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 [2]Adelphi University [3]Syracuse University

ISSP 2020
Dec 14-18, 2020

Introduction 1

- Production-related parameters have been developed in order to compare fricatives – mechanical models vs. human speech[7,9], normal vs. clinical population[1,4,10,12], effect of size and gender[4,9], effect of voicing[3] and phonetic context[4,9,11], etc.
- When articulatory/aerodynamic data exist, assumptions underlying parameters are tested[2,7,9,12], strengthening predictions made from acoustic-only data [3,10]
- Study of /s/ in adolescents (typical and with RSSEs) [4,10] led to followup: record adults with same methods. **How much of variability is due to age, how much to characteristics of corpus, recording, etc.?**

Definition Of Parameters -- Sibilants

4 frequency bands: for /s/, male, Low = 1-3, mid = 3-7, high = 7-11, xhigh = 11-15 kHz

Low: find minimum; Mid: find maximum

F_M = frequency of max; related to Length of front cavity (place, lip rounding)

AmpD = difference between max and min. Degree of sibilance

Slope = fit line from F_M to 14 kHz. Related to noise source spectrum

Level = sound power level within mid, high, xhigh bands

-- Non-Sibilants

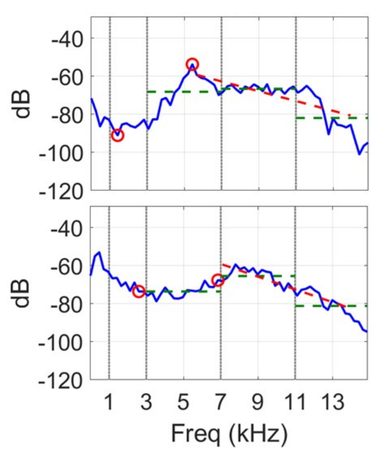
Use frequency bands as for /s/

Low-freq min, 1-7 kHz: FminA; Overall max, 1-13 kHz: FmaxA

AmpRange = difference between overall max and min

Max of broad peak, 5-12 kHz: F_h. Free jet noise spectrum

Slope, Levels as for /s/. Noise source – changes during fricative, & with stress



Method 2

CORPORA

Words were drawn from a study of children with residual speech sound errors (RSSEs) [6]

- Isolated words, elicited by picture naming: 65 words
- Rapid picture naming: name 6 words slowly, then rapidly in random order, 4-5 times each. Strawberries, spaghetti, elephant, thermometer, helicopter, umbrella
- Other tasks involving connected speech:
  - Recalling sentences: repeat 20 sentences that are successively longer
  - Grandfather Passage

PARTICIPANTS & PROCESSING

7 adults, ages 21 – 58; 3 men 4 women; Normal speech, hearing

Recorded in anechoic chamber

Shure WH30 headset Mic, Icicle Amplifier, recorded to Praat at 44 kHz; limiting frequency range is amplifier at 15 kHz

Speech is force-aligned [5] with manual correction

Analyzed each speaker separately

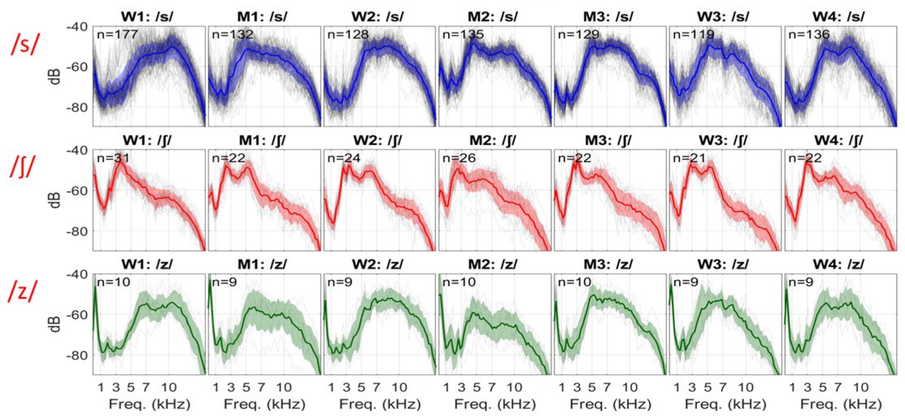
For some parameters, heuristic frequency bands differ for men and women

Multitaper spectra computed in 30 ms frames, 7 tapers, beginning, middle and end of each fricative [4]

