ACOUSTIC AMBIGUITY LEADS TO ARTICULATORY RE-ANALYSIS: AFFRICATE PALATALIZATION IN HONG KONG CANTONESE

Ping Hei YEUNG

Jonathan HAVENHILL

The University of Hong Kong

NACCL 33

24 June 2021

INTRODUCTION

Hong Kong Cantonese (HKC) has three sibilant phonemes: /s, ts, ts^h/

Studies suggest HKC affricates /ts/ and /ts^h/ (but not the fricative /s/) are becoming palatalized in certain linguistic environments (Yu 2016, Chan 2007, Liu 2010, Lan 2017)

Sound change in progress?

INTRODUCTION

Old system:

- All three sibilants have only one allophone [ts, tsʰ, s]

New system:

- /ts/ and /tsʰ/ have two allophones while/s/ has one:
  - /ts, tsʰ/: Alveolar [ts, tsʰ] vs. post-alveolar [tʃ, tʃʰ]
  - /s/: only alveolar [s]

What is the **conditioning environment** for palatalization?

INTRODUCTION

Conditioning environment?

- Before **rounded** vowels: /ts, tsʰ/ > [tʃ, tʃʰ] / _ [+round] (Chan 2007, Liu 2010)
- Before **back** vowels: /ts, tsʰ/ > [tʃ, tʃʰ] / _ [+back] (Lan 2017)

However, not supported by any acoustic or articulatory data

- Evidence for palatalization comes from impressionistic judgements
- Contradictory findings in articulatory study: no difference in place of articulation across different vowel contexts, always alveolar (Lee & Zee 2010)

OBJECTIVES

- Investigate **affricate palatalization**, an ongoing sound change in HKC, through **acoustic** and **articulatory** analysis
- Research questions:
  - Is affricate palatalization indeed an ongoing sound change in HKC?
  - What is the conditioning environment of the affricate allophones?
  - How are the allophones realized in terms of their articulation?
  - What are the possible phonetic motivations of such a change?

METHODOLOGY

1. Audio recording
   - Acoustic analysis
   - Does the acoustics of affricates differ by **phonological environment** and by **age and gender** of speakers?
   - Examine whether there is a sound change

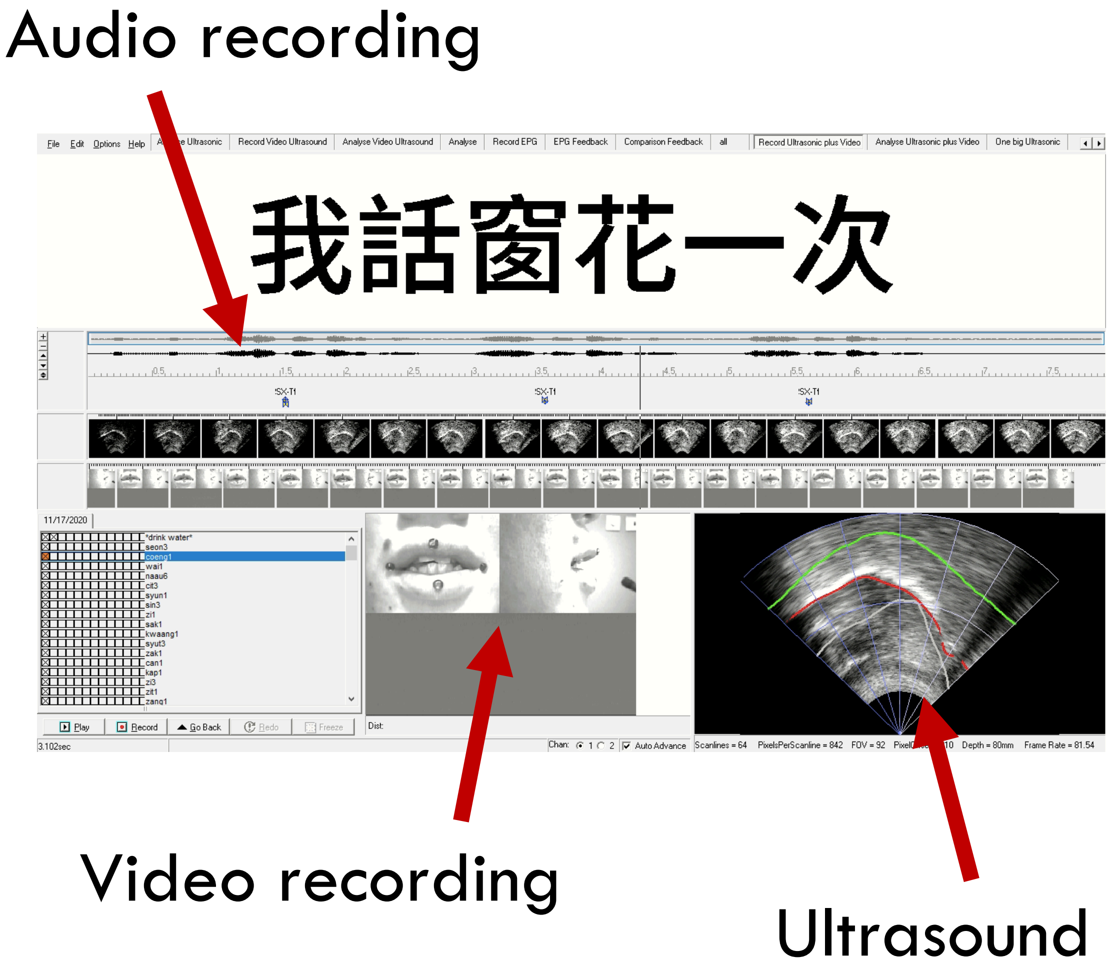
Audio recording

Video recording

Ultrasound

METHODOLOGY

2. Ultrasound tongue imaging
   - Measure tongue position
   - Examine lingual articulation of affricate allophones
3. Video recording
   - Measure lip movements
   - Examine labial articulation of affricate allophones



METHODOLOGY

Materials:

- 126 disyllabic words that start with sibilants: 42 /s/, 45 /ts/, 39 /tsʰ/
- All attested **sibilant** + monophthong CV combinations tested
  - Vowels: [aː, ɐ, ɛː, œː, ɵ, iː, ɪ, yː, ʊ, ɔː]
  - Coda consonants: /#, t, n, k, ŋ/
  - Tone: 1 /˥/ (high level) or 3 (mid level) /˧/
- 44 filler items to obscure aim of study

METHODOLOGY

Participants:

- 12 native speakers of HKC, born between 1956-2001
- No previous history of speech or hearing impairments

Four groups by age and gender:

- 6 older speakers (born 1956-1962)
- 6 younger speakers (born 1997-2001)
- 6 female, 6 male

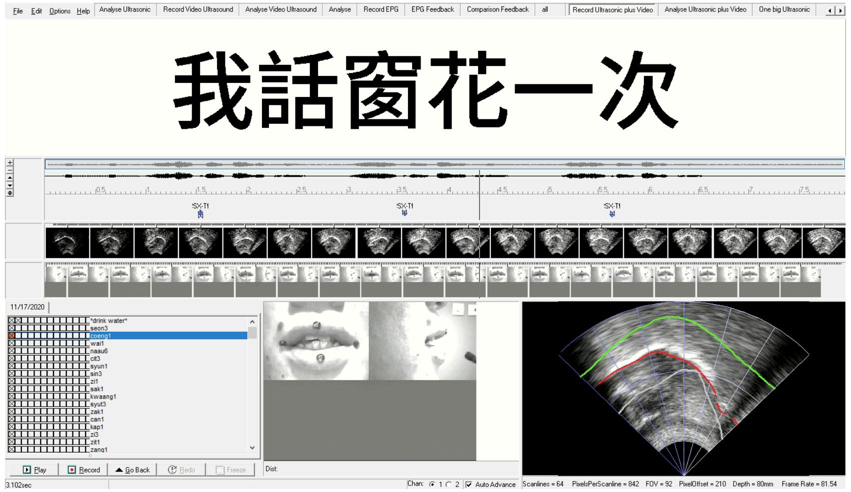
Age group	Female	Male
Older	3	3
Younger	3	3

METHODOLOGY

Procedure:

- Data collected at the University of Hong Kong
- Words embedded in carrier phrase [ŋɔ:˩˧ wa:˨] ___ [jɐt˥ tsʰi:˧] “I say ___ once”
- Words presented to each participant in unique randomized order
- Each word presented once, repeated three times
- 378 target tokens per participant, total of 4,536 tokens across all participants prior to exclusion of mispronounced words etc.

METHODOLOGY



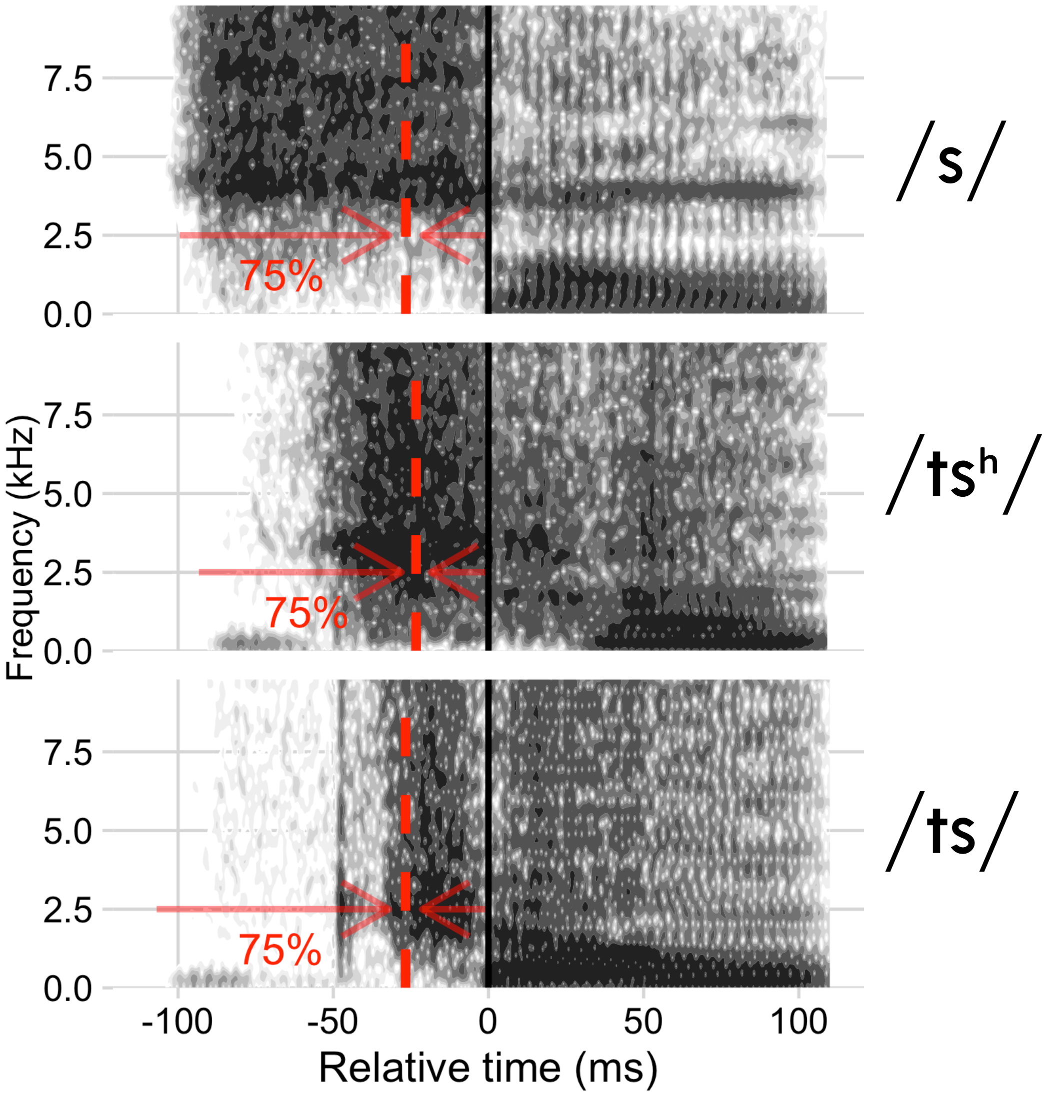
Recording:

