Revealing covert articulation in s-retraction

Stephen Nichols & George Bailey
University of Manchester

Annual Meeting of the LAGB
University of Sheffield
12 September 2018

INTRODUCTION

- We investigate the realisation of the sibilant in the word-initial clusters **/stɹ/** and **/stj/**, which is often more **[ʃ]**-like, using both acoustic and articulatory data
- We address the following questions:
    - Categoricity v. gradience in *s*-retraction, i.e. is the surface realisation of **/s/** in **/stɹ/** and **/stj/** identical to an underlying **/ʃ/**?
        - not just with respect to acoustics but also articulation
    - What degree of inter-speaker variation do we find? To what extent do we find different "systems" of *s*-retraction?
    - What happens in **/stj/** (e.g. *stupid*) and how comparable is it to **/stɹ/** (e.g. *street*)?
    - What does this suggest about the mechanisms that trigger this process?

Background

- Attested in various varieties of English (see e.g. Shapiro 1995, Lawrence 2000, Durian 2007, Bass 2009, Sollgan 2013, Phillips 2016, Wilbanks 2016, 2017, Wilson 2018)
- Focus has often been sociolinguistic rather than phonetic aspect
  - But see Stevens & Harrington (2016) for work on the phonetic origins
- Well-studied with **/stɹ/** in AmE but relatively under-studied in BrE
- BrE also has **/stj/**, which is absent in AmE (at least in these contexts)
- Has been characterised as **retraction**, based primarily on acoustic data
  - Notable exceptions being ultrasound studies by Mielke et al. (2010) and Baker et al. (2011)
- However, acoustics doesn't necessarily have a one-to-one mapping with articulation
  - See e.g. Mielke et al. (2016) on covert articulation of **/ɹ/**

Phonetic motivations

- The rôle of **/ɹ/** has been foregrounded in many studies:
  - Shapiro (1995) claims *s*-retraction is triggered non-locally by **/ɹ/**
  - Baker et al. (2011) find that even "non-retractors" show coarticulatory bias towards retraction in clusters containing **/ɹ/**, e.g. **/spɹ/**
- However, some have argued that **/ɹ/**'s influence may be more indirect:
  - Lawrence (2000) claims that this is local assimilation with **/ɹ/** causing affrication of **/t/** to **/tʃ/** leading to the retraction of **/s/**
  - This could be particularly appropriate for BrE where **/t/** undergoes a similar process before **/j/** for most speakers
    - e.g. *tune* /tjʉːn/ > [tʃʉːn]   *stupid* /stjʉːpɪd/ > [ʃtʃʉːpɪd]?
  - But Magloughlin & Wilbanks (2016) suggest otherwise for Raleigh English

METHODOLOGY

Design of Stimuli

- 9 word-initial contexts embedded in the carrier sentence **'I know [...] is a word'**

Baselines for comparison:
underlying /s, ʃ/

/s/
e.g. *seep*

/ʃ/
e.g. *sheep*

Retracting environments:

/stɹ/
e.g. *street*

/stj/
e.g. *stupid*

\+ /st/
e.g. *steep* ?

Pseudo distractors:

/tʃ/
e.g. *cheap*

/ɹ/
e.g. *read*

/tɹ/
e.g. *treat*

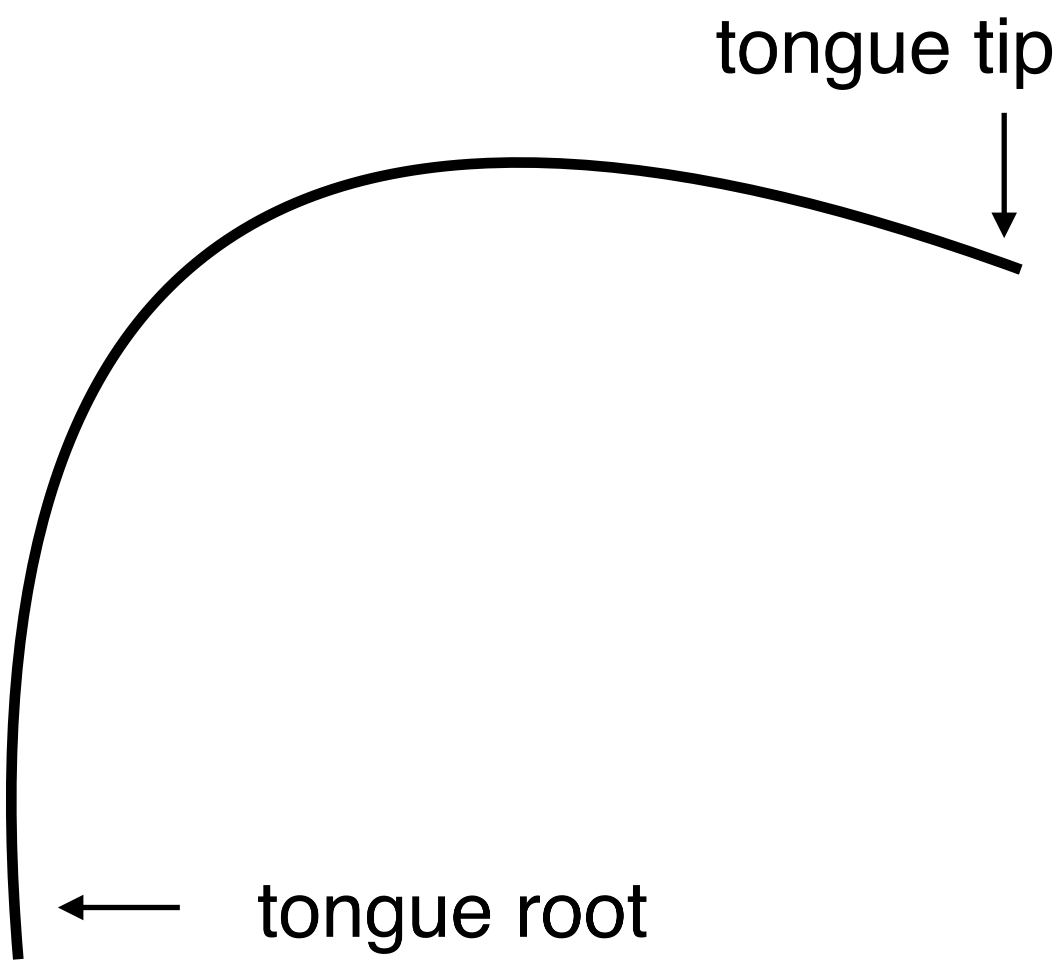
/tj/
e.g. *tune*

Useful for independent evidence of what happens to /tɹ/ and /tj/ outside of post-/s/ environments

- All contexts precede **[iː]**, **[ʉː]** and **[ɒ]** (except **/stj/**, which only occurs before **[ʉː]**)
- 5 repetitions per token giving a total of 130 sentences per speaker

