Revisiting the Role of the Sublingual Cavity in the /s/-/ʃ/ Distinction

Haskins Laboratories
300 George St, New Haven, Connecticut, USA.

Christine H. Shadle, Michael I. Proctor, Maria Berezina, Khalil Iskarous
(shadle@haskins.yale.edu)

157th Meeting of the ASA
Portland OR
18-22 May 2009

2aSC10

Abstract

MRI and acoustic investigation into the role of the sublingual cavity (SC) in fricative production have revealed differing patterns. For one subject, [ʃ] is characterized by SC presence and lip protrusion, resulting in marked spectral differences from [s]. Another subject exhibits variability in SC shape and size, and tongue and lip position in both fricatives; acoustic spectra likewise vary. Mechanical model experiments show that noise generation changes as the sublingual cavity is filled in; the main resonance frequency does not necessarily change. It appears, therefore, that the sublingual cavity is just one of many articulatory variables available to the speaker, and may not be the principal cause of observed acoustic differences in all subjects.

Sibilant Production

Noise source: located at teeth for all sibilants (Shadle 1991)
Filter: excitation of front cavity resonances
Sublingual Cavity: may be important for filter differences between [s] and [ʃ]:

- Perkell et al. (1979): sudden increase in SC volume results in shift [s] → [ʃ], lowering frequency of main resonance abruptly
- Perkell et al. (2004): subjects who make (for [s]) or avoid (for [ʃ]) tongue-lower teeth contact have clearest separation of sibilants (judge by spectral moment M1 and % identification metrics)

but: articulation & acoustics differ amongst speakers (Shadle et al. 2008a):

- *constriction location* varies (so +/- room for SC in [s], [ʃ])
- [s] *tongue configuration* varies (laminal vs. apical)
- amount and kind of *vowel context effects* vary
- *acoustic spectra* of sibilants can vary widely with vowel context, resulting in apparent overlap in some subjects (eg. [us:u] & [iʃ:i] spectra very similar)
- SC differences *not always clean or categorical* between sibilants

⇒ *is SC volume really main contributor to* [s]-[ʃ] *spectral differences?*

Goals

- compare MRI and acoustic data for two subjects who differ the most
- study the effect of varying SC volume using mechanical models
- examine acoustic consequences of targeted articulatory changes

Method - MRI

Corpus:
- 2 sibilants: [s] – [ʃ]
- 3 vocalic contexts: [i_i] – [a_a] – [u_u]
- 5 American English subjects (3W, 2M)

Scan Sequence:
- Siemens Sonata 1.5T, 2D True-FISP: Tr=200ms, Te=3.3ms, Flip=70°
- 4mm slices, 4.8mm inter-slice spacing (resolution vs. acquisition time)
- three orientations: sagittal, axial, oblique

Procedure:
- each fricative imaged 9 times (3 orientations, 3 contexts)
- elicited with IPA stimuli (eg. [piʃ:i]), sustained 36sec: [pi ʃ ʃ ʃ ….. ʃ i]
- teeth imaged separately using hi-res 3D scans of dental impressions
- midsagittal dentition from soft-tissue boundaries, registered with landmarks
- production monitored real time: PhonOr noise-cancelling microphone

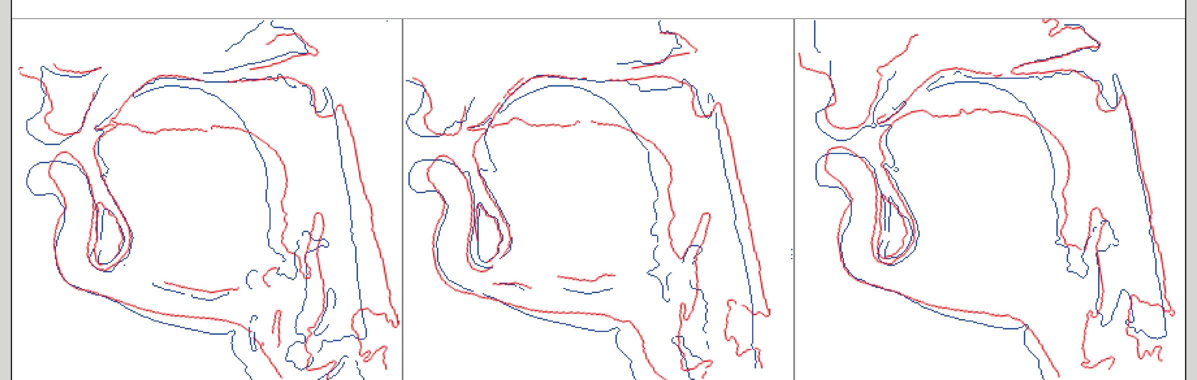
Method – Acoustic Analysis

Recording:
- corpora reproduced in anechoic chamber
- far-field recording @ 44kHz : B&K 4190 + Nexus, National Instruments A/D