- Simultaneous audio, ultrasound, and video recording
- Ultrasound: High speed (84 fps) SonoSpeech Micro ultrasound system, 20mm radius probe
- Audio: Recorded at 22kHz/16-bit with Sennheiser MKE2-P-C Microphone
- Video: coronal and sagittal view of speaker's lips recorded at 60 fps
- Synchronized in Articulate Assistant Advanced (AAA) (Articulate Instruments Ltd. 2012)

METHODOLOGY: ACOUSTICS

Audio:

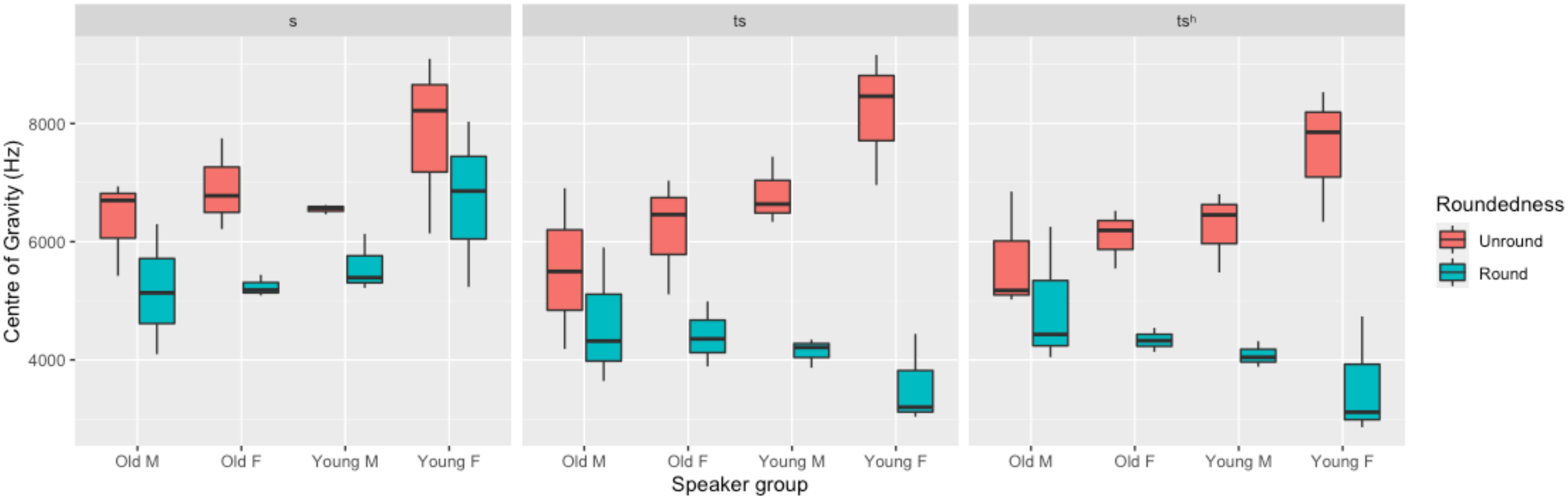
- High-pass filtered at 500Hz in Praat (Boersma & Weenink 2021)
- 20ms Hamming window at 75% of sibilant duration
- Centre of gravity (COG) measured: indicator of sibilant place of articulation (Jongman et al. 2000)
  - ↑ Higher value: more fronted alveolar place, lip spreading
  - ↓ Lower value: more backed post-alveolar place, lip rounding



METHODOLOGY: ACOUSTICS

- Mixed-effect regression model in R with *lmerTest* (Kuznetsova et al. 2020)
- Fixed effects: following vowel (roundedness, height, backness), type of sibilant, coda, tone, duration, age group, gender
  - Random effects: speaker, word, repetition
  - *M* ~ type * roundedness * age group * gender + height + backness + tone + coda + duration + (1|speaker) + (1|word) + (1|repetition)

RESULTS: ACOUSTICS



COG (in Hz) of HKC sibilants by age group and gender

RESULTS: ACOUSTICS

Fixed effects:	Estimate	Std. Error	df	t value	Pr(>|t|)	
roundedness (round)	-1362.8225	106.7232	371.0189	-12.77	**< 2E-16**	***
height (low)	-435.5149	83.1643	104.9933	-5.237	**8.45E-07**	***
height (mid)	11.155	48.3761	107.7861	0.231	0.818072	
backness (front)	-188.7609	60.667	106.9728	-3.111	**0.002388**	**

Effect of following vowel:

- Significant effect of vowel roundedness, **rounded** vowels **lower** COG
- Significant effect of vowel height and backness, low vowels and front vowels lower COG

RESULTS: ACOUSTICS

Fixed effects:	Estimate	Std. Error	df	t value	Pr(>|t|)
round * age (young)	135.7872	119.8423	4240.6129	1.133	0.257258
* gender (F)	-526.562	120.584	4245.8832	-4.367	**1.29E-05** ***
* age (young) * gender (F)	395.282	169.7052	4251.4845	2.329	**0.019894** *

Significant interaction between roundedness of following vowel and gender

- COG of sibilants **lower** for **female** speakers before **rounded** vowels

Significant interaction between roundedness, age and gender

- Across all sibilants, **higher** COG for **younger female** speakers before **rounded** vowels

RESULTS: ACOUSTICS

Fixed effects:	Estimate	Std. Error	df	t value	Pr(>|t|)	
round * /ts/	244.8829	137.1287	519.1839	1.786	0.074717	
* /tsʰ/	351.8602	141.9413	509.4451	2.479	**0.013501**	*
round * age (young) * /ts/	-1897.3421	166.2484	4240.1867	-11.413	**< 2E-16**	***
* /tsʰ/	-1515.3525	172.7982	4239.9851	-8.769	**< 2E-16**	***
* gender (F) * /ts/	-373.8218	166.634	4243.1425	-2.243	**2.49E-02**	*
* /tsʰ/	-465.1248	173.0502	4241.8786	-2.688	**0.00722**	**
* age (young) * gender (F) * /ts/	-1432.938	234.6275	4244.6985	-6.107	**1.10E-09**	***
*/tsʰ/	-1260.5438	244.2515	4245.9021	-5.161	**2.57E-07**	***

Significant interaction between **type of sibilant**, roundedness and age/gender:

- COG of /ts, tsʰ/ **lower** before rounded vowels for **younger** speakers, **female** speakers, **younger female** speakers

DISCUSSION: ACOUSTICS

- **Roundedness** of the following vowel conditions sibilant COG
  - Consistent with Chan (2007) and Liu (2010)
- In general, sibilants more palatalized/labialized before **rounded** vowels
  - Palatalization, or co-articulatory effect of vowel rounding?

DISCUSSION: ACOUSTICS

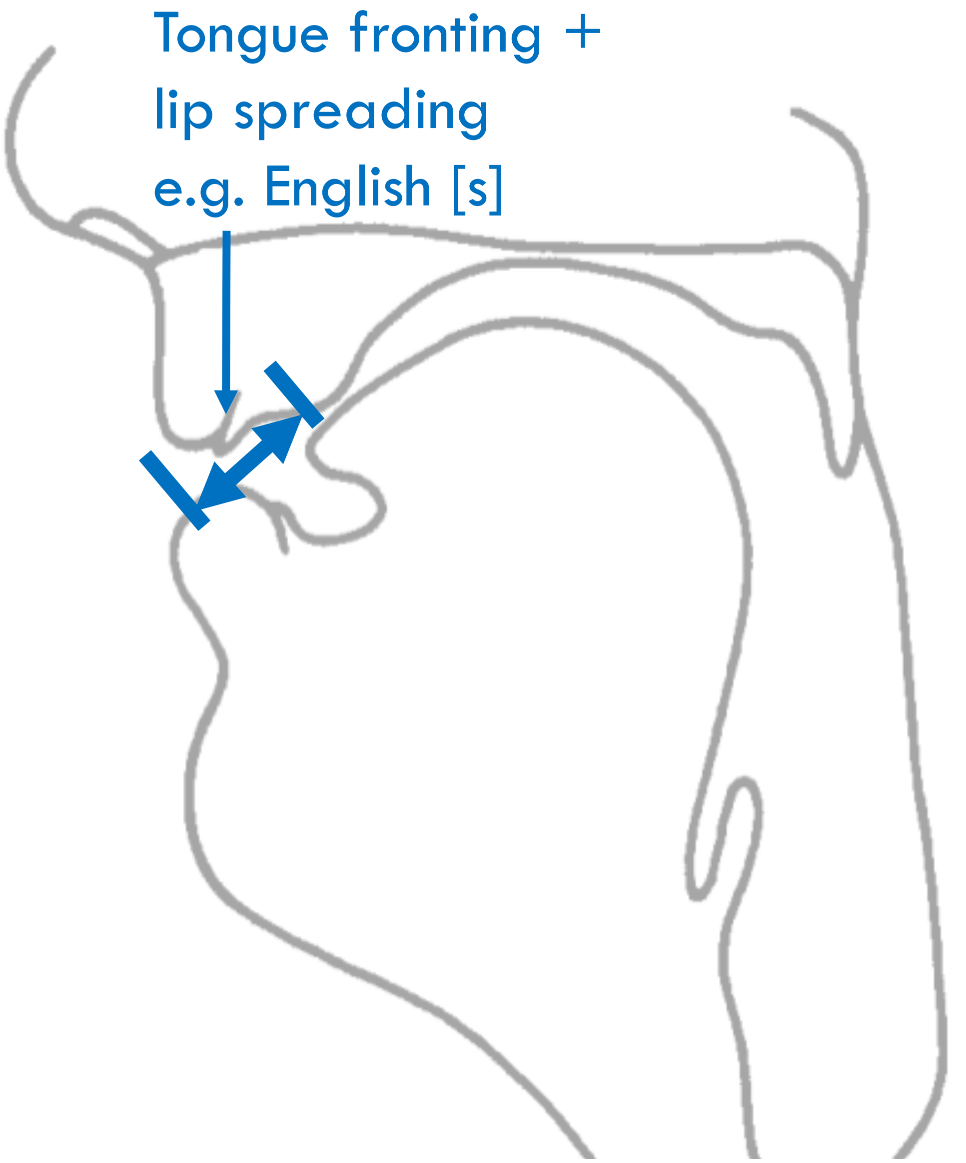
- **Affricates** /ts, ts^h/ more palatalized/labialized than fricative /s/ before **rounded** vowels for **younger** and **female** speakers
- Younger females tend to be innovative speakers. Their speech pattern may indicate direction of sound change (Labov 1990)
  - Consistent with Chan (2007) and Liu (2010)
  - Change from co-articulatory effect to allophonic split?
  - Palatalization or labialization?