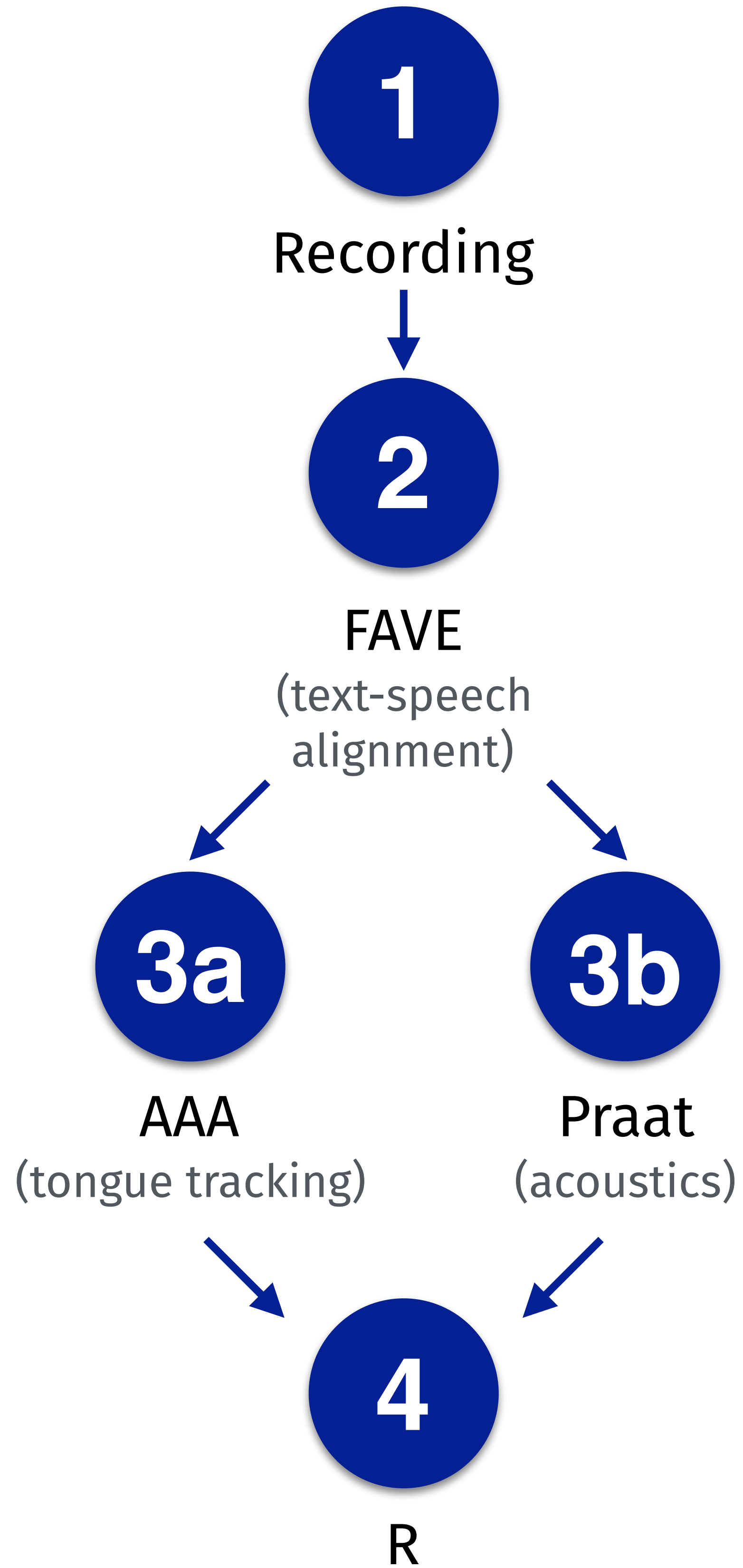
DATA COLLECTION

- Synchronised UTI (60fps) and audio recording (lavalier mic)
- Mid-sagittal view
- Stabilised with headcage
- Currently 8 speakers (3M; 5F) aged 18-26
    - All born (or at least raised from age 4) in Greater Manchester, but in some cases parents aren't from Manchester (or even England)



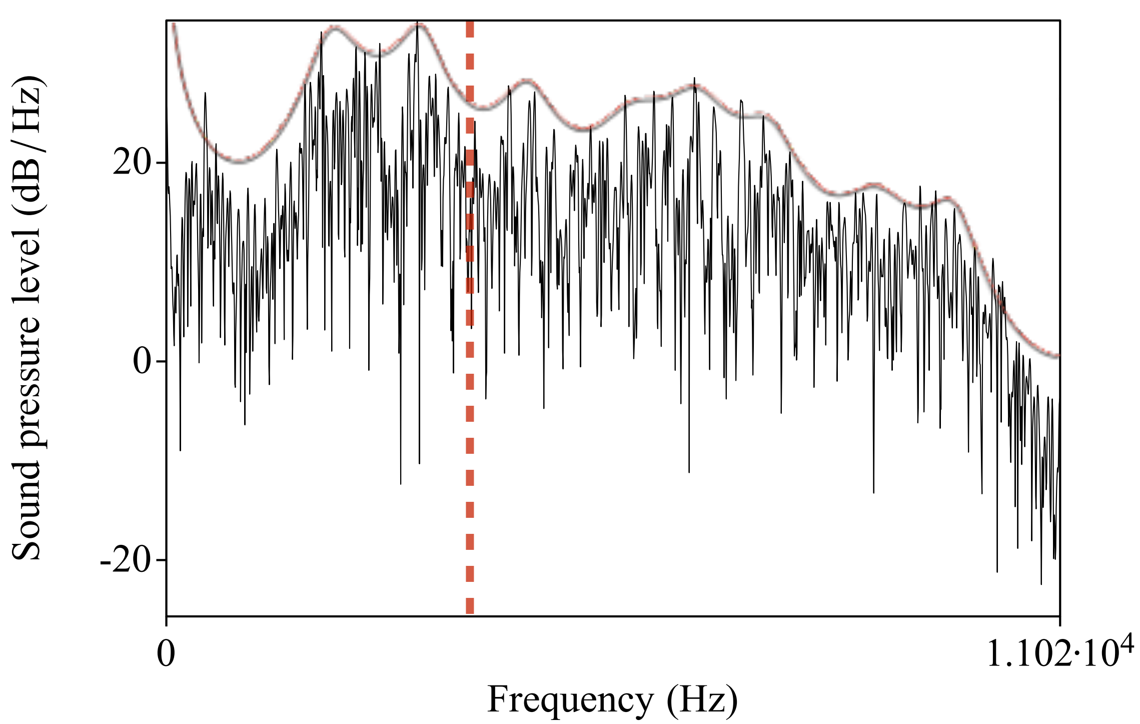
DATA ANALYSIS

- Forced-alignment using FAVE (Rosenfelder et al. 2011)
  - Manually-corrected, with further sub-segmentation
    - e.g. *tree* T R IY1 > T CH R IY1
- Tongue splines tracked and exported using AAA (Articulate Instruments Ltd. 2011)
  - 3 keyframes per segment - analysis conducted on keyframe 2 (segment mid-point)
  - Data read into R with `rticulate` (Coretta 2017) package

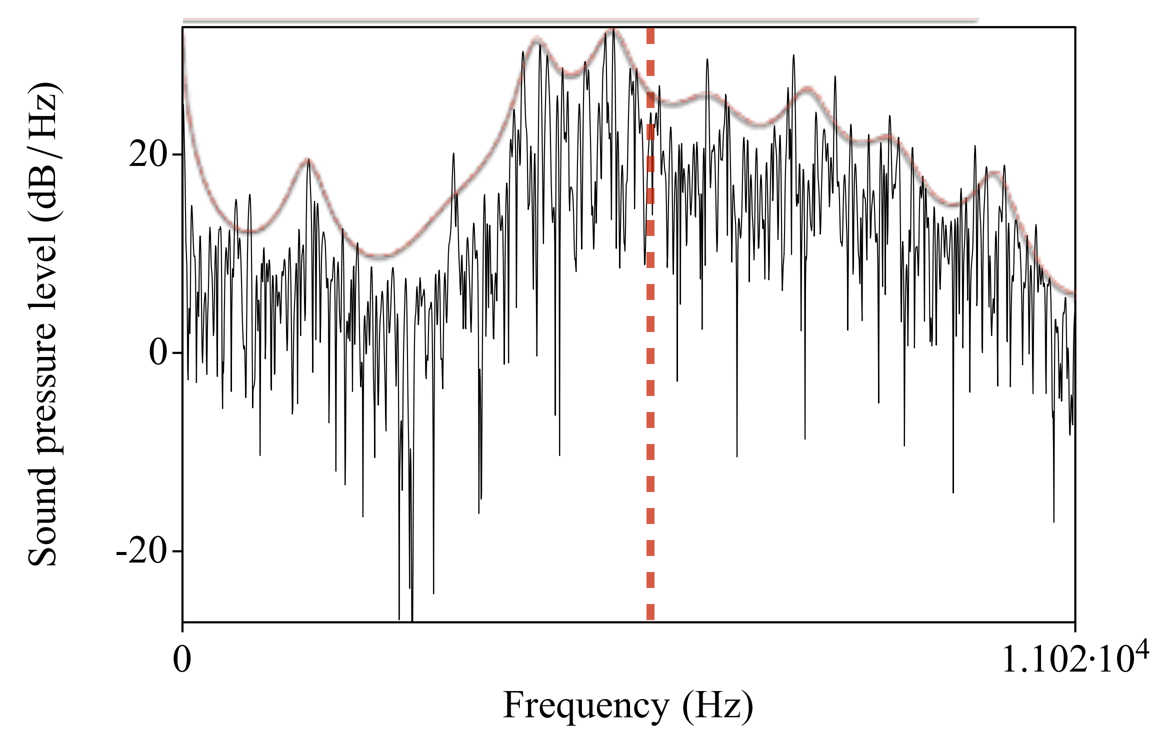


DATA ANALYSIS

- To complement ultrasound data, acoustic analysis was performed in Praat using two scripts adapted from DiCanio (2017)
- For each fricative (and affricate), we extract:
  - **Centre of gravity (CoG)**
    - lower value = more /ʃ/-like; higher value = more /s/-like (Jongman et al. 2000, Baker et al. 2011)
  - **LPC-smoothed spectral slice**
    - 10 peaks