Analysis:
- spectra generated from 5 second segments of sustained fricative signal
- PSD: 512-pt FFTs (100ms windows, 50% overlap, 3dB pre-emphasis)

Results – MRI

Midsaggital comparison of sibilant production:
- sublingual cavity generally present for [ʃ:], not for [s:]
- SC larger in low ([aʃ:a]) and back ([uʃ:u]) vowel contexts
- SC less prominent in high front vowel context ([iʃ:i])
- SC more pronounced for subject M2 than for M1



[is:i] – [iʃ:i] (M1) [as:a] – [aʃ:a] (M1) [us:u] – [uʃ:u] (M1)

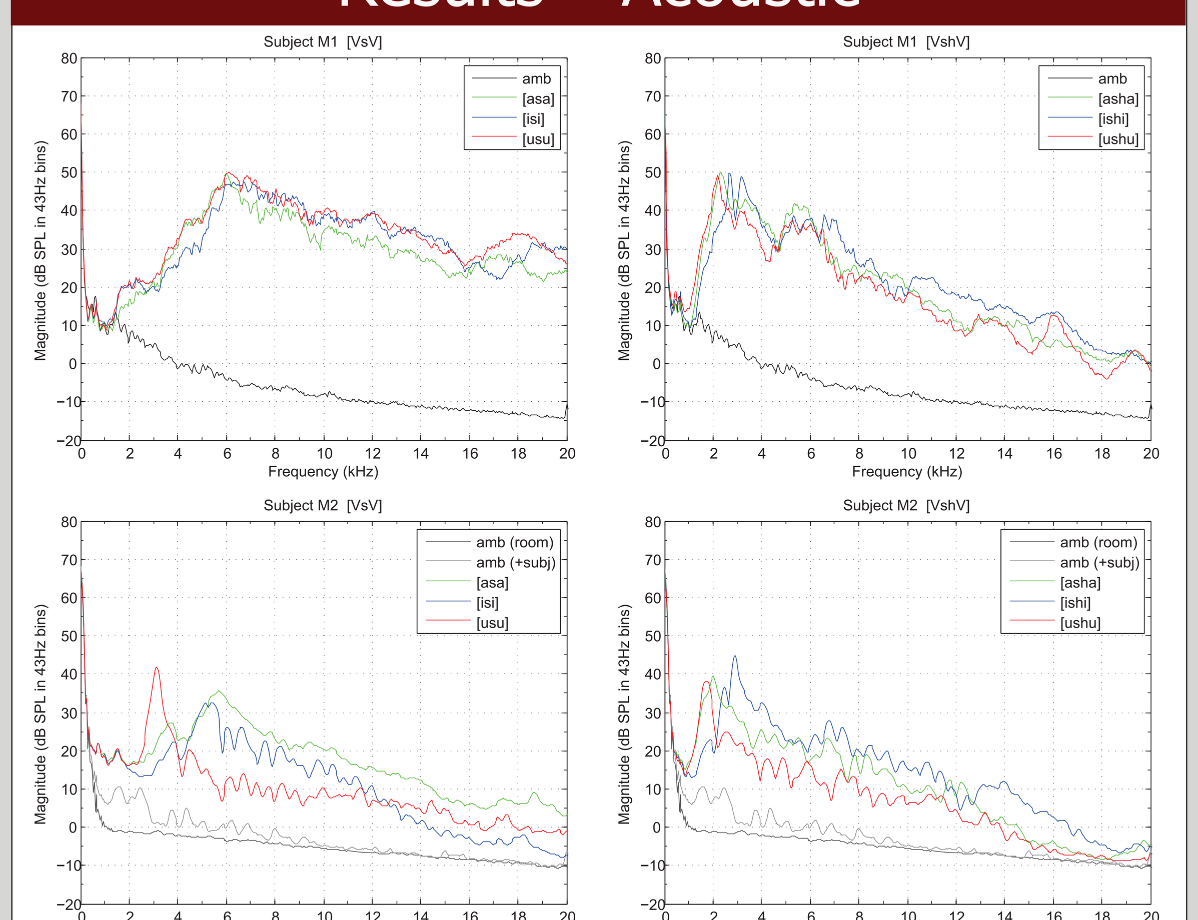
[is:i] – [iʃ:i] (M2) [as:a] – [aʃ:a] (M2) [us:u] – [uʃ:u] (M2)