DISCUSSION: ACOUSTICS

Differences in COG can be result of any gestures that changes length of front cavity of vocal tract

Both tongue and lip configuration change length of front cavity:

- **Tongue fronting** or **lip spreading,** front cavity shortens, **higher** COG



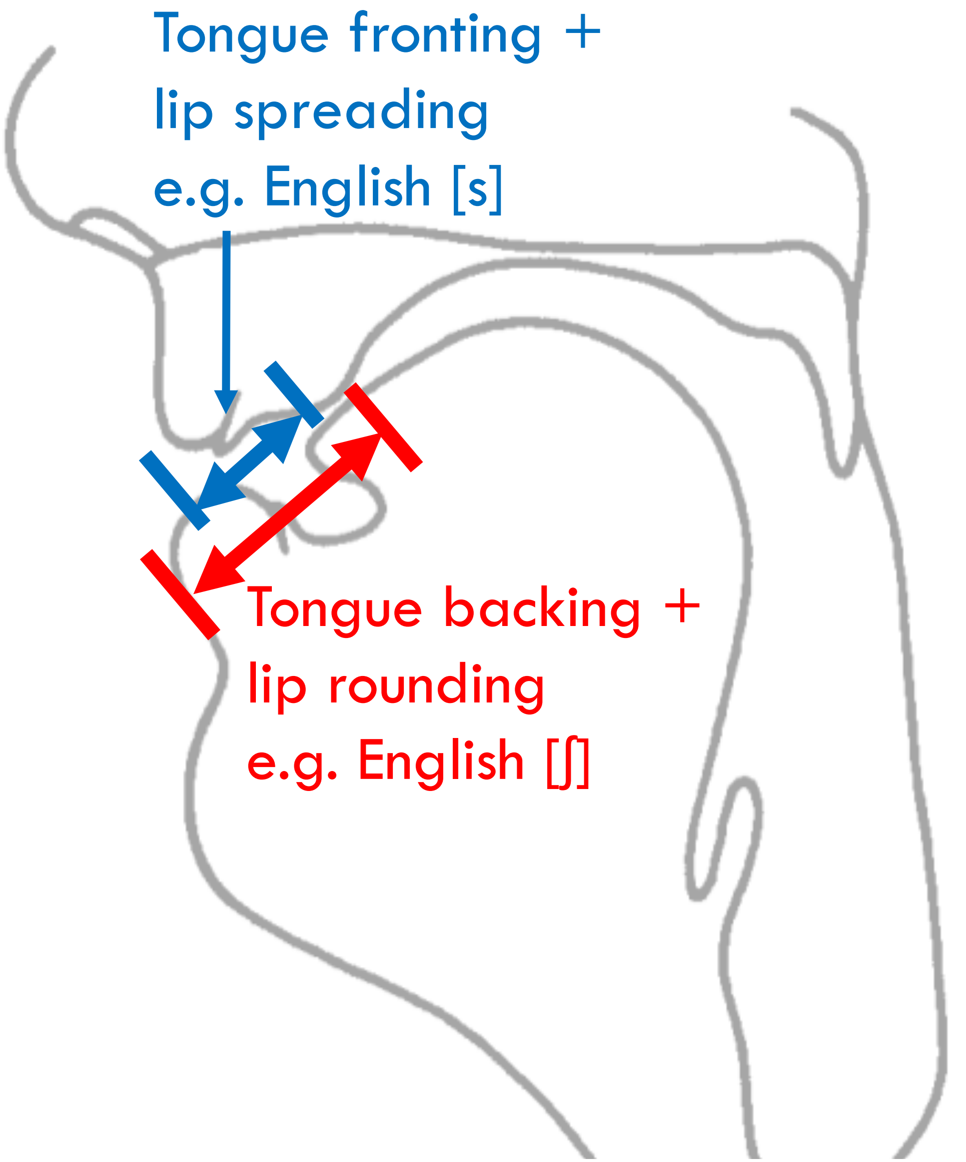
Figures adapted from Ladefoged & Johnson (2010)

DISCUSSION: ACOUSTICS

Differences in COG can be result of any gestures that changes length of front cavity of vocal tract

Both tongue and lip configuration change length of front cavity:

- **Tongue fronting** or **lip spreading**, front cavity shortens, **higher** COG
- **Tongue backing** or **lip rounding**, front cavity lengthens, **lower** COG



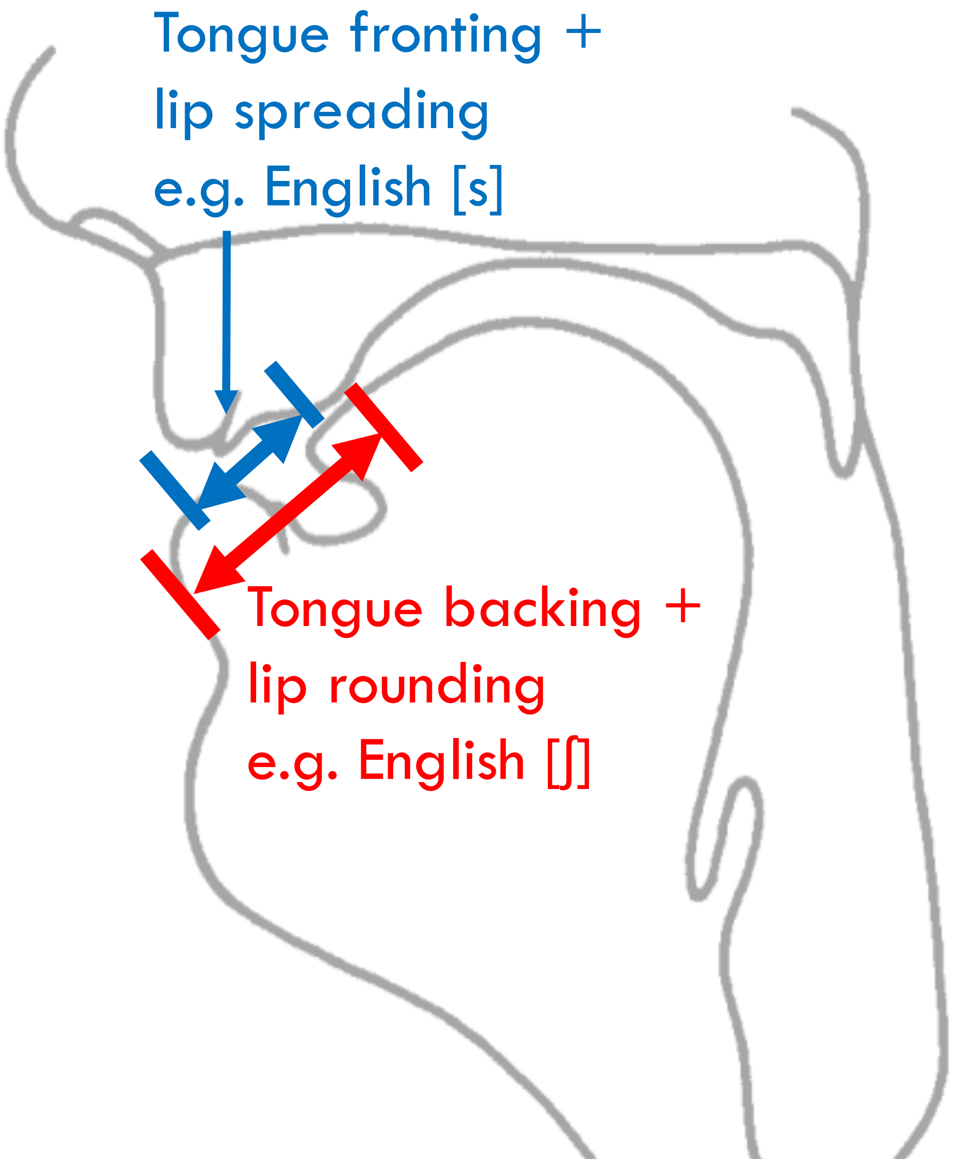
Figures adapted from Ladefoged & Johnson (2010)

DISCUSSION: ACOUSTICS

Acoustic analysis alone **cannot** disambiguate contributions of tongue and lip in the articulation of sibilants

What causes age and gender differences in sibilant COG before rounded vowels?