/ʃ/ CoG: 3749 Hz



/s/ CoG: 5743 Hz

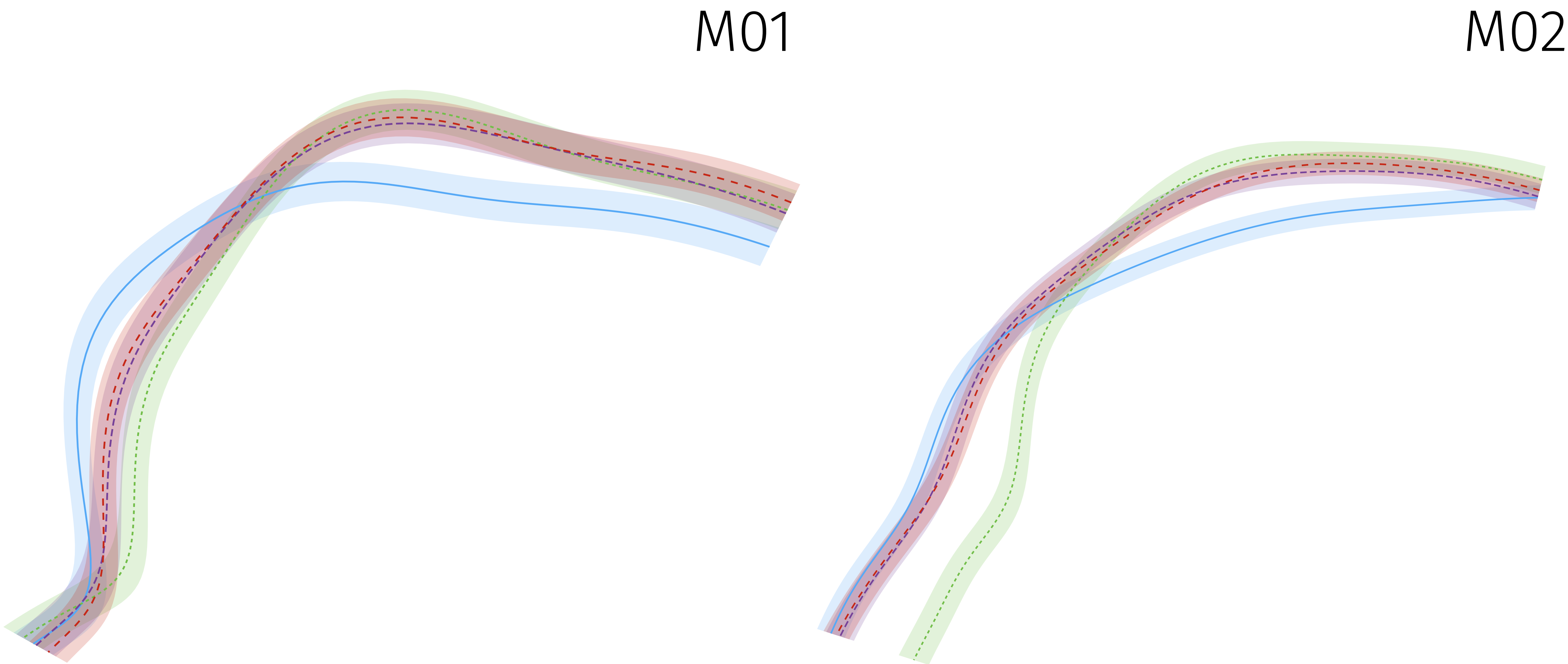
Statistical methods

- **Ultrasound**
  - Modelled with *GAMMs* (*generalised additive mixed models*) using `tidymv` and `rticulate` packages (Coretta 2017, 2018)
  - Ideal for modelling non-linear effects in dynamic (time/space) data (see Sóskuthy 2017 and references therein)
- **Acoustics**
  - *Mixed-effects linear regression* for CoG measures with `lme4` package (Bates et al. 2015)
  - Supplemented with *functional principle components analysis* for LPC-smoothed spectral slices using `fda` package (Ramsay et al. 2013)
    - see Appendix