Individual variation in sibilant production:

Subject M1: more lip protrusion for [ʃ:] regardless of vowel context
SC variation: largest in [aʃ:a]

Subject M2: more lip protrusion/rounding in [u] context for both [s:], [ʃ:]
tongue shape varies with vowel
SC variation: largest in [aʃ:a], smallest in [us:u]

Results – Acoustic



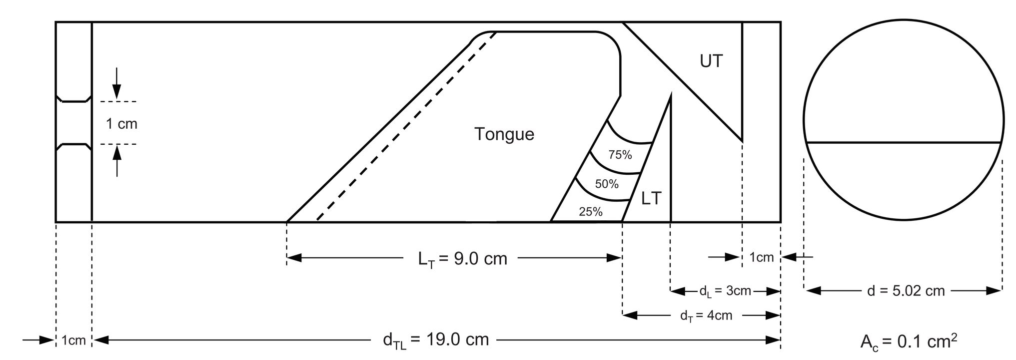
Subject M1: *little acoustic variation* with vowel context in [s:] and [ʃ:]
peak freqs lower (~300 Hz) in tokens with larger SC

Subject M2: *large variation* with vowel context:
[us:u] peak at 2-2.5kHz < [is:i], [as:a] peak freqs
[uʃ:u] peak at 1 kHz < [iʃ:i] peak

Conclusion: [us:u], [uʃ:u] spectral shape differ at all frequencies
⇒ *can't be explained by larger SC* in [us:u]

Method – Mechanical Modeling

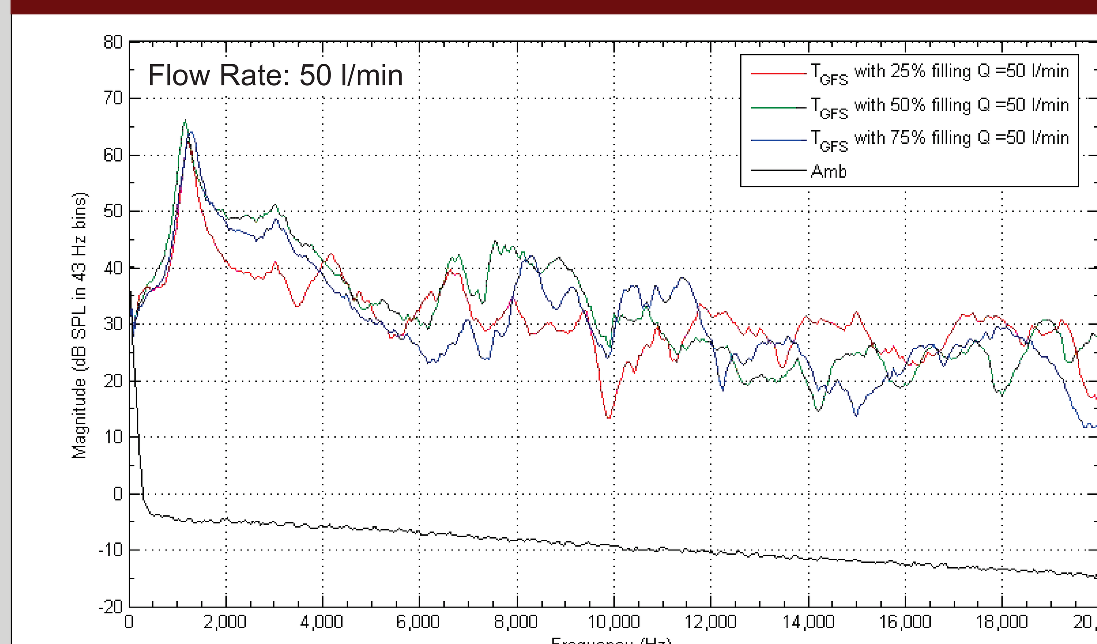
Model Construction:
- tract: cylindrical polycarbonate tube (∅ = 5.02 cm, L = 19.0 cm)
- articulators: machinable wax, sealed with silicon adhesive
- trachea: o-ring sealed PVC tubing (∅ = 1.8 cm, A = 2.55 cm^2)
- glottis: wax ring tapered orifice (∅ = 1.0 cm, A = 0.79 cm^2)