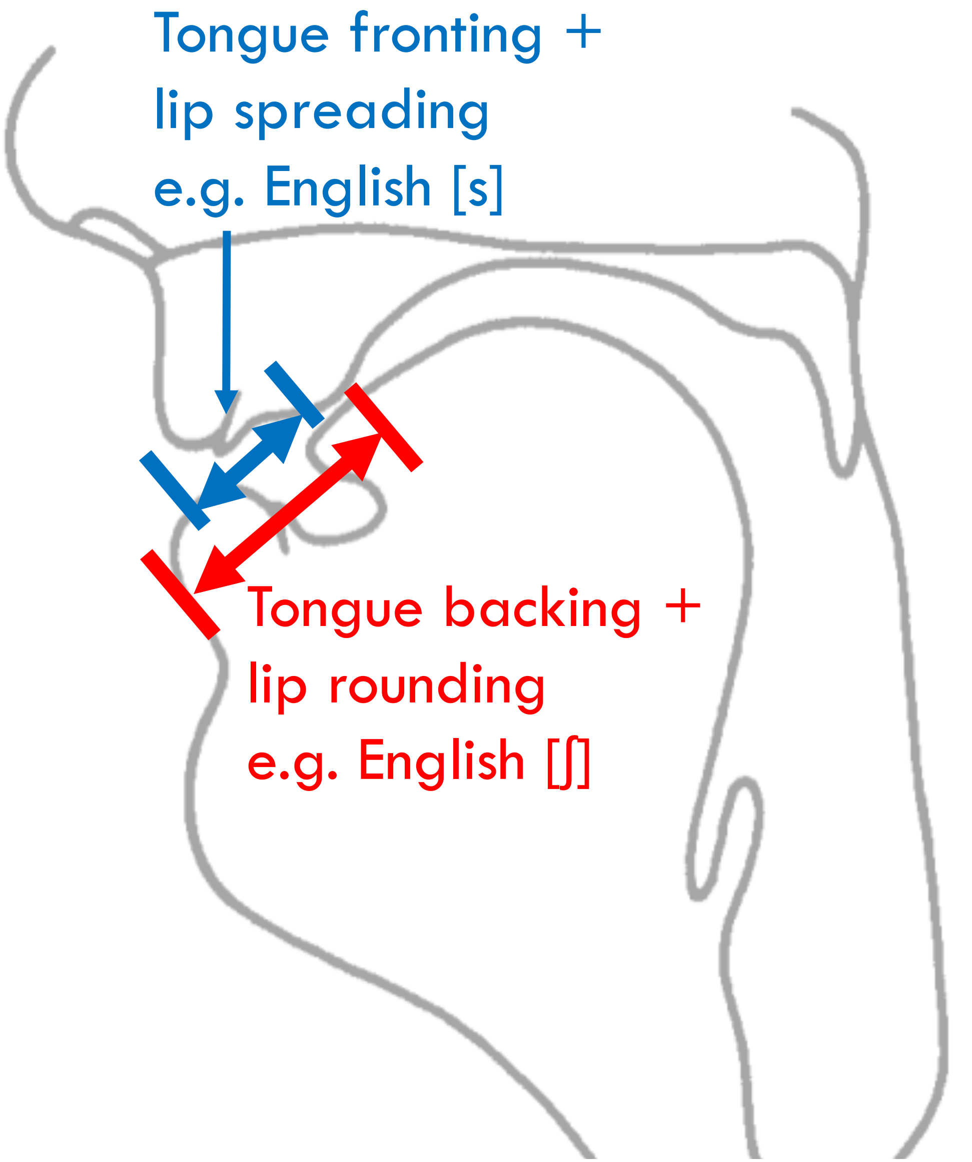
Articulatory measurement with **ultrasound tongue imaging** and **lip video** necessary



Figures adapted from Ladefoged & Johnson (2010)

DISCUSSION: ACOUSTICS

- Lip rounding only?
  - Hypothesis 1: co-articulatory effect, anticipatory lip rounding
  - Hypothesis 2: phonologization of labialized allophones [ts^w, ts^{hw}]
- Tongue backing?
  - Hypothesis 3: phonologization of post-alveolar allophones [tʃ, $tʃ^h$]



Figures adapted from Ladefoged & Johnson (2010)

METHODOLOGY: LIP MEASUREMENT

Beads affixed to upper lip, lower lip and two corners of the mouth to track movement of lips

Every other video frame from onset of sibilant to mid-point of following vowel extracted and measured

33.3ms interval between measured frames (30 fps)

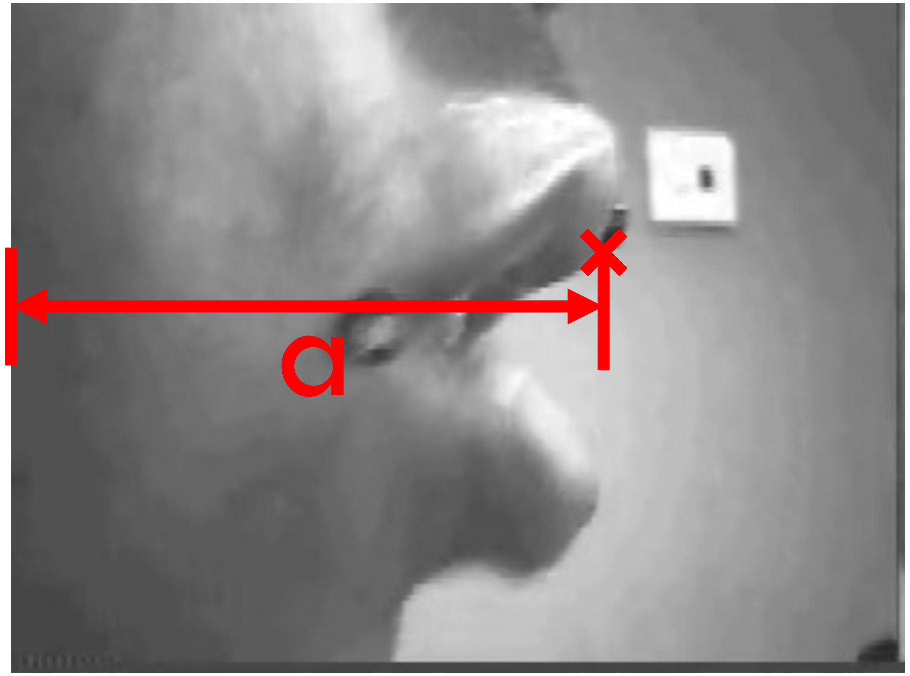
29,969 frames entered analysis in total

METHODOLOGY: LIP MEASUREMENT

Measure of lip protrusion: Horizontal distance between left edge of video frame and upper lip

- Longer distance: more protruded (round)
- Shorter distance: less protruded (unround)

Movement of upper lip independent of that of other articulators, better reflects labial gestures (Farnetani 1999)



Articulatory measurements for lip protrusion. The degree of protrusion is defined as (a)

METHODOLOGY: LIP MEASUREMENT

Upper lip protrusion was z-score normalized by speaker

Generalized additive mixed modelling (GAMMs) in R with mgcv (Wood 2017) and itsadug (van Rij et al. 2020), model criticism performed with compareML() and gam.check():

```
bam(upper.x.norm ~ vowelrounding +
    s(time.rel.fric) + s(time.rel.fric, by = vowelrounding) +
    onset + s(time.rel.fric, by = onset) +
    s(duration, by = vowelrounding) +
    ti(time.rel.fric, duration, by = vowelrounding) +
    age.gender + s(time.rel.fric, by = age.gender) +
    s(time.rel.fric, speakerid, by = vowelrounding, bs = 'fs', m = 1),,
    AR(1) error model, family = scaled-t)
```

vowel rounding

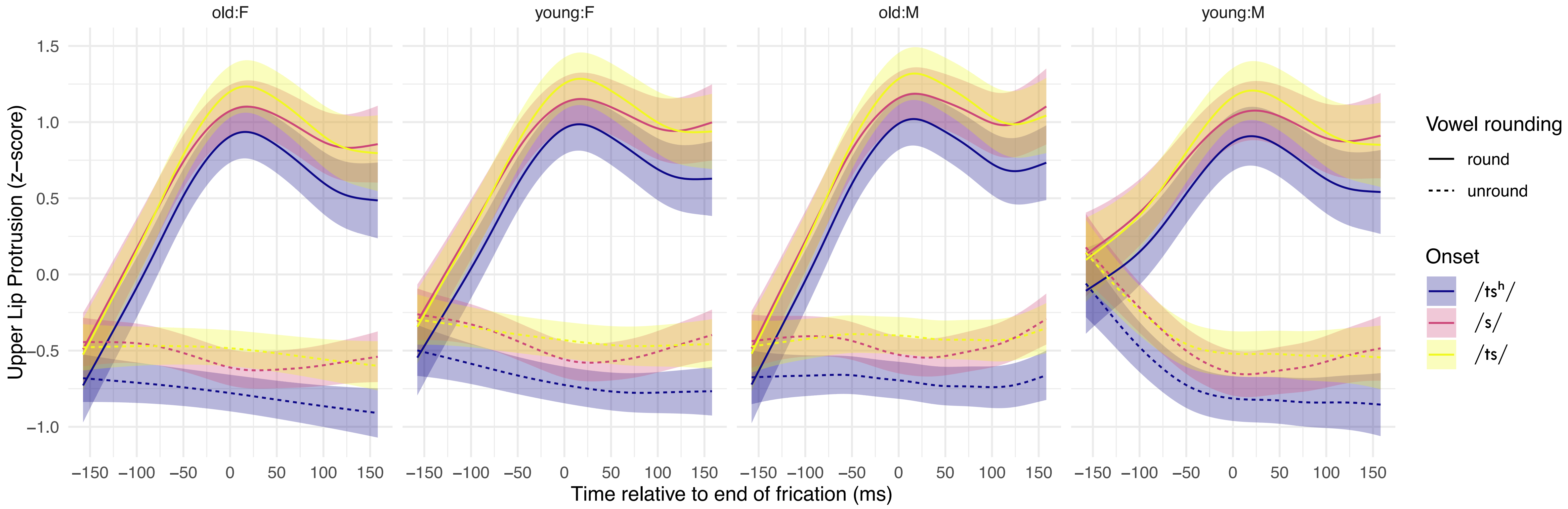
onset (s, ts, or ts^h)

duration

age * gender

random speaker effect

RESULTS: LIP MEASUREMENT



Fitted smooths with 95% CI for upper lip protrusion for HKC sibilants by age, gender, and rounding

RESULTS: LIP MEASUREMENT

Significant difference in labial coarticulation for sibilants in round vs. unround environments