RESULTS

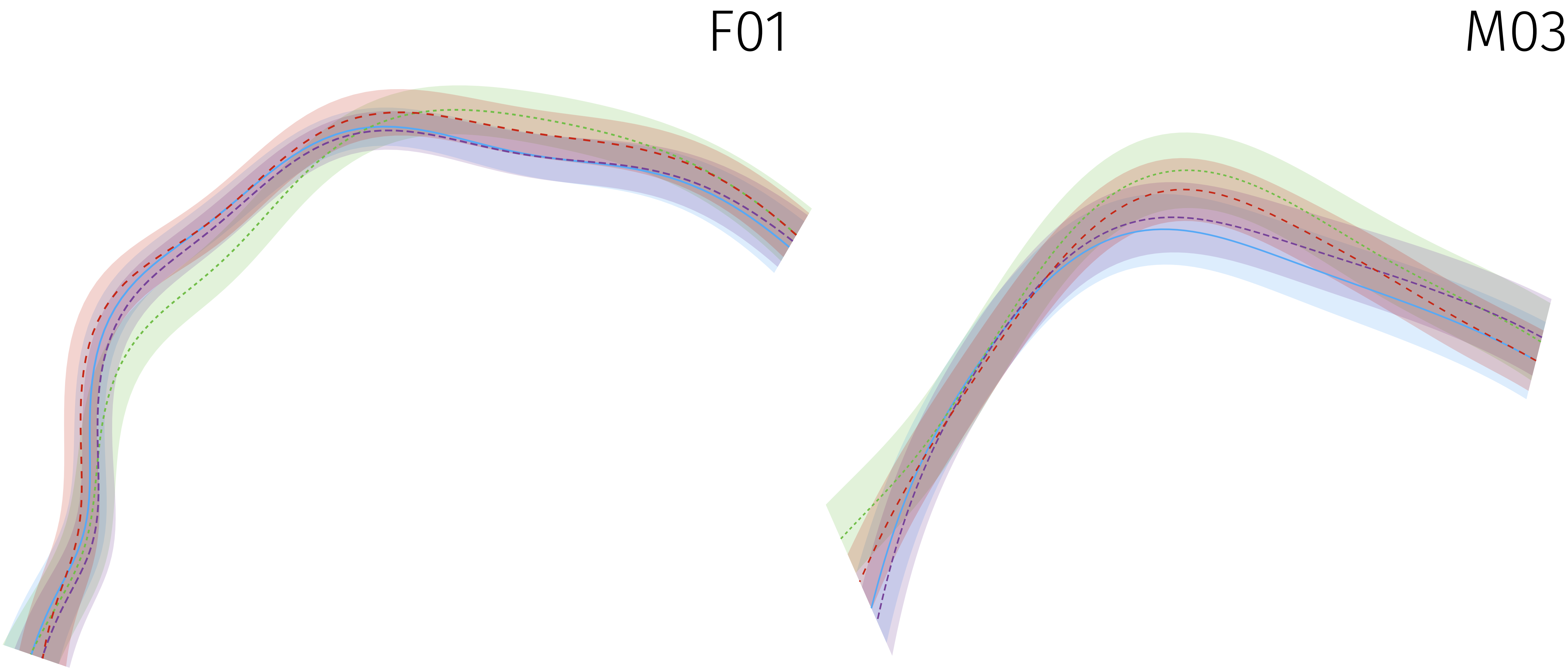
ARTICULATION

Articulation



- Clear bimodality for tongue body: /ʃ/-/stɹ/-/stj/ v. /s/

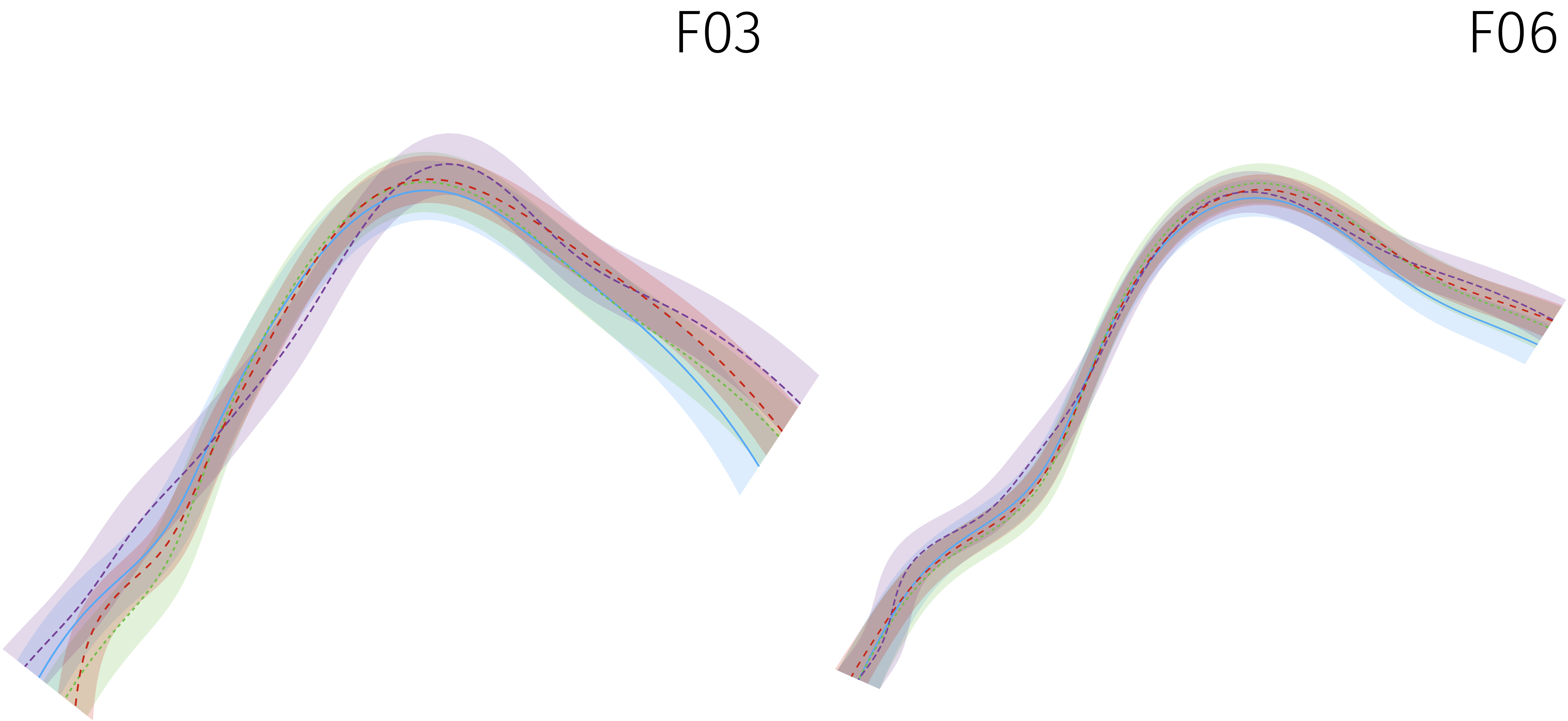
Articulation



- Tongue body for **/stj/** largely overlapping with **/ʃ/**
- But **/stɹ/** much more similar to **/s/** than **/ʃ/**

ARTICULATION

(also F07 and F08)



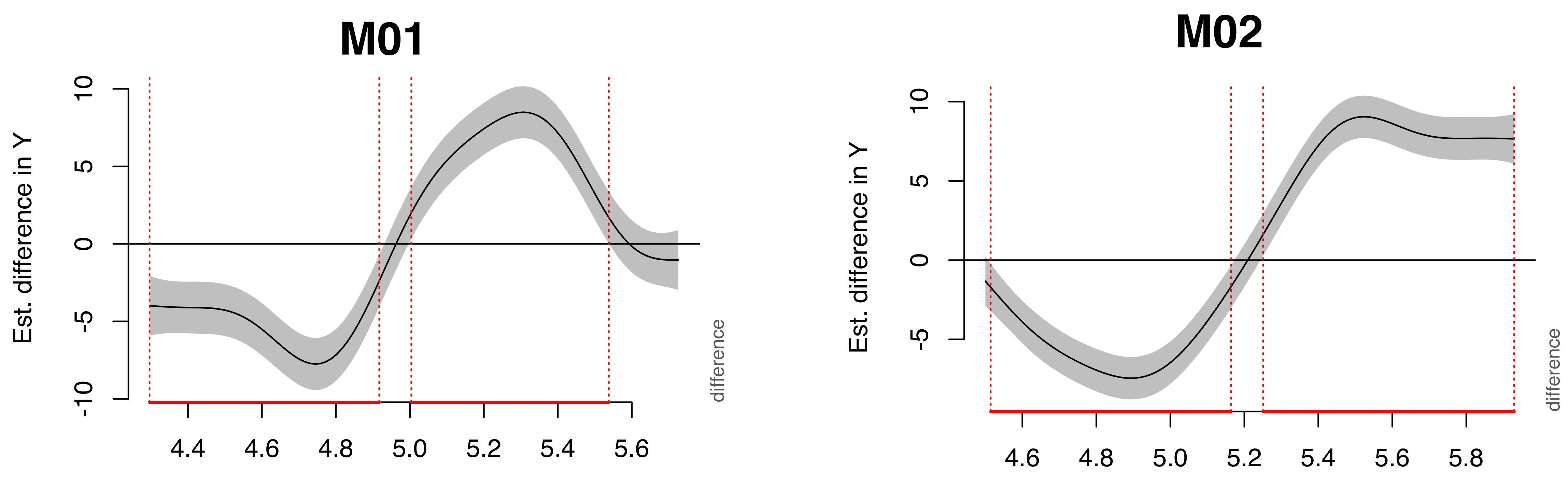
- Almost complete overlap between all four contexts, even /s/ and /ʃ/
- More differentiation at tongue tip (but confidence intervals also wider)

Interim summary: articulation

- Some speakers exhibit clear tongue body retraction, such that there are two groups:
  - /s/ v. /ʃ/-/stɹ/-/stj/
- Others show a more intermediate pattern where /stj/ is closer to /ʃ/ but /stɹ/ is more similar to /s/
- Finally, other speakers have no apparent lingual difference, even between /s/ and /ʃ/

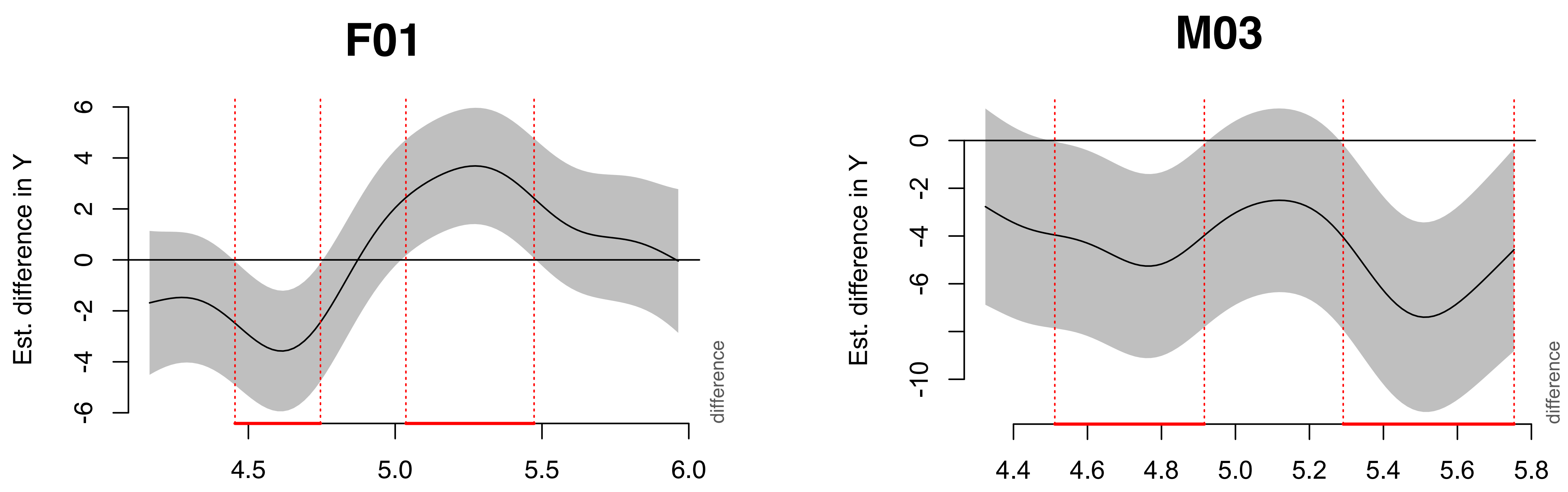
Difference smooths

- In addition to visual inspection of the splines, difference smooths can be used for pairwise comparisons of tongue shapes
  - Differences between the two curves are highlighted in red (where confidence interval of difference smooth does not contain 0)
  - More red = more differentiation in tongue shape
  - /s/ and /ʃ/ completely different for M01 and M02