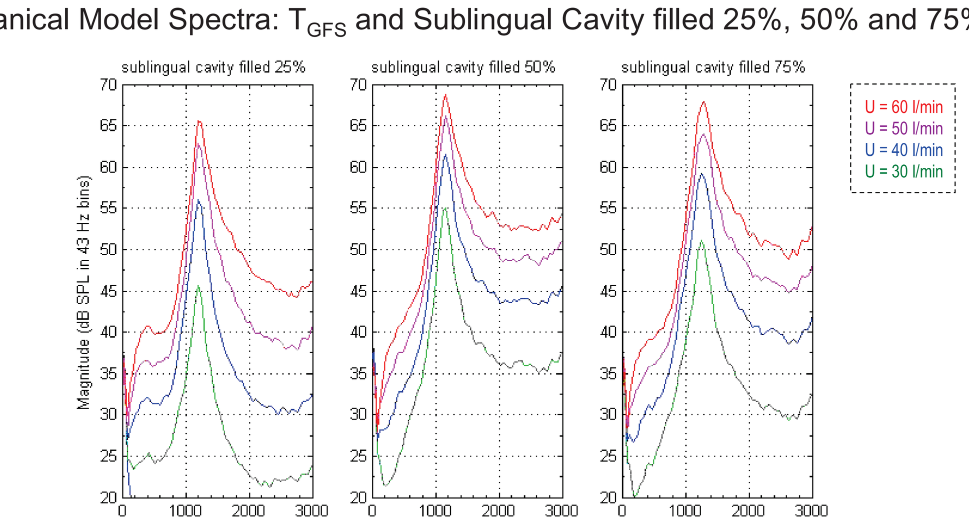
Aerodynamic Setup:
- source: Gast constant pressure air compressor
- regulation: controlled volume velocity monitored by rotameter
- conditioning: laminar flow element (LFE; Meriam 50MJ10-9)
- monitoring: Honeywell pressure sensor measuring instantaneous flow velocity just upstream of model

Acoustic Recording:
- LFE & model placed inside anechoic chamber
- baffle (Plexiglas square 60x60cm) installed at end of 'tract'
- far-field sound pressure measured off-centre, 0.5 m from 'lips'

Results – Mechanical Models



Mechanical Model Spectra: T_{GFS} and Sublingual Cavity filled 25%, 50% and 75%



Observations:
- frequency of main peak does not change with size of sublingual cavity (unexpected)
- spectral shape above main peak varies with SC size
- rate of change of amplitude with flowrate varies with SC size

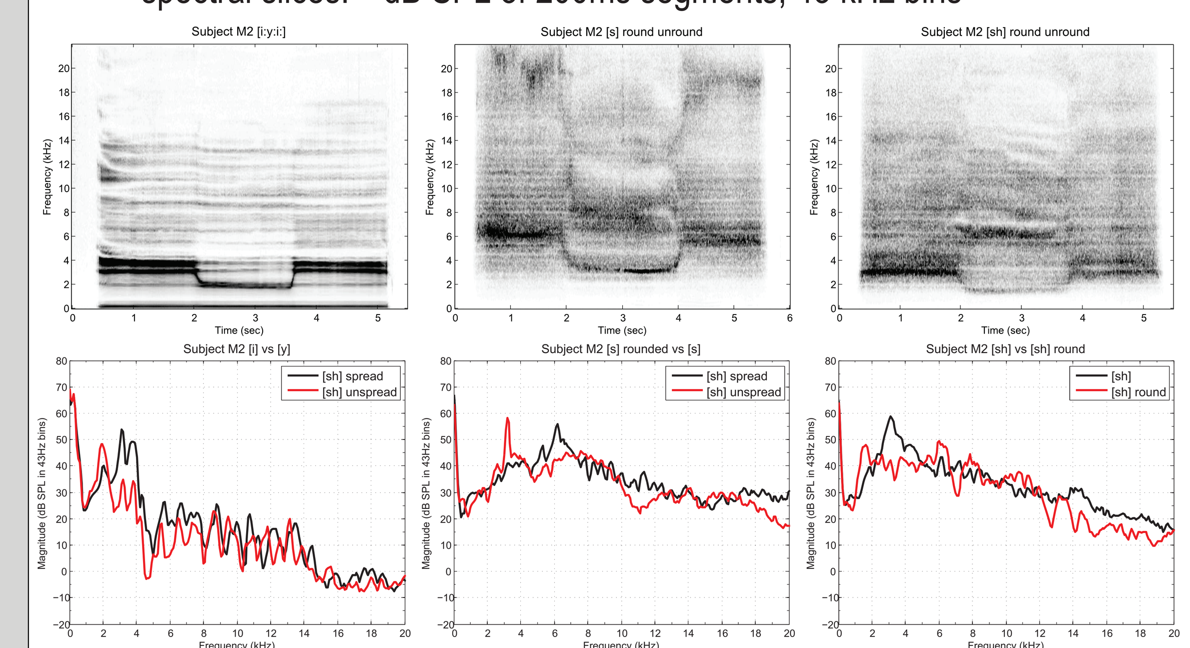
Questions:
- does lack of lips in model cause atypical result?

