Acoustic Variability of Fricatives in Normal Adults

Christine H. Shadle[1], Wei-rong Chen[1], Laura L. Koenig[1,2] and Jonathan L. Preston[3]
[1]Haskins Laboratories, New Haven, CT; [2]Adelphi University, Garden City, NY;
[3]Syracuse University, Syracuse, NY

In order to characterize and compare fricatives, different sets of acoustic parameters have been used. These have demonstrated, for example, the acoustic correlates of fricative place and voicing within and across languages; the effects of phonetic context and speaker age and gender; and differences between typical and misarticulated fricatives. One set of parameters describes the shape of the spectrum, i.e. spectral moments. Another set was devised in studies employing mechanical models, for which the articulatory and aerodynamic variables were known precisely (Shadle 1985). These parameters include acoustic characteristics that are strongly related to articulatory and/or aerodynamic aspects of production, such as peak and trough frequencies, spectral tilt and dynamic amplitude. When such parameters can be shown to correspond to articulatory and aerodynamic conditions in human speech production, they can be used to estimate the articulatory and aerodynamic production variables in situations where such data are unknown or only partially measurable.

The parameters have been adapted over the years to deal with constraints of recording conditions (e.g. sampling rate, acoustic calibration), corpus (e.g. all fricatives, or /s/ only), and speaker populations (e.g. can nonsense words be used; how many tokens are possible?). Some studies included both acoustic and articulatory data, allowing the parameters to be validated. Acoustic-only studies have manipulated phonetic variables such as phonetic context, effort, and lexical stress. These can be predicted to affect, e.g., lip rounding, constriction area, and airflow, and therefore resonance frequencies and noise source parameters; the inferences can be tested using the acoustic parameters (e.g. Jesus & Shadle 2002; Koenig et al. 2013).

These studies have raised some questions that could not be answered given the constraints of each study. First, what is the variability of these parameters for normal speakers? Are the adapted parameters, such as level differences in bands up to 11 kHz, as useful in distinguishing fricatives from each other, or in predicting the underlying articulatory parameters, as the original parameters, such as spectral slopes up to 17 kHz? If accuracy is lost as a result of the adaptations, how much?

This study computes the parameters for acoustic recordings of seven normal adults producing all English fricatives. Key parameters discussed in this abstract are F_M, the frequency of the main peak, which is determined by the resonances of the front cavity; and AmpD, the amplitude difference between the main peak and the minimum at low frequencies, which is related to the type and strength of the noise source. Four women and three men with normal speech and hearing were recorded producing corpora used in a study of adolescents (Koenig et al. 2013). Three of the corpora used pictures to elicit words (originally designed for use with children). The fourth corpus was the Grandfather Passage (Reilly & Fisher 2012). Participants were recorded in an anechoic chamber. Multitaper analysis was performed at the midpoint of all fricatives at least 30 ms long; 31 tapers were used to compute a spectrum of every token. Following previous studies, frequency bands were defined for each fricative and participant gender to compute F_M and AmpD parameters for every token.

Following Koenig et al. (2013), results for /s/ tokens were split into three groups according to context: nonlabial, labial (rounded vowel, labial consonant and/or lip approximant, but excluding /str-/), and /str-/; F_M trended lower in that order, but there was sizeable overlap for most participants (see Fig. 1).

One participant, W1, produced /str-/ words such that they were perceived as [ʃtr-] (Stevens & Harrington 2016). The average spectrum of those /s/ tokens was similar to her /ʃ/ spectra. However, her F_M for /ʃ/ tokens were lower on average and less variable than her F_M for /s/ in /str-/. Her AmpD values for /ʃ/ were on average higher, although the difference was not as marked, and were also less variable than for /s/ in /str/. These results are similar to those of Stuart-Smith et al.(2019), though they used the first spectral moment as their measure.

Subject W3 overall had the lowest F_M values of /s/ of the women. Average F_M for her /s/ in /str-/ was almost as low as W1's, but without her /str-/ words being perceived as [ʃtr-]. Formal perceptual results will be compared to the acoustic parameters to further investigate these data regarding this ongoing /str-/ to /ʃtr-/ sound change.

For both men and women, F_M for /ʃ/ is less variable than for /s/ (see Fig. 3), though note that the phonetic contexts are less variable for /ʃ/. F_M in sibilants is lower on average for men than women, consistent with previous studies (Shadle et al. 2017); the gender difference does not hold for the nonsibilants /f/ and /θ/.

AmpD is designed to differentiate sibilants from nonsibilants, being larger for sibilants. This prediction holds for women and men, though for these data, the difference is larger for women. For speakers M1 and M2, AmpD distributions of /ʃ/ overlap with those for /f/ and /θ/.

Overall, F_M was higher on average for /f/ than /θ/, though the size of this difference varied by subject. Many studies (Harris 1958, Shadle, Mair & Carter 1996) have indicated that the transitions matter more in distinguishing /f/ from /θ/; these parameters based on mid-fricative spectra do not include transition information.

In general, these results indicate that the parameters differentiate the fricatives in the expected ways. The data also provide needed context with which to interpret results from clinical populations. Analyses are underway to assess change throughout the fricatives, and characteristics of voiced fricatives.