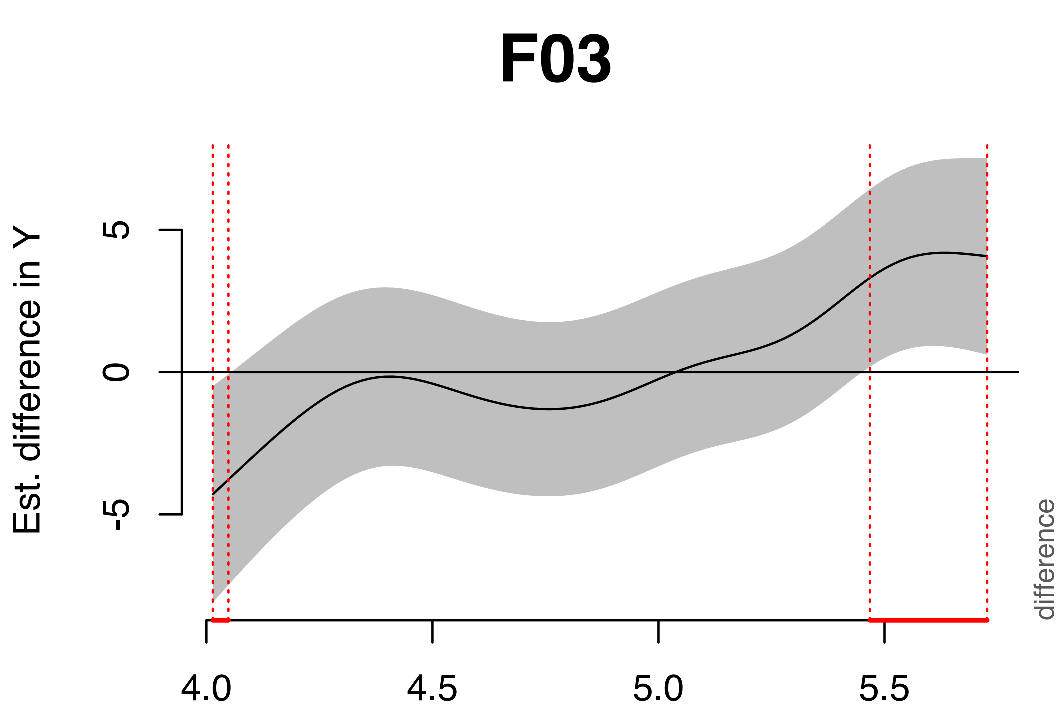
DIFFERENCE SMOOTHS

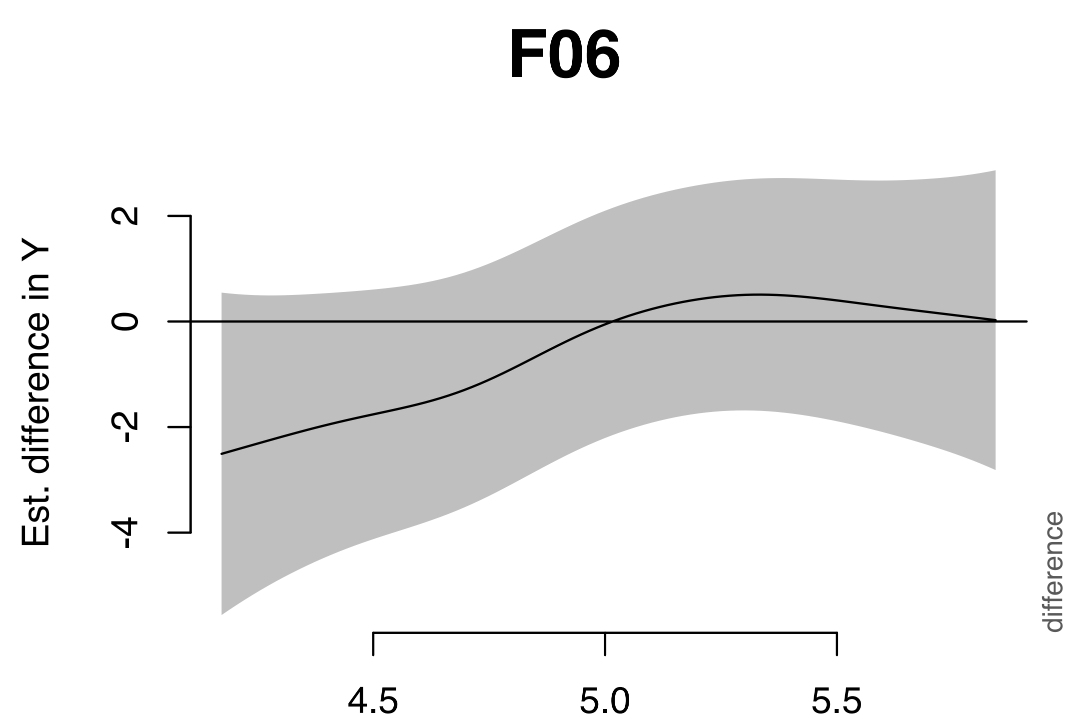
- In addition to visual inspection of the splines, difference smooths can be used for pairwise comparisons of tongue shapes
    - Differences between the two curves are highlighted in red (where confidence interval of difference smooth does not contain 0)
    - More red = more differentiation in tongue shape
    - /s/ and /ʃ/ largely distinct (but to a lesser extent) for F01 and M03



DIFFERENCE SMOOTHS

- In addition to visual inspection of the splines, difference smooths can be used for pairwise comparisons of tongue shapes
  - Differences between the two curves are highlighted in red (where confidence interval of difference smooth does not contain 0)
  - More red = more differentiation in tongue shape
  - /s/ and /ʃ/ not at all different for F03 and F06 (as well as F07 and F08)