⇒ investigate role of lip rounding using human subject 'acrobatics'

Acoustic 'Acrobatics'

Objective:
- investigate acoustic consequences of controlled, targeted changes in articulation using corpus with extreme articulatory changes

Method:
- subjects trained, then recorded producing nonsense words
- stimuli presented in IPA, recorded in anechoic chamber
- spectrograms: 512 point FFT from 10ms windows (50% overlap)
- spectral slices: dB SPL of 200ms segments, 43 kHz bins



[i:y:i:]: formant frequencies stay the same
amplitude relationship changes when rounding, protrusion added
changes similar to M2 sibilants in [uFu] context

[s:] ±round: peak frequencies stay approximately the same
amplitude relationship changes

[ʃ:] ±round: [ʃ] peak changes from 3 kHz to 1.5 kHz [+rnd]; spectral shape flatter. Resonances mostly not at same frequencies

Subject M2: in cases shown, constriction stays same and so (we assume) does SC; lip rounding/protrusion causes spectral changes

Discussion

Subject M1: displays a 'classic' and consistent [s-ʃ] difference in lip rounding and SC presence. Peak frequencies differ by 3-4 kHz from [s] to [ʃ]. SC variations in [s] cause peak differences of ~ 300 Hz.

Subject M2: displays significant lip differences in [u] vowel context; SC cavity noticeable in [ʃ], changes somewhat with vowel. Peak freqs vary comparably with vowel context and with sibilant. Clean SC difference but large acoustic variation.

Mech Model: SC changes ⇒ more source- than filter-related.

Acrobatics: show effect of lip shape on a given tongue-SC config (M2). Spectral shape can vary considerably with rounding only.

Overall: tongue groove, constriction relative to teeth, and lip position offer many ways to manipulate source and filter properties of sibilants. Sublingual cavity is only one factor affecting [s-ʃ] acoustic differences. Spectral shape offers many clues to articulatory-acoustic relationship not captured by first spectral moment.

References

Perkell, J. S., Matthies, M. L., Tiede, M., Lane, H., Zandipour, M., Marrone, N. (2004). The distinctness of speakers' /s/-/ʃ/ contrast is related to their auditory discrimination and use of an articulatory saturation effect. *JSLHR* 47(6): 1259-1269

Perkell, J.S., Boyce, S.E., Stevens, K.N. (1979) "Articulatory and acoustic correlates of the [s-sh] distinction," *Speech Communication Papers*, eds. J.J.Wolf & D.H.Klatt, New York: ASA: 109-113

Proctor, M.I., Shadle, C.H. and Iskarous, K. (2008) A method of co-registering multiple magnetic resonance imaged vocal tract volumes for fricatives. *Proc. Acoustics '08*, Paris: 5093-5098

Shadle, C.H. (1991) "The effect of geometry on source mechanisms of fricative consonants." *J.Phon.* 19: 409-424

Shadle, C.H., Proctor, M.I. and Iskarous, K. (2008a) An MRI study of the effect of vowel context on English fricatives. *Proc. Acoustics '08*, Paris: 5099-5104

Shadle, C.H., Berezina, M.A., Proctor, M.I. and Iskarous, K. (2008b) Mechanical models of fricatives based on MRI-derived vocal tract shapes. *Proc. 8th Int'l Sem. Speech Prod.*, Strasbourg, France: 413-416

Acknowledgements

Work supported by NIH grants NIDCD R01 DC 006705 (Haskins) & NIDCD T32 DC00038 (MIT-SHBT)