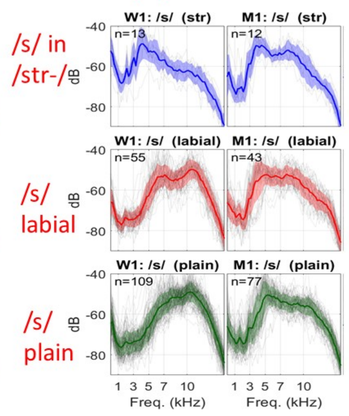
Parameters computed for each spectrum

Results -- sibilants 3

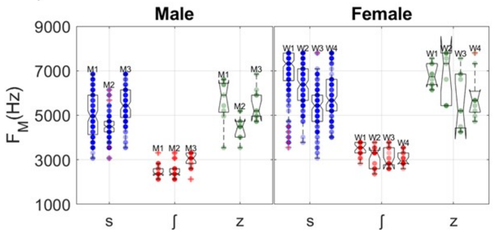
Mid-Fricative MT Spectra: Speakers



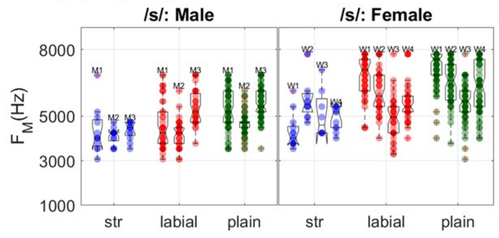
3 /s/ contexts



Graphs above: solid line = mean of all MT mid-fricative spectra; shaded area = +/- 1 std. dev. Note W1 /s/; she produced sound change in words such as "strap, strawberries". Upper right, /s/ separated into 3 phonetic contexts. Both W1 and M1 show difference in lowest peak for /s/ in /str-/.

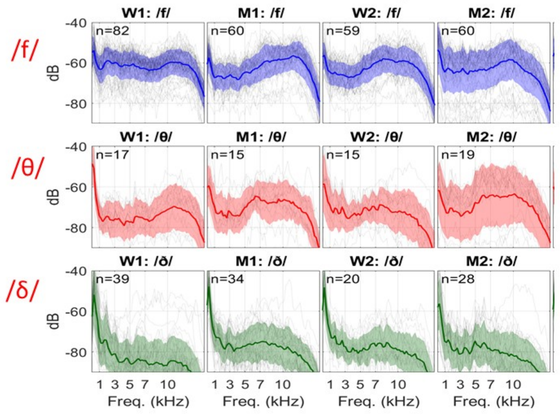


Left, note F_M – higher and more variable for /s/ than /ʃ/, for women than for men. Right, /s/ in /str-/ has lower F_M for W1 than other contexts.