- More protruded before round vowels than unround vowels

Across three sibilants, **no difference in timing or degree of lip protrusion**

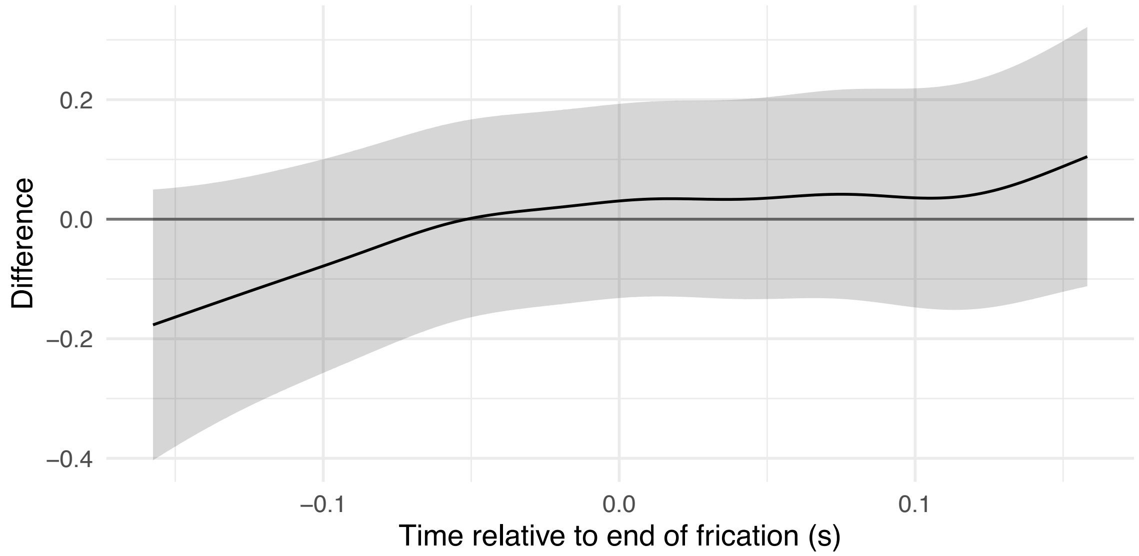
- Sibilants followed by round vowels exhibit labial coarticulation, with lip protrusion maximum aligned with onset of vowel
- For sibilants followed by unrounded vowels, lip remains retracted

DISCUSSION: LIP MEASUREMENT

No diachronic change in timing or magnitude of lip rounding gesture

Results do not support Hypothesis 2

- No evidence that labialization has been phonologized as [ts^w, ts^{hw}]
- Rounding remains the result of anticipatory lip rounding before round vowels



Non-significant difference in fitted lip protrusion smooths for [ts^h] as produced before round vowels by Older Male vs. Younger Female speakers

METHODOLOGY: ULTRASOUND

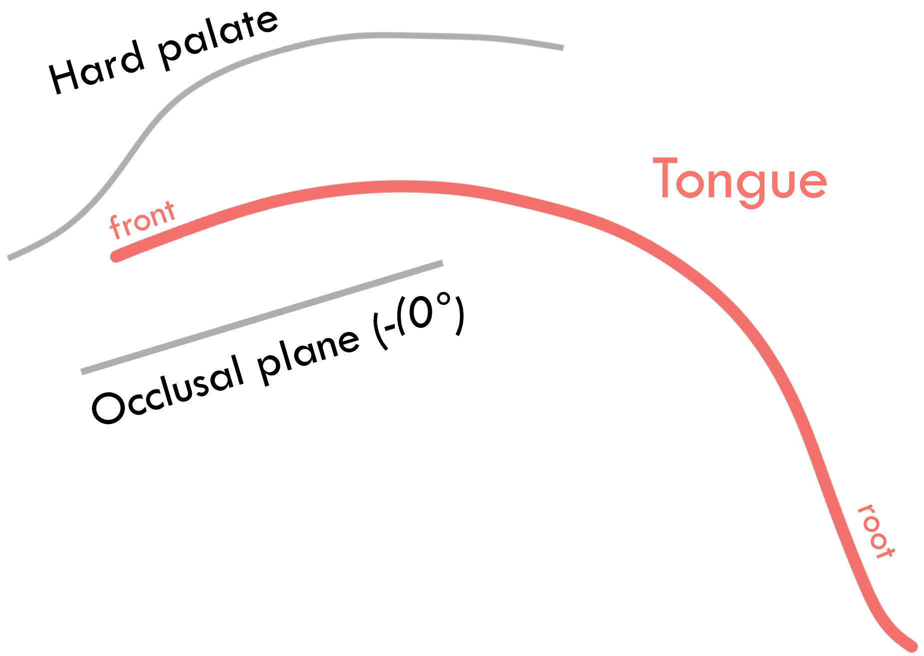
All ultrasound frames from onset of sibilant to end of frication batch splined in AAA (Articulate Instruments Ltd. 2012)

11.9ms interval between frames (84 fps)

55,732 frames entered analysis in total

Tongue splines rotationally corrected relative to horizontal bite plane (Scobbie et al. 2011), distance from probe z-score normalized

Generalized additive mixed modelling (GAMMs) and polar SSANOVA (Gu 2002, Davidson 2006, Mielke 2015) in R

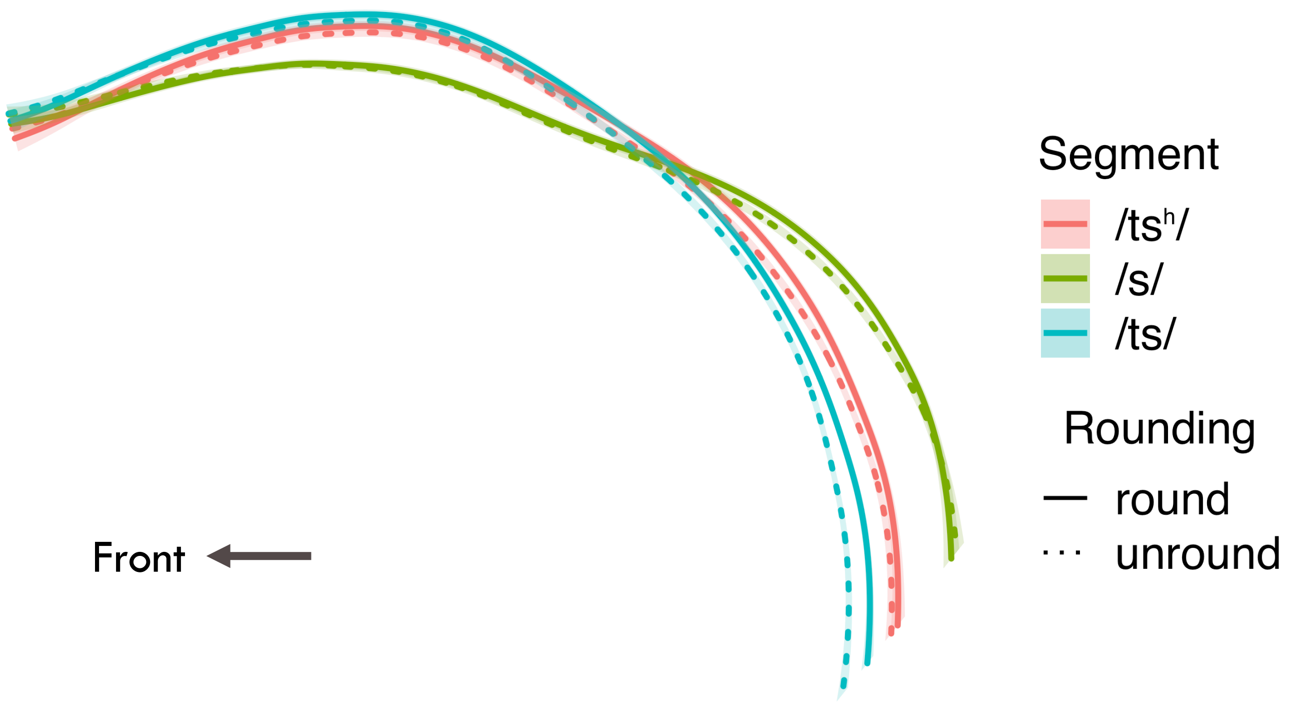


Tongue contour, palate trace, and occlusal plane for speaker 3, rotated 17°

RESULTS: ULTRASOUND SSANOVA

Speaker 9 (**Older male**):

- **Same** tongue position before rounded vs. unrounded vowels for all three sibilants
- **Different** tongue positions for affricates and fricative
- /s/ more **fronted**, /ts, ts^h/ more **backed**

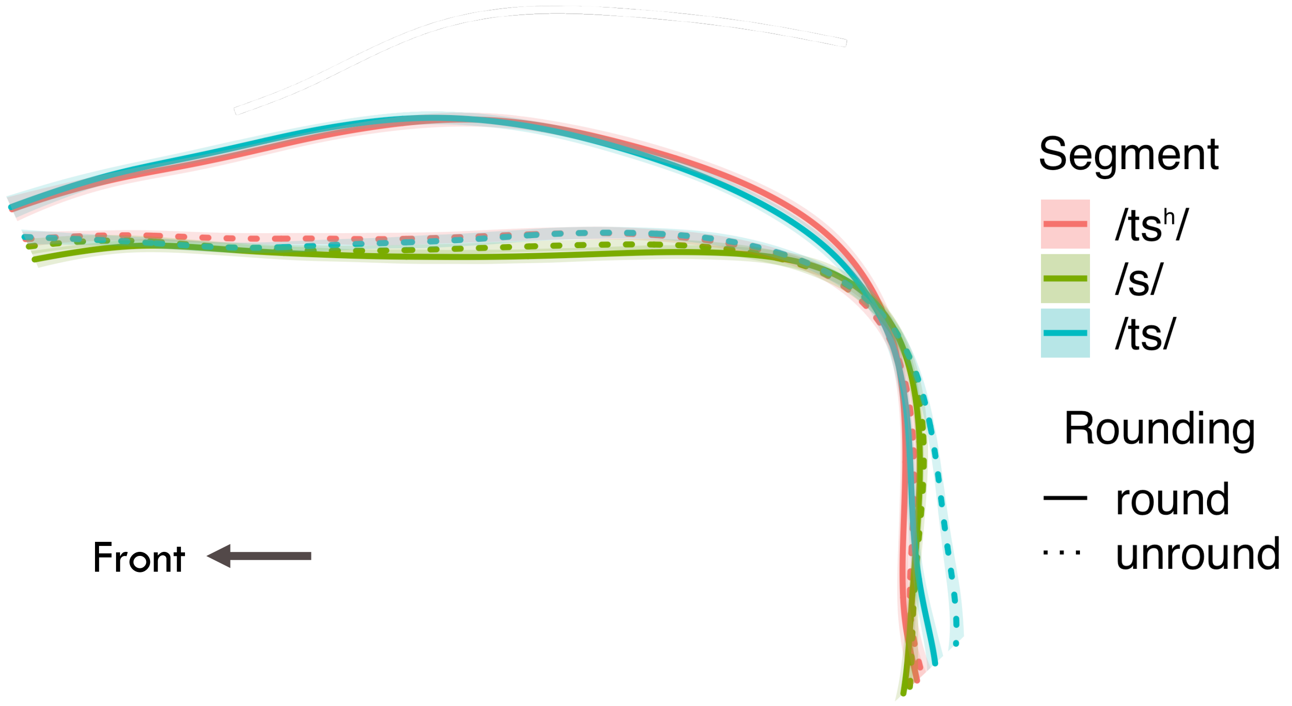


Polar SSANOVA tongue contours of sibilants by rounding for speaker 9 (Older male)

RESULTS: ULTRASOUND SSANOVA

Speaker 5 (**Younger female**):