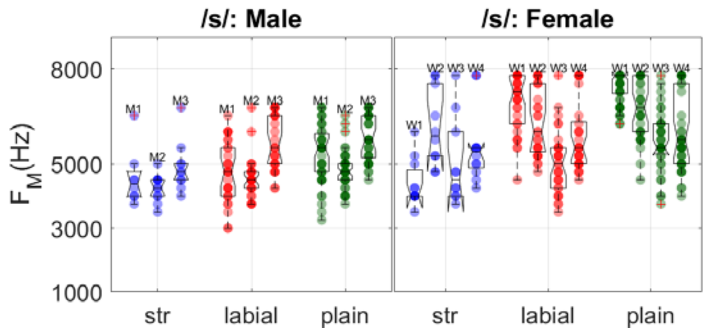

Figure 1. F_M for /s/, 1 circle per token. Horizontal bar is median; surrounding hourglass indicates central 50% of distribution.

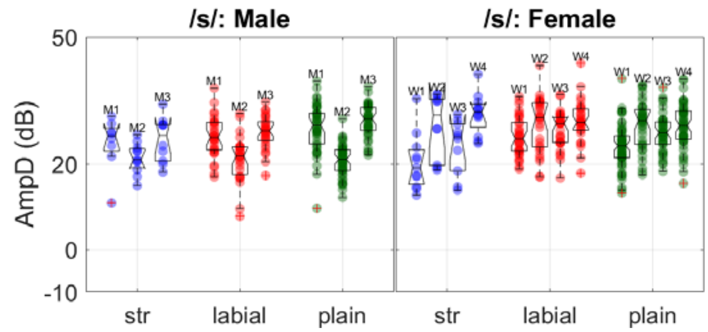

Figure 2. AmpD for /s/.

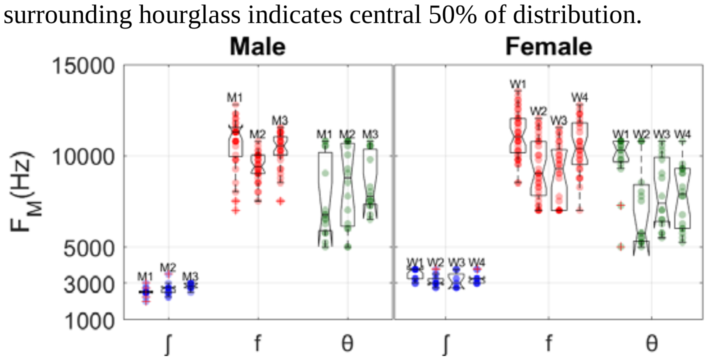

Figure 3. F_M for / ʃ , f , θ /.

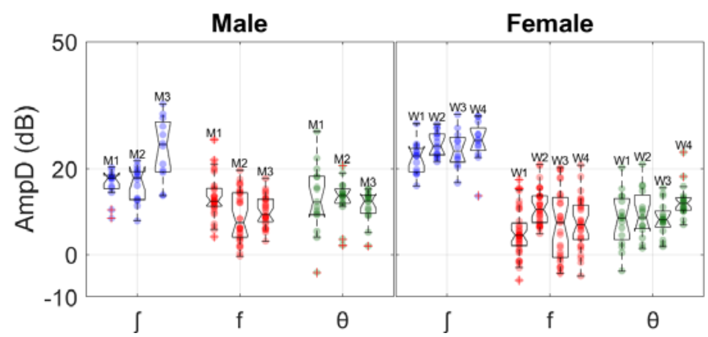

Figure 4. AmpD for / ʃ , f , θ /.

References

Harris, K.S. (1958) Cues for the discrimination of American English fricatives in spoken *syllables Lang and Speech* 1, 1-7.

Jesus, L.M.T. & Shadle, C.H. (2002) A parametric study of spectral characteristics of European Portuguese fricatives. *J. Phonetics* 30, 437-464.

Koenig, L.L., Shadle, C.H., Preston, J.L. & Mooshammer, C.R. (2013) Toward improved spectral measures of /s/: results for adolescents. *JSLHR* 56, 1175-1189.

Reilly, J., & Fisher, J. L. (2012). Sherlock Holmes and the strange case of the missing attribution: A historical note on "The Grandfather Passage". *JSLHR 55*, 84-88. doi: 10.1044/1092-4388(2011/11-0158).

Shadle, C.H. (1985) The acoustics of fricative consonants. Ph.D. thesis, Mass. Inst. Tech; RLE Rpt. 506.

Shadle, C.H., Chen, W-R., & Whalen, D.H. (2017) Articulatory-acoustic relationships for [s] in the XRMB database. Proc. ISSP, Tianjin, 16-19 October.

Shadle, C.H., Mair, S.J. & Carter, J.N. (1996) Acoustic characteristics of the front fricatives [f, v, θ, δ]. 1st ETRW – 4th Seminar on Speech Production, Autrans, 193-196.

Stevens, M., & Harrington, J. (2016). The phonetic origins of /s/-retraction: Acoustic and perceptual evidence from Australian English. *Journal of Phonetics*, *58*, 118-134. doi: https://doi.org/10.1016/j.wocn.2016.08.003.

Stuart-Smith, J., Sonderegger, M., Macdonald, R., Mielke, J., McAuliffe, M., & Thomas, E. (2019) Large-scale acoustic analysis of dialectal and social factors in English/s/-retraction. Paper 872 ICPhS, Melbourne, Australia.