence of sex-specific spectral characteristics, where spectral mean was higher and spectral skewness was lower for boys compared to girls. This result is contrary to the results reported by Fox and Nissen [6], and is also surprising given that /f/ is less likely than /s/ or /ʃ/ to be a socio-indexical cue. However, the children may indeed be producing a gender-specific feature evident in the speech of adult speakers of AusEng in their community, but this cannot be explored further without sociophonetic investigations of fricatives produced by AusEng-speaking adults.

Age-related effects on fricative spectral characteristics were also observed here, most notably in M1. These were particularly evident for the sibilant fricatives. A significant decrease in the spectral mean of /s/ and /ʃ/ was found between Prep and Year Six. There was some evidence of an increase in spectral mean of /θ/ with an increase in age for the speakers overall, however this increase was not statistically significant. There appeared to be no change in /f/ spectral mean with an increase in age.

Overall, the results found here indicate that the sibilant fricatives /s/ and /ʃ/ appear to be more sensitive to sex and age effects in comparison to the non-sibilant fricatives /f/ and /θ/. This result is consistent with previous work on sociophonetic variation of fricatives in the speech of both adults and children in several varieties of English. Sibilants, particularly /s/, are robust sociophonetic cues to a range of social factors. Some evidence of sex-specific production of /f/ was found, however explanations of this result and its importance in the social context of AusEng cannot be explored without additional sociophonetic investigations. More work also needs to be carried out on the acquisition and use of sociophonetic variation by speakers of AusEng from a variety of social backgrounds in order to further investigate these patterns.

5. REFERENCES

[1] Busby, P.A., Plant, G.L. 1995. Formant frequency values of vowels produced by preadolescent boys and girls. *J. Acoust. Soc. Am.* 97, 2603–2606.

[2] Fitch, W.T., Giedd, J. 1999. Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *J. Acoust. Soc. Am*. 106, 1511–1522.

[3] Flipsen, P., Shriberg, L., Weismer, G., Karlsson, H., McSweeny, J. 1999. Acoustic Characteristics of /s/ in Adolescents. *J. Speech, Lang. Hear. Res*. 42, 663–677.

[4] Ford, C., Tabain M. 2018. Gender differences in spectral characteristics of voiceless sibilants produced by Australian English-speaking children. *Proc. 17th Australasian International Conference on Speech Science and Technology.* Coogee, 97–100.

[5] Forrest, K., Weismer, G., Milenkovic, P., Dougall, R.N. 1988. Statistical analysis of word-initial voiceless obstruents: preliminary data. *J. Acoust. Soc. Am.* 84, 115–123.

[6] Fox, R.A., Nissen, S.L. 2005. Sex-related acoustic changes in voiceless English fricatives. *J. Speech, Lang. Hear. Res.* 48, 753–765.

[7] Harrington, J. 2010. *Phonetic Analysis of Speech Corpora*. Malden: Blackwell.

[8] Jongman, A., Wayland, R., Wong, S. 2000. Acoustic Characteristics of English Fricatives. *J. Acoust. Soc. Am.* 108, 1252–1263.

[9] Kisler, T., Schiel, F., Sloetjes, H. 2012. Signal Processing via web services: the use case WebMAUS. *Proceedings of Digital Humanities 2012*. 30–34.

[10] Kuznetsova, A., Bruun Brockhoff, P., Haubo Bojesen Christensen, R., Brockhoff, P.B., Christensen, R.H.B. 2017. lmerTest Package: Tests in Linear Mixed Effects Models, *J. Stat. Softw*. 82, 1–26.

[11] Labov, W. 2007. Transmission and Diffusion. *Language.* 82, 344–387.

[12] Lee, S., Potamianos, A., Narayanan, S. 1999. Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *J. Acoust. Soc. Am*., 105, 1455– 468.

[13] Lenth, R. 2016. Least-Squares Means: The R Package lsmeans. *J. Stat. Softw*. 69, 1–33.

[14] Li, F., Rendall, D., Vasey, P.L., Kinsman, M., Ward-Sutherland, A., Diano, G. 2016. The development of sex/gender-specific /s/ and its relationship to gender identity in children and adolescents. *J. Phon*. 57, 59–70.

[15] Munson, B., McDonald, C., Deboe, N.L., White, A.R. 2006. The acoustic and perceptual bases of judgements of women and men's sexual orientation from read speech. *J. Phon*. 34, 202–240.

[16] Nissen, S.L., Fox, R.A. 2005. Acoustic and spectral characteristics of young children's fricative productions: a developmental perspective. *J. Acoust. Soc. Am.* 118, 2570–2578.

[17] Podesva, R.J., Van Hofwegen, J. 2015. /s/exuality in small-town California: Gender normativity and the acoustic realization of /s/. In: Levon, E., Mendes, R.B. (eds), *Language, Sexuality, and Power: Studies in Intersectional Sociolinguistics.* New York: Oxford University Press, 168–188.

[18] R Core Team. 2014. A language and environment for statistical computing. R Foundation for statistical computing. Vienna, Austria.

[19] Schwartz, M.F. 1968. Identification of Speaker Sex from Isolated, Voiceless Fricatives. *J. Acoust. Soc. Am.* 43, 1178–1179.

[20] Stevens, K.N. 1998. *Acoustic Phonetics*. Cambridge: MIT Press.

[21] Stuart-Smith, J. 2007. Empirical evidence for gendered speech production. In: Cole, J., Haulde, J.-I. (eds), *Laboratory Phonology 9*. Berlin: Mouton, 65-86.

[22] Vorperian, H.K., Wang, S., Schimek, E.M., Durtschi, R.B., Kent, R.D., Gentry, L.R., Chung, M.K. 2011. Developmental Sexual Dimorphism of the Oral and Pharyngeal Portions of the Vocal Tract: An Imaging Study. *J. Speech, Lang. Hear. Res*. 54, 995–1010.

[23] Zimman, L. 2017. Variability in /s/ among transgender speakers: Evidence for a socially-grounded account of gender and sibilants. *Linguistics*. 55, 993–1019.