- **Different** tongue positions before rounded vs. unrounded vowels for /ts, ts^h/
- **Backed** before **rounded** vowels
- **Fronted** before **unrounded** vowels

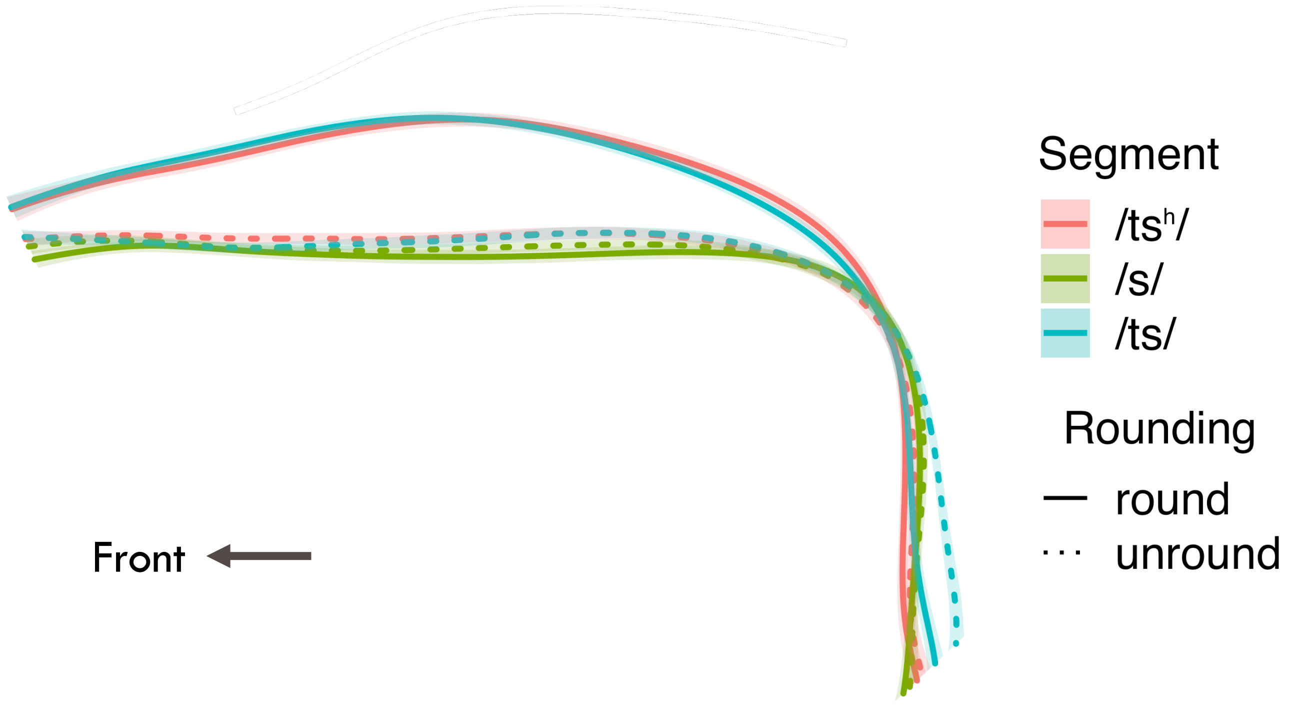


Polar SSANOVA tongue contours of sibilants by rounding for speaker 5 (Younger female)

RESULTS: ULTRASOUND SSANOVA

Speaker 5 (**Younger female**):

- **Same** tongue position for affricates before unrounded vowel and fricative
- **Same** tongue position for /s/ regardless of vowel rounding



Polar SSANOVA tongue contours of sibilants by rounding for speaker 5 (Younger female)

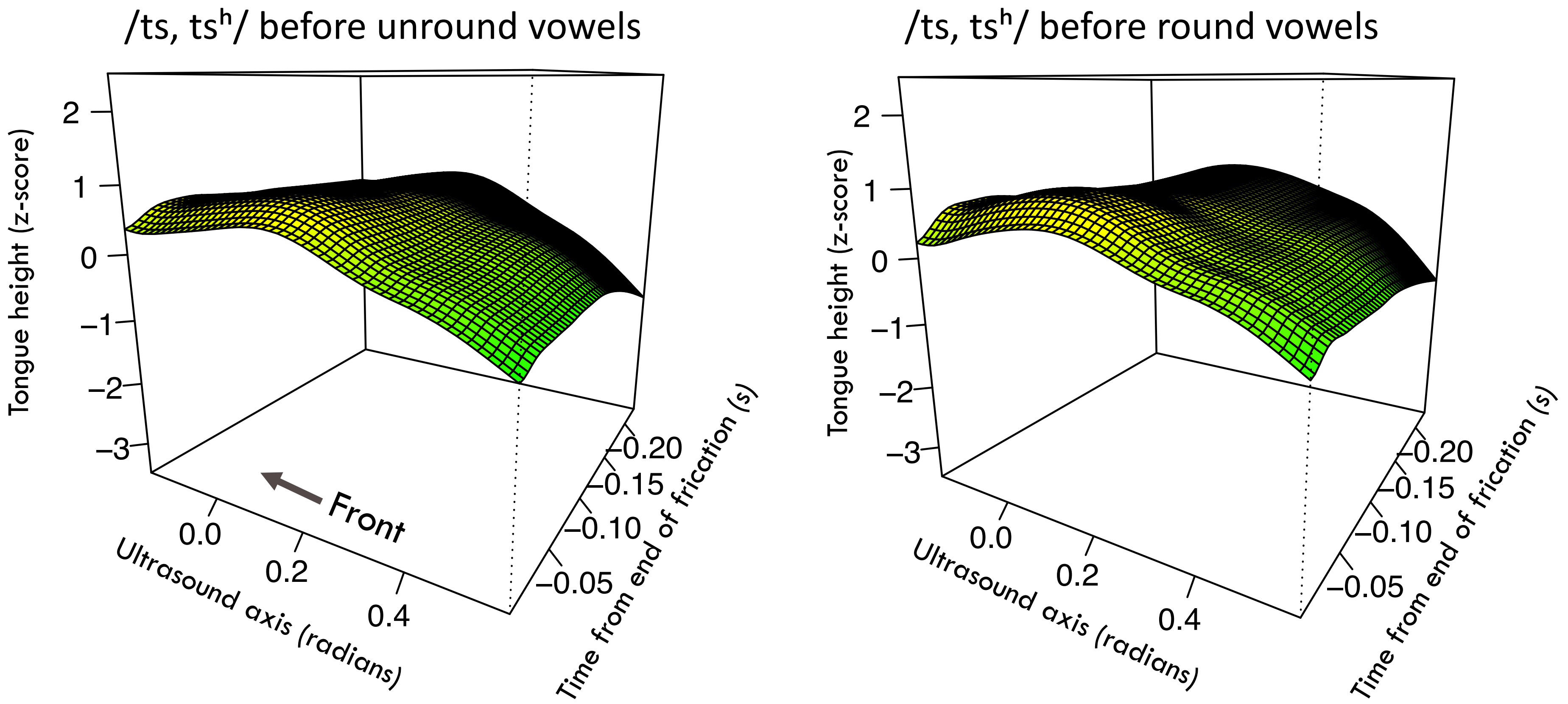
METHODOLOGY: ULTRASOUND GAMM

Spatiotemporal GAMMs (Carignan et al. 2020) in R with mgcv (Wood 2017) and itsadug (van Rij et al. 2020):

```
bam(tongue.height ~ interaction(vowelrounding, sibilanttype) +
    te(time, tongue.location) +
    te(time, tongue.location, by = interaction(vowelrounding, sibilanttype)) +
    s(time.rel.fric, speakerid, by = interaction(vowelrounding, sibilanttype), bs="fs", m=1) +
    s(theta.rot, speakerid, by = interaction(vowelrounding, sibilanttype), bs="fs", m=1),
    AR(1) error model, family = scaled-t)
```

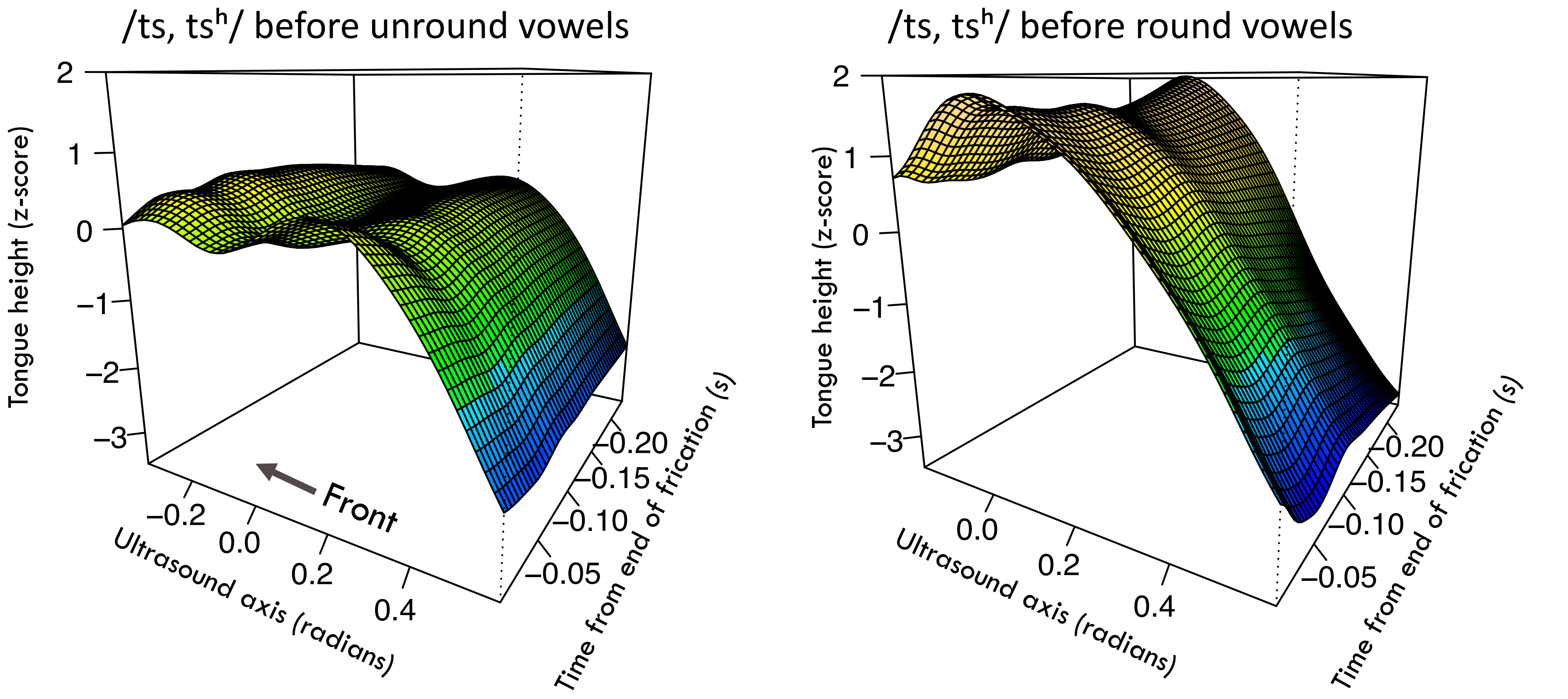
vowel rounding, time, tongue location, sibilant type (fricative vs. affricate), random speaker effect