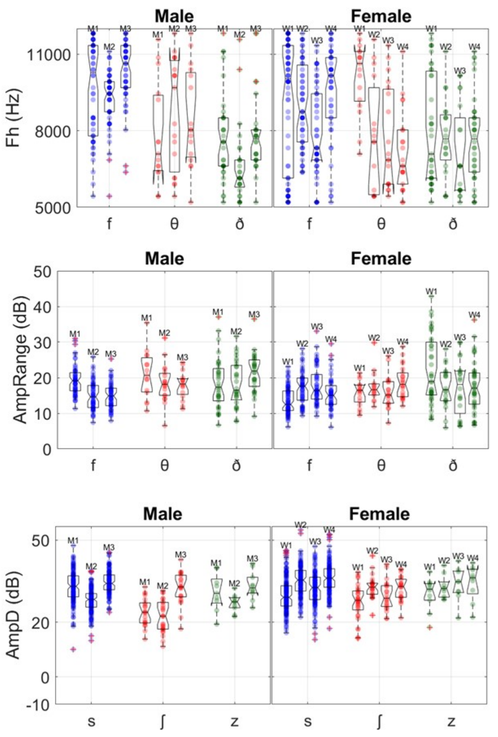


Results – non-sibilants 4

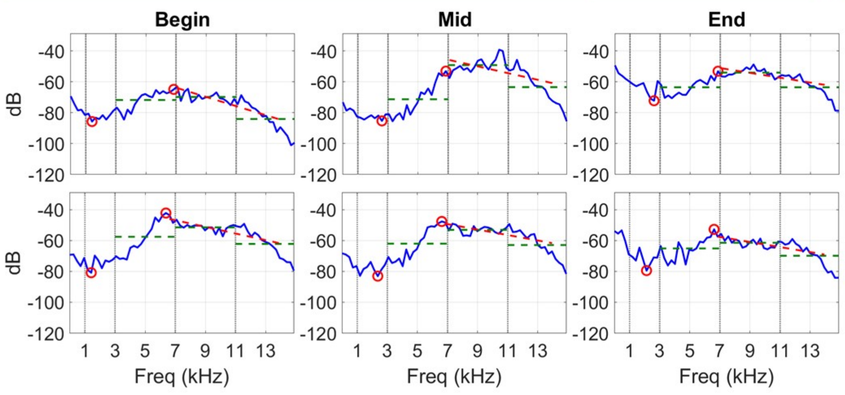
Mid-Fricative MT Spectra: Speakers



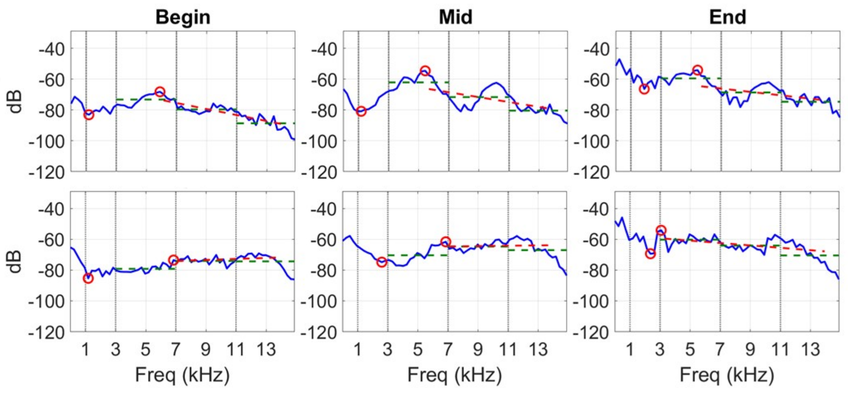
- F_h = max in range 5 -12 kHz
- Variable, but F_h (/f/) > F_h (/θ/) for most
- AmpRange = amp difference from 1 – 13 kHz
- AmpRange < AmpD as predicted; localized source for sibilants generates bigger difference



Results – speaking style 5



M1: /s/ in /s æ../
Top isolated word "sandwich"
Bottom connected speech "Saturday"
More high-frequency energy in isolated word
Slopes similar for all



M1: /θ/ in /θɝ../
Top isolated word "thirsty"
Bottom connected speech "thermometer"

AmpD > 20 dB isol. word – Sibilant-like for this context
Slopes similar for all

Conclusions 6

- Parameters F_M and AmpD work for sibilants even in these corpora; AmpD slightly smaller than past studies using sustained fricatives [3,11]
- Sound change demonstrated in two subjects with F_M parameter mid-/s/ in /str-/ words; F_M ([ʃtr-]) < F_M ([str-])
- For non-sibilants, AmpRange demonstrates relative flatness of spectra; AmpRange (f, θ) < AmpD (s, ʃ, z/), but in some contexts /θ/ appears more sibilant-like. F_h measures frequency of broad peak, not specific resonance. F_h (/f/) > F_h (/θ/) for most speakers
- Spectral slope without any amplitude intercept does not differentiate fricatives, or position within fricative
- Sound power spectrum levels in three bands shift as expected, though less in connected speech; need to devise ways of expressing that
- Production-related parameters offer way to characterize variability, for use in comparing adults to adolescents, normal to clinical populations, etc. Speaking style differences bear further scrutiny

References

1. Cox, S., McNicholl, K., Shadle, C.H., and Chen, W.-R. (2020) Variability of electrolarynx speech intelligibility in multi-talker babble. AJSLP 29, 2012-2022.
2. Iskarous1, K., Shadle, C.H. and Proctor, M.I. (2011) Articulatory-acoustic kinematics: the production of American English /s/. *J. Acoust. Soc. Am.* 129:2, 944-954.
3. Jesus, L.M.T. and Shadle, C.H. (2002) A parametric study of spectral characteristics of European Portuguese fricatives. *J. Phonetics* 30, 437-464.
4. Koenig, L., Shadle, C.H., Preston, J.L. and Mooshammer, C.R. (2013) Toward improved spectral measures of /s/: results from adolescents. *J.Speech Lang. Hearing Res.* 56, 1175-1189.
5. McAuliffe, M., Socolof, M., Mihuc, S., Wagner, M. and Sonderegger, M. (2017) Montreal Forced Aligner: trainable text-speech alignment using Kaldi. Proc. Interspeech, Stockholm, 498-502.
6. Preston, J. L., & Edwards, M. L. (2007). Phonological processing skills of adolescents with Residual Speech Sound Errors. *Language, Speech, and Hearing Services in Schools, 38*, 297–308.
7. Shadle, C.H. (1985) The Acoustics of Fricative consonants. Ph.D. thesis, MIT, Cambridge, MA; RLE Tech. Rpt. TR-506.
8. Shadle, C.H. (2012) Acoustics and Aerodynamics of fricatives, in *Handbook of Laboratory Phonology*, eds. A. Cohn, C. Fougeron and M. Huffman. Oxford Univ. Press, Oxford, 511-526.
9. Shadle, C.H., Chen, W.-R. and Whalen, D.H. (2017) Articulatory-acoustic relationships for [s] in the XRMB database. Proc. ISSP, Tianjin, 16-19 October.
10. Shadle, C.H., Koenig, L.L. and Preston, J.L. (2014) Acoustic characterization of /s/ spectra of adolescents: moving beyond moments. POMA vol. 12, 060006
11. Shadle, C.H. and Mair, S. (1996) Quantifying spectral characteristics of fricatives. Proc. ICSLP, Philadelphia, 1521-1524.
12. Shadle, C.H., Stone, M.L., and Chen, W.-R. (2018) The acoustic consequences for sibilants produced by glossectomies compared to healthy controls. ASA poster 5aSC23, Victoria, Canada

Acknowledgements

This work is partially supported by NIH grant DC-002717 to Haskins Laboratories.

NIH National Institute on Deafness and Other Communication Disorders (NIDCD)