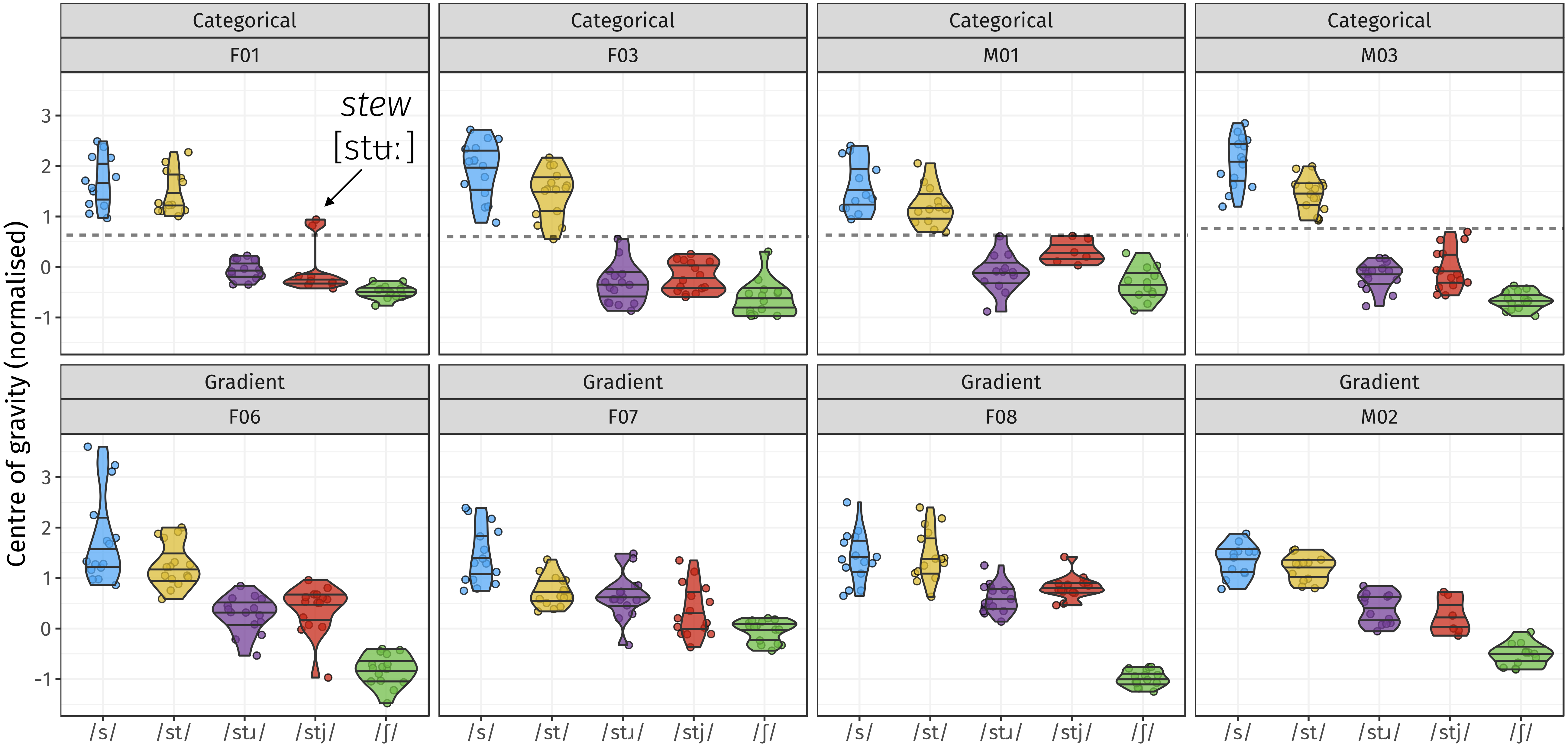




RESULTS

ACOUSTICS

Centre of gravity



- **All** speakers maintain an acoustic contrast between /s/ and /ʃ/
- Categoricity/gradience determined by Tukey contrasts for post-hoc pairwise significance tests in linear regression models (i.e. whether or not **/stɹ/** or **/stj/** are significantly different from **/ʃ/**)

Centre of gravity

- The acoustic analysis reveals that:

  1. **All** speakers do have an acoustic contrast between **/s/** and **/ʃ/**

  2. **All** speakers exhibit some degree of acoustic "retraction" in **/stɹ/** and **/stj/**

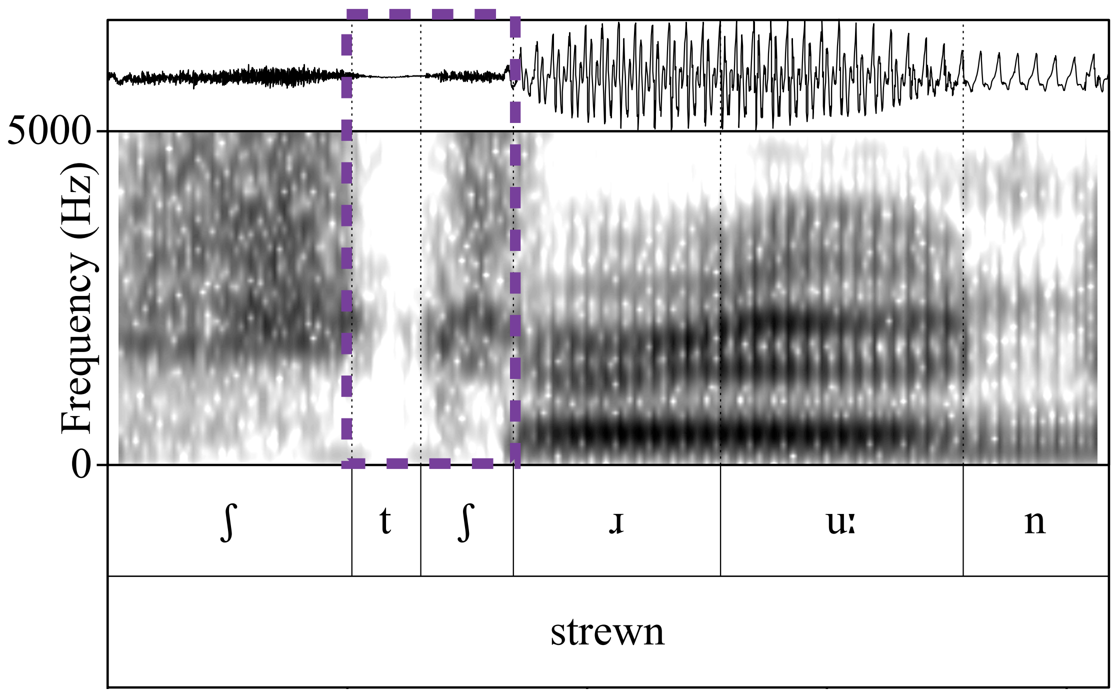
- This may be categorical for some and gradient for others but crucially:

  - Speakers are either categorical in both or gradient in both - there is no evidence that for a single speaker retraction is more advanced in one than the other

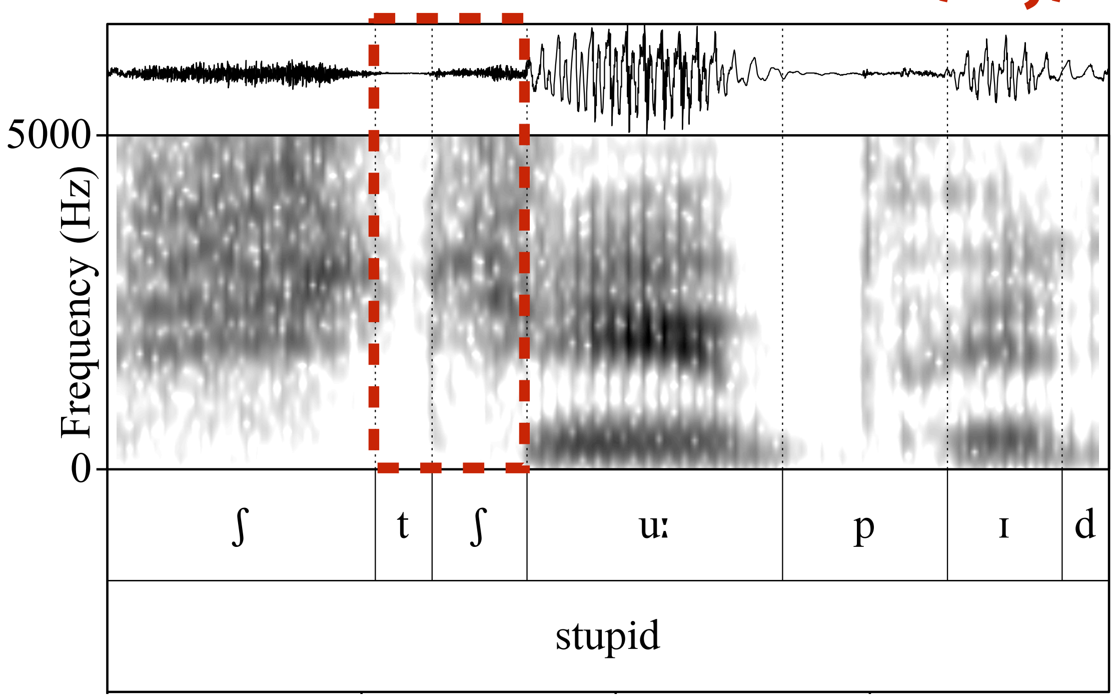
  - Suggests that retraction in both environments is governed by the same underlying process, or at least the same phonetic motivations

AFFRICATION?