RESULTS: ULTRASOUND GAMM



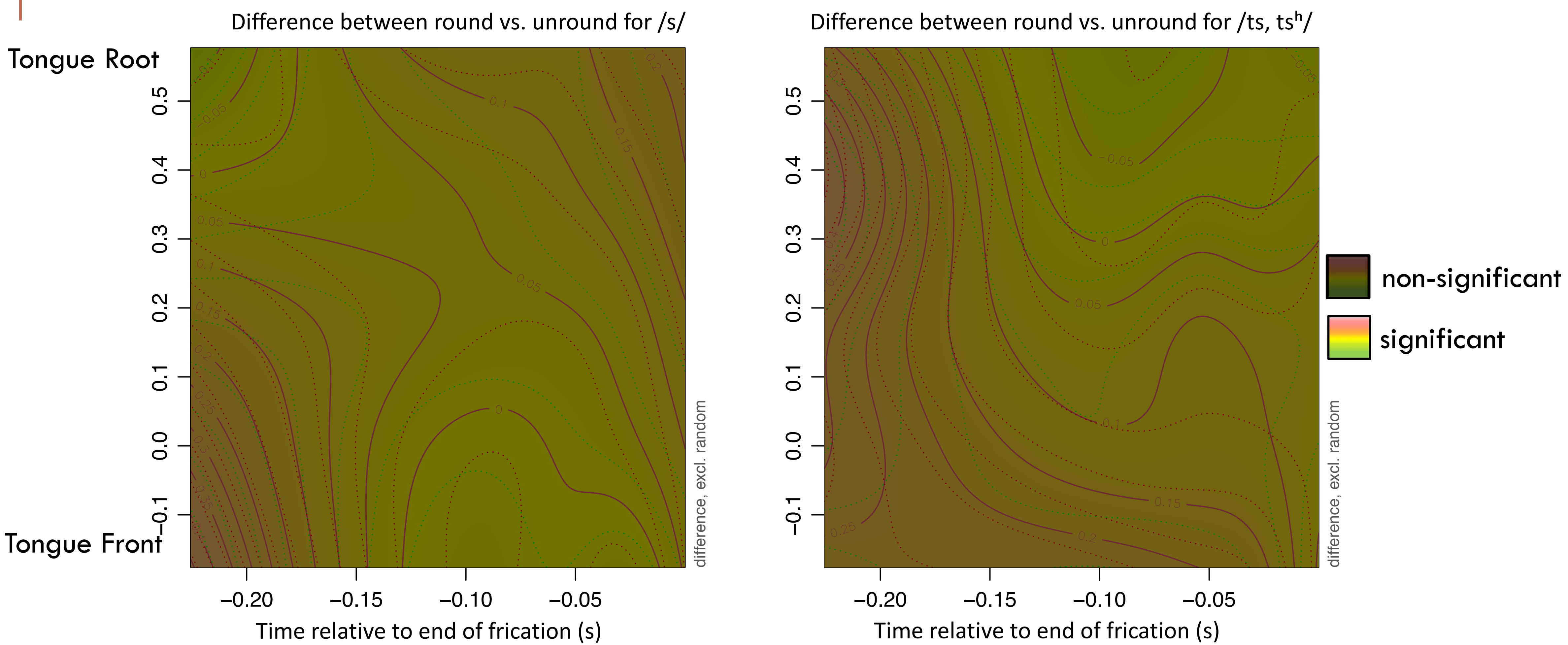
Fitted surfaces for /ts, ts^h/ before round vs. unround vowels, Older Male speakers

RESULTS: ULTRASOUND GAMM



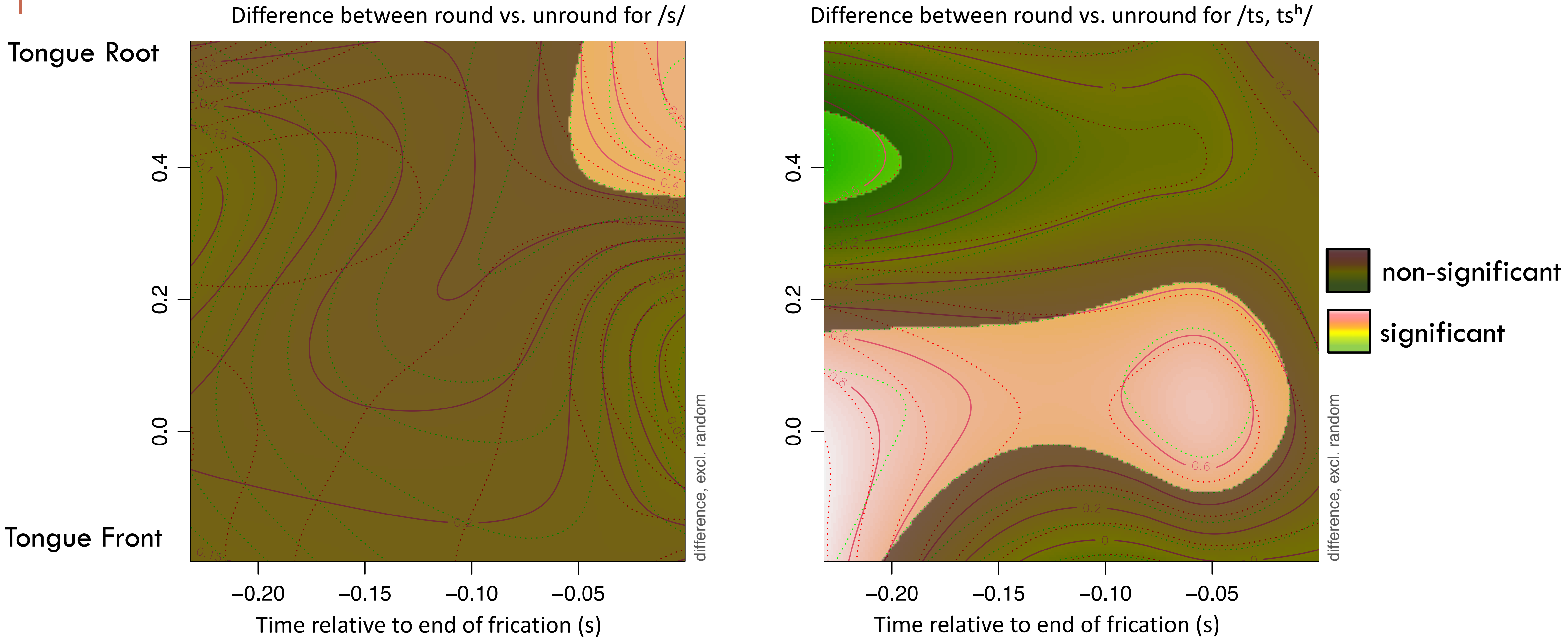
Fitted surfaces for /ts, tsʰ/ before round vs. unround vowels, Younger Female speakers

RESULTS: ULTRASOUND GAMM



GAMM heatmap of tongue height difference (color) over time (x-axis) across the tongue surface (y-axis) for Older Male speakers

RESULTS: ULTRASOUND GAMM



GAMM heatmap of tongue height difference (color) over time (x-axis) across the tongue surface (y-axis) for Younger Female speakers

DISCUSSION: ULTRASOUND

For most advanced **younger female** speakers: tongue position for affricates /ts, ts^h/ varies by roundedness of following vowel, consistent with acoustic measurements

- Before unrounded vowels: alveolar [ts, ts^h]
- Before rounded vowels: (palatalized) post-alveolar [tʃ, $tʃ^h$]

For most conservative **older** male speakers: same tongue position for affricates regardless of roundedness of following vowel

DISCUSSION: ULTRASOUND

Older Female and **Younger Male** speakers are intermediate; some speakers show an allophonic split, some do not.

No palatalization of fricative /s/

- For all speakers, /s/ produced at the same alveolar place of articulation, regardless of roundedness of following vowel
- Consistent with Yu (2016)

GENERAL DISCUSSION

- Hypothesis 1: co-articulatory effect, anticipatory lip rounding — **rejected**
- Hypothesis 2: phonologization of labialized allophones [tsʷ, tsʰʷ] — **rejected**
- Hypothesis 3: phonologization of post-alveolar allophones [tʃ, tʃʰ] — **supported**

GENERAL DISCUSSION

Results of both **acoustic** and **articulatory** analyses show evidence for a phonologized allophonic split in HKC

Conservative system: /ts, ts^h/ as [ts, ts^h] in all contexts

Innovative system: /ts, ts^h/ > [tʃ, $tʃ^h$] / _ [+round]

GENERAL DISCUSSION

Is affricate allophonic restructuring induced by **language contact**?

Liu (2010) and Lan (2017): Influence of English

- Onset of change coincide with the period when Cantonese-English bilingualism became widespread in HK (late 20^{th} century)
- The only affricates in English are post-alveolar /tʃ, dʒ/

GENERAL DISCUSSION

Yet, does not explain why palatalization only occurs **before rounded vowels** but not in other environments

- English /tʃ, dʒ/ post-alveolar in all linguistic environments
- e.g. cheese vs. choose, check vs. chalk
- Phonetic motivation for restructuring of affricate allophonic system?