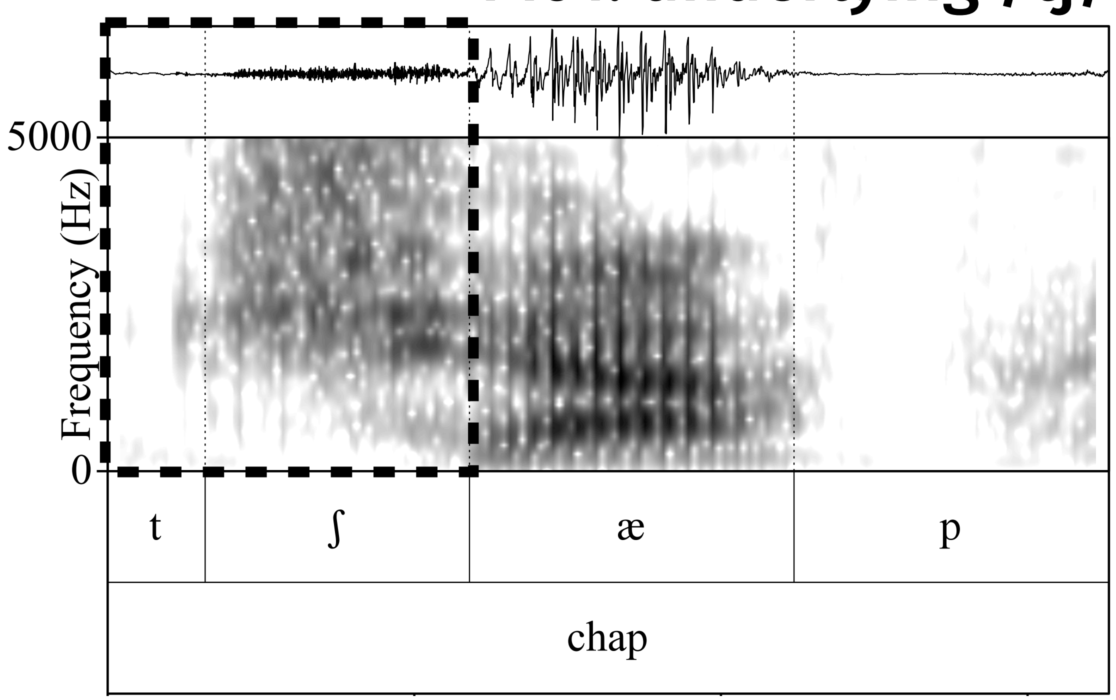
M01: /stɹ/



M01: /stj/



M01: underlying /tʃ/



- Comparable affrication of **/t/** across both **/stɹ/** and **/stj/** environments
- Phonetically similar to underlying **/tʃ/** (just shorter in duration)
- **Some** speakers do differentiate the affricated **/t/** (w.r.t. CoG) depending on whether it is followed by **/j/** or **/ɹ/** (see Appendix)

AFFRICATION?

- Crucially, **all** speakers affricate **/t/** - it's only the spectral properties of the fricated portion that are variable
- Some evidence that a speaker can affricate **/t/** with only minimal retraction of **/s/** (e.g. F08)
    - But no evidence that speakers retract **/s/** without affricating **/t/**
        - e.g. *[ʃtɹiːt], *[ʃtjʉːpɪd]

DISCUSSION

THE ARTICULATION-ACOUSTICS MAPPING

Covert articulation

- Even though some speakers show no apparent articulatory difference even between underlying /s/ and /ʃ/, the acoustic contrast is still maintained
- Rutter (2011) highlights the three phonetic parameters that define the /s/-/ʃ/ contrast:
  - **TONGUE PLACEMENT** - alveolar for /s/, post-alveolar for /ʃ/
  - **TONGUE SHAPE** - grooved for /s/, slit/flat for /ʃ/
  - **LIP SHAPE** - slight labialisation for /s/, strong labialisation for /ʃ/

'It is also worth noting that changes in one of the phonetic parameters discussed above may not necessarily co-occur with changes in the other two' (Rutter 2011:31)

  - **TONGUE TIP** - laminal vs. apical constriction
- Speakers achieving the same acoustic output through different articulatory means?
  - Similar to covert articulation in **/ɹ/** (Delattre & Freeman 1968, Mielke et al. 2016)

The articulation–acoustics mapping

	articulation (UTI)		acoustics (CoG)
M01	categorical	↔	categorical
M02	categorical	↔	gradient
M03	gradient	↔	categorical
F01	gradient	↔	categorical
F03	none	↔	categorical
F06	none	↔	gradient
F07	none	↔	gradient
F08	none	↔	gradient

The articulation-acoustics mapping

- No one-to-one mapping between articulation (ultrasound) and acoustics (CoG)
- We find all but one of the six possible mappings (using these categories)
  - With a larger sample size we would likely find examples of this

- **categorical ↔ categorical**
  - M01
- **categorical ↔ gradient**
  - M02
- **gradient ↔ categorical**
  - F01, M03
- **none ↔ categorical**
  - F06, F07, F08
- **none ↔ gradient**
  - F03
- **gradient ↔ gradient**
  - ...

CONCLUSIONS

Conclusions

- The /stɹ/ and /stj/ contexts behave similarly in terms of acoustic *s*-retraction and *t*-affrication
- This lends support to the idea that retraction is triggered by affrication and not by **/ɹ/** directly
- Evidence that the articulatory mechanisms behind the /s/-/ʃ/ contrast are more complicated than a simple retraction of the place of articulation
  - highlights the need for a more nuanced approach to the articulation of "retraction"
  - and calls into question the suitability of "retraction" as a label for this phenomenon:
    - *s*-hushing? (i.e. hissing /s/ > hushing /ʃ/)
- Speakers could be hitting an acoustic target rather than articulatory target (Boersma 2011:§4)
- Lends support to the older idea that distinctive features should be defined primarily in acoustic terms (Jakobson et al. 1952, Durand 1990:§2.5)
- Highlights importance of (ideally simultaneous) articulatory **and** acoustic studies
- Although, in this case, even capturing midsagittal ultrasound does not tell the whole story

FUTURE WORK

- **Further avenues for articulatory exploration:**
  - Look more closely at the tongue shape of **/ɹ/** with midsagittal UTI
  - Video recording for lip-rounding (rather than using F3-F2 as a proxy)
  - Electropalatography (EPG), electromagnetic articulography (EMA) and parasagittal UTI to investigate the other articulatory mechanisms of sibilant production, e.g. tongue tip, grooving/slitting

- **Acoustic work to be done:**
  - Investigate word-internal retraction and the effect of morpheme boundaries, e.g. *posture, registry* etc.
  - Investigate phrase-level retraction, e.g. *pass treats*, and the effect of prosodic boundaries and speech rate
  - Collect **/ʃɹ/** data (e.g. *shriek*, *shrew*, *shrapnel*) to compare with **/stɹ/**
  - Look at pre-**[p]** and pre-**[k]** environments, e.g. *spoon, spring; school, screw*
  - Perform acoustic analysis on conversational data (existing corpus of 32 sociolinguistic interviews from Manchester and other North West cities)

Acknowledgements

Thanks to **Stefano Coretta** for help with ultrasound; **Patrycja Strycharczuk** and **Ricardo Bermúdez-Otero** for their feedback; **Michele Gubian** for help with FPCA; and **Jane Scanlon** for agreeing to be our first victim while we tried fitting the headcage.

🌎 http://personalpages.manchester.ac.uk/staff/stephen.nichols/
✉️ stephen.nichols@manchester.ac.uk

🌎 http://personalpages.manchester.ac.uk/staff/george.bailey/
✉️ george.bailey@manchester.ac.uk
🐦 @grbails

References

Articulate Instruments Ltd. 2011. Articulate Assistant Advanced. Version 2.17.02. URL: `http://www.articulateinstruments.com/aaa/`.

Baker, Adam, Diana Archangeli & Jeff Mielke. 2011. Variability in American English *s*-retraction suggests a solution to the actuation problem. *Language Variation and Change* 23(3). 347-74.

Bass, Michael. 2009. Street or shtreet? Investigating (str-) palatalisation in Colchester English. *Estro: Essex Student Research Online* 1(1). 10-21.

Boersma, Paul. 2011. A programme for bidirectional phonology and phonetics and their acquisition and evolution. In Anton Benz & Jason Mattausch (eds.), *Bidirectional Optimality Theory*, 33-72. Amsterdam: John Benjamins.

Coretta, Stefano. 2017. `rticulate`: Ultrasound Tongue Imaging in R. R package version 1.3.1. URL: `https://github.com/stefanocoretta/rticulate`.

Coretta, Stefano. 2018. `tidymv`: Tidy Model Visualisation. R package version 1.3.1. URL: `https://github.com/stefanocoretta/tidymv`.

Delattre, Pierre & Donald C. Freeman. 1968. A dialect study of American R's by X-ray motion picture. *Linguistics* 44. 29-68.

DiCanio, Christian. 2017. Time averaging for fricatives. Praat script. Haskins Laboratories & SUNY Buffalo. URL: `https://www.acsu.buffalo.edu/~cdicanio/scripts/Time_averaging_for_fricatives_2.0.praat`.

Durand, Jacques. 1990. Generative and non-linear phonology. London: Longman.

Durian, David. 2007. Getting [ʃ]tronger every day?: More on urbanization and the socio-geographic diffusion of (str) in Columbus, OH. *University of Pennsylvania Working Papers in Linguistics* 13(2). 65-79.

Gylfadottir, Duna. 2015. Shtreets of Philadelphia: An acoustic study of /str/-retraction in a naturalistic speech corpus. *University of Pennsylvania Working Papers in Linguistics* 21(2). 89-97.

Haley, Katarina L., Elizabeth Seelinger, Kerry Callahan Mandulak & David J. Zajac. 2010. Evaluating the spectral distinction between sibilant fricatives through a speaker-centered approach. *Journal of Phonetics* 38(4). 548-54.

Jakobson, Roman, Gunnar Fant & Morris Halle. 1952. *Preliminaries to Speech Analysis*. Cambridge, MA: MIT Press.

Jongman, Allard, Ratree Wayland & Serena Wong. 2000. Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America* 108(3). 1252-63.

Lawrence, Wayne P. 2000. /str/ → /ʃtr/: Assimilation at a distance? *American Speech* 75. 82-7.

Magloughlin, Lyra & Eric Wilbanks. 2016. An Apparent Time Study of (str) Retraction and /tɹ/ - /dɹ/ Affrication in Raleigh, NC English. Presentation given at New Ways of Analyzing Variation 45, Vancouver, BC, Canada, 3-6 November.

Mielke, Jeff, Adam Baker & Diana Archangeli. 2010. Variability and homogeneity in American English /ɹ/ allophony and /s/ retraction. In Barbara Kühnert (ed.), Variation, detail, and representation. *LabPhon* 10, 699–729. Berlin: Mouton de Gruyter.

Mielke, Jeff, Adam Baker & Diana Archangeli. 2016. Individual-level contact limits phonological complexity: Evidence from bunched and retroflex /ɹ/. *Language* 92(1). 101-40.

Phillips, Jacob B. 2016. Phonological and prosodic conditioning of /s/-retraction in American English. Presentation given at the 15th Conference on Laboratory Phonology, Ithaca, NY, United States, 3-6 November.

Ramsay, J. O., Hadley Wickham, Spencer Graves & Giles Hooker. 2013. `fda`: Functional Data Analysis. R package version 2.4.0. URL: `https://CRAN.R-project.org/package=fda`.

Rosenfelder, Ingrid, Josef Fruehwald, Keelan Evanini & Jiahong Yuan. 2011. FAVE (Forced Alignment and Vowel Extraction) program suite. URL: `http://fave.ling.upenn.edu`.

Rutter, Ben. 2011. Acoustic analysis of a sound change in progress: The consonant cluster /stɹ/ in English. *Journal of the International Phonetic Association* 41(1). 27-40.

Shapiro, Michael. 1995. A case of distant assimilation: /str/ → /ʃtr/. *American Speech* 70. 101-7.

Sollgan, Laura. 2013. STR-palatalisation in Edinburgh accent: A sociophonetic study of a sound change in progress. MSc dissertation, University of Edinburgh.

Sóskuthy, Márton. 2017. Generalised additive mixed models for dynamic analysis in linguistics: a practical introduction. ArXiv preprint: `https://arxiv.org/abs/1703.05339`.

Stevens, Mary & Jonathan Harrington. 2016. The phonetic origins of s-retraction: Acoustic and perceptual evidence from Australian English. *Journal of Phonetics* 58. 118-34.

Wilbanks, Eric. 2016. (str) Retraction in Raleigh: "Identical" variants implicated in Two Separate Sound Changes. Presentation given at Penn Linguistics Conference 40, Philadelphia, PA, United States, 18-20 March.

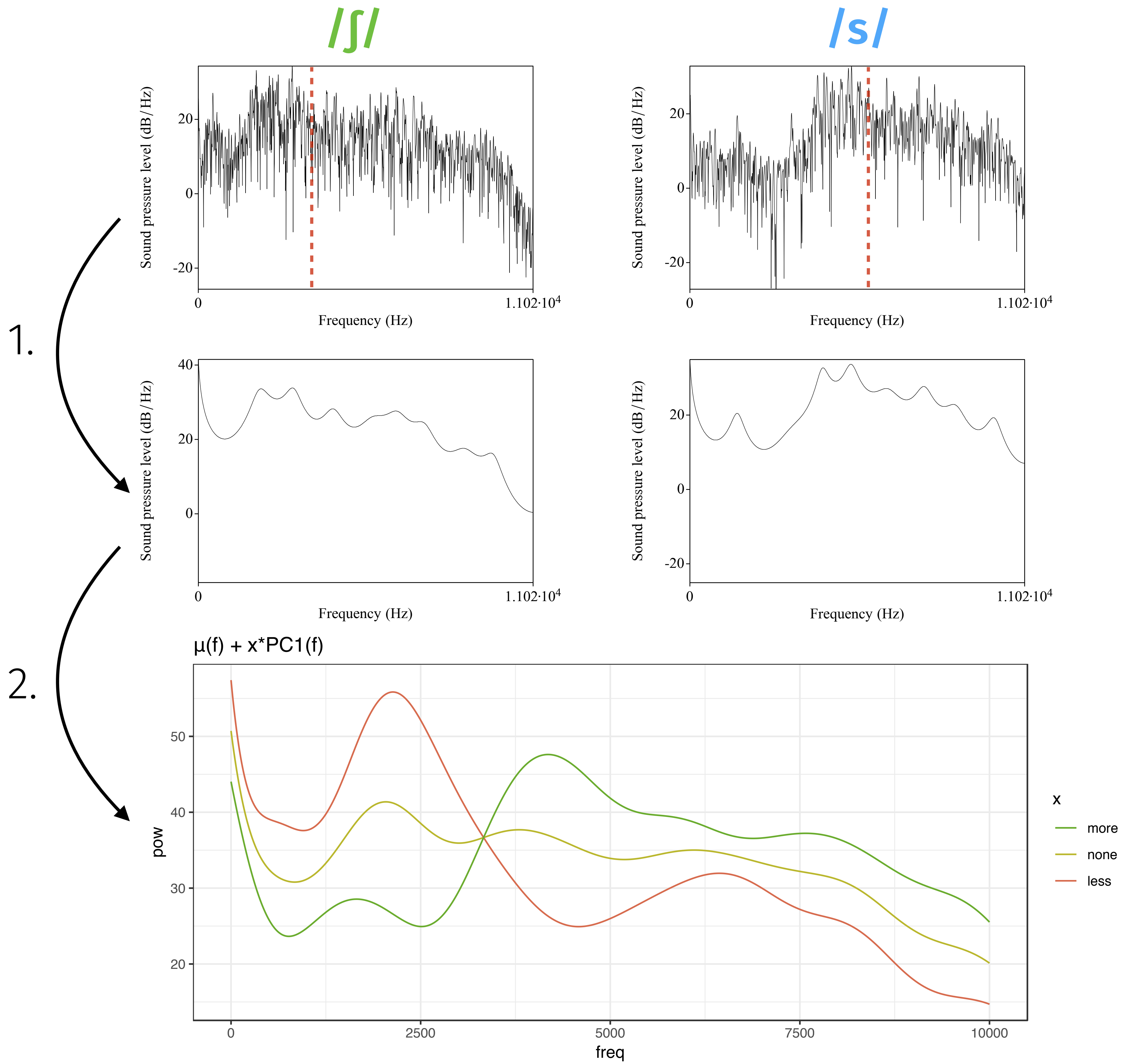
Wilbanks, Eric. 2017. Social and structural constraints on a phonetically-motivated change in progress: (str) retraction in Raleigh, NC. U*niversity of Pennsylvania Working Papers in Linguistics* 23(1). 301-10.

Wilson, Sophie. 2018. A midsagittal ultrasound tongue imaging study to investigate the degree of /s/-retraction in /stɹ/ onset clusters in British English. BA dissertation, Newcastle University.