GENERAL DISCUSSION

Lip rounding and tongue backing share similar acoustic effect

- Both lower spectral peak and COG

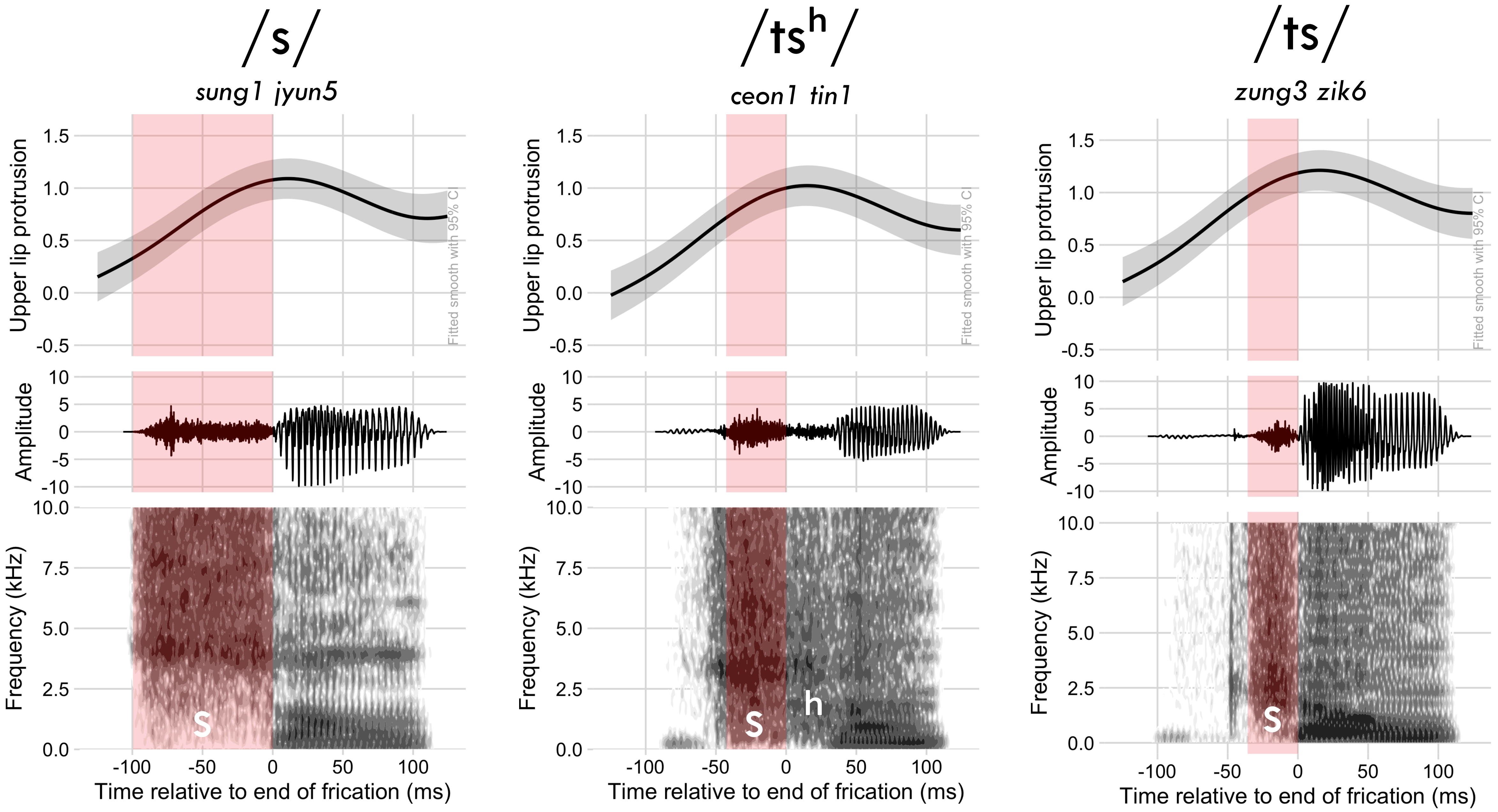
- Younger speakers reinterpret lowering of COG from anticipatory lip rounding as tongue backing (palatalization)
- Hypo-correction: accentuate articulatory difference between affricates before rounded vs. unrounded vowels
- Palatalization became feature of affricates, two allophones formed

GENERAL DISCUSSION

Why does palatalization target the **affricates** but not the fricative?

- Frication phase of affricates less than half that of fricatives
- Anticipatory lip rounding overlaps with entire audible portion of affricates

Sibilant	Mean overall duration	Mean frication duration
/s/	115 ms	115 ms
/ts/	118 ms	44 ms
/ts^h/	153 ms	49 ms



Spectrogram and waveform for representative pre-round sibilant tokens, with GAMM smooths for upper lip protrusion

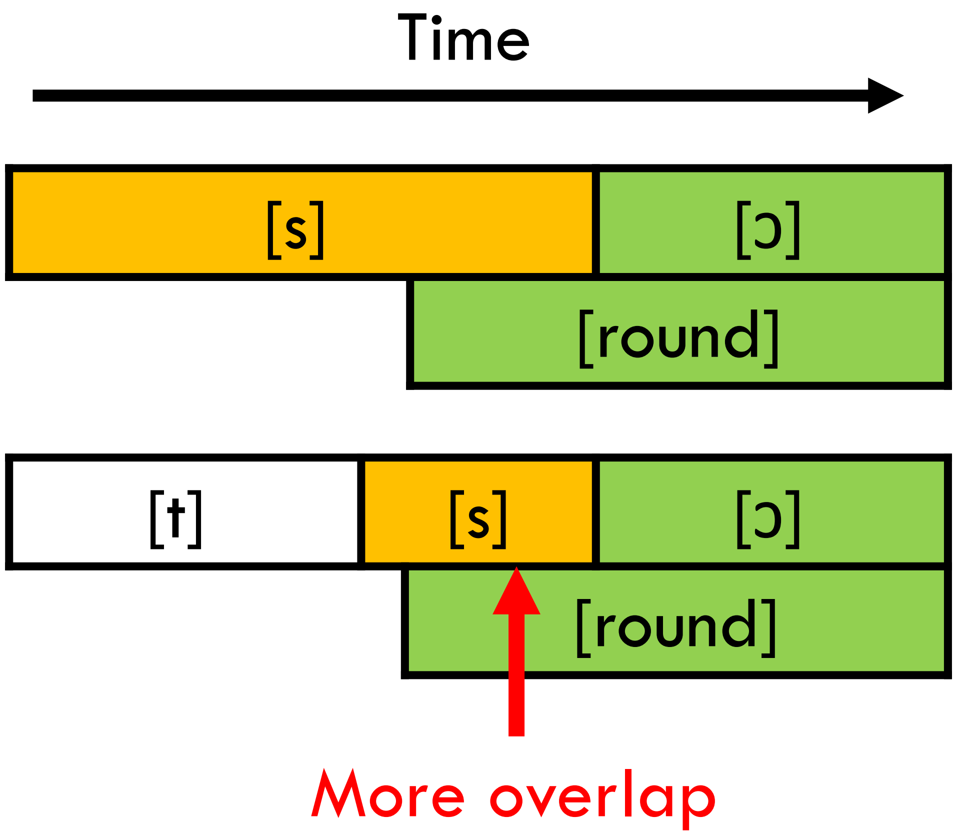
GENERAL DISCUSSION

Affricates and fricatives exhibit differing degrees of gestural overlap with the vowel rounding gesture (cf. Browman & Goldstein 1989)

As a result, acoustic ambiguity is greater for affricates than for fricatives

- Listeners may recover place of articulation cues from first half of /s/, while the first half of /ts, ts^h/ is silent

This ambiguity has led to the introduction of an allophonic split among younger speakers



THANK YOU!

Contact:

Ping Hei Yeung

py101@georgetown.edu

Jonathan Havenhill

jhavenhill@hku.hk

REFERENCES

Articulate Instruments Ltd. (2012). Articulate Assistant Advanced user guide: Version 2.14.

Boersma, P., & Weenink, D. (2021). Praat: doing phonetics by computer [Computer program]. Version 6.1.38, retrieved 2 January 2021 from http://www.praat.org/

Browman, C. P., & Goldstein, L. (1989). Articulatory Gestures as Phonological Units. *Phonology, 6*(2): 201-251.

Carignan, C., Hoole, P., Kunay, E., Pouplier, M., Joseph, A., Voit, D., Frahm, J. and Harrington, J., 2020. Analyzing speech in both time and space: Generalized additive mixed models can uncover systematic patterns of variation in vocal tract shape in real-time MRI. *Laboratory Phonology: Journal of the Association for Laboratory Phonology, 11*(1).

Chan, K. (2007). *The Sound Change of [ts, tsʰ, s] to [tʃ, tʃʰ, ʃ] in Hong Kong Cantonese* (MPhil thesis, The Hong Kong University of Science and Technology, Hong Kong).

Davidson, Lisa. 2006. Comparing Tongue Shapes from Ultrasound Imaging Using Smoothing Spline Analysis of Variance. *The Journal of the Acoustical Society of America 120* (1): 407–415.

Farnetani, E. (1999). Labial coarticulation. In W. J. Hardcastle & N. Hewlett (Eds.), *Coarticulation: Theory, Data and Techniques (pp. 144–163)*. chapter, Cambridge: Cambridge University Press.

REFERENCES

Gu, Chong. 2002. *Smoothing Spline ANOVA Models*. New York: Springer.

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America, 108*(3), 1252-1263.

Kuznetsova, A., Brockhoff, P. B., Christensen, R., & Pødenphant, S. (2020). “lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). R package version 3.1-2,” available at http://CRAN.R-project.org/package.lmerTest.

Ladefoged, P., & Johnson, K. (2010). *A course in phonetics*. Boston, MA: Cengage Learning.

Labov, W. (1990). *The intersection of sex and social class in the course of linguistic change. Language Variation and Change, 2*(2), 205-254.

Lan, Y. (2017). Foreign Influence and Sound Change: A Case Study of Cantonese Alveolar Affricates. In R.E. Roxas (Ed.), *Proceedings of the 31st Pacific Asia Conference on Language, Information and Computation* (pp. 155-160). The National University (Philippines).

Lee, W., & Zee, E. (2010). Articulatory characteristics of the coronal stop, affricate, and fricative in Cantonese. *Journal of Chinese Linguistics, 38*(2), 336-372.

REFERENCES

Liu, Q. (2010). A Research on the Affricate and the Fricative Consonants of Hong Kong Cantonese - A Statistical Analysis of Cantonese Songs in the Past Thirty Years. *Journal of Jinan University (Philosophy and Social Sciences), 32*(6), 123-129.

Mielke, Jeff. 2015. An Ultrasound Study of Canadian French Rhotic Vowels with Polar Smoothing Spline Comparisons. *The Journal of the Acoustical Society of America 137* (5): 2858–2869.

Scobbie, J. M., Lawson, E., Cowen, S., Cleland, J., and Wrench, A. A. (2011). A common co-ordinate system for mid-sagittal articulatory measurement. QMU CASL Working Papers WP-20, available at http://eresearch.qmu.ac.uk/3597/.

van Rij, J., Wieling M., Baayen R., van Rijn H. (2020). itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs. R package version 2.4

Wood S (2017). *Generalized Additive Models: An Introduction with R*, 2nd edition. Boca Raton, FL: Chapman and Hall/CRC.

Yu, A. C. (2016). Vowel-dependent variation in Cantonese /s/ from an individual-difference perspective. *The Journal of the Acoustical Society of America, 139*(4), 1672-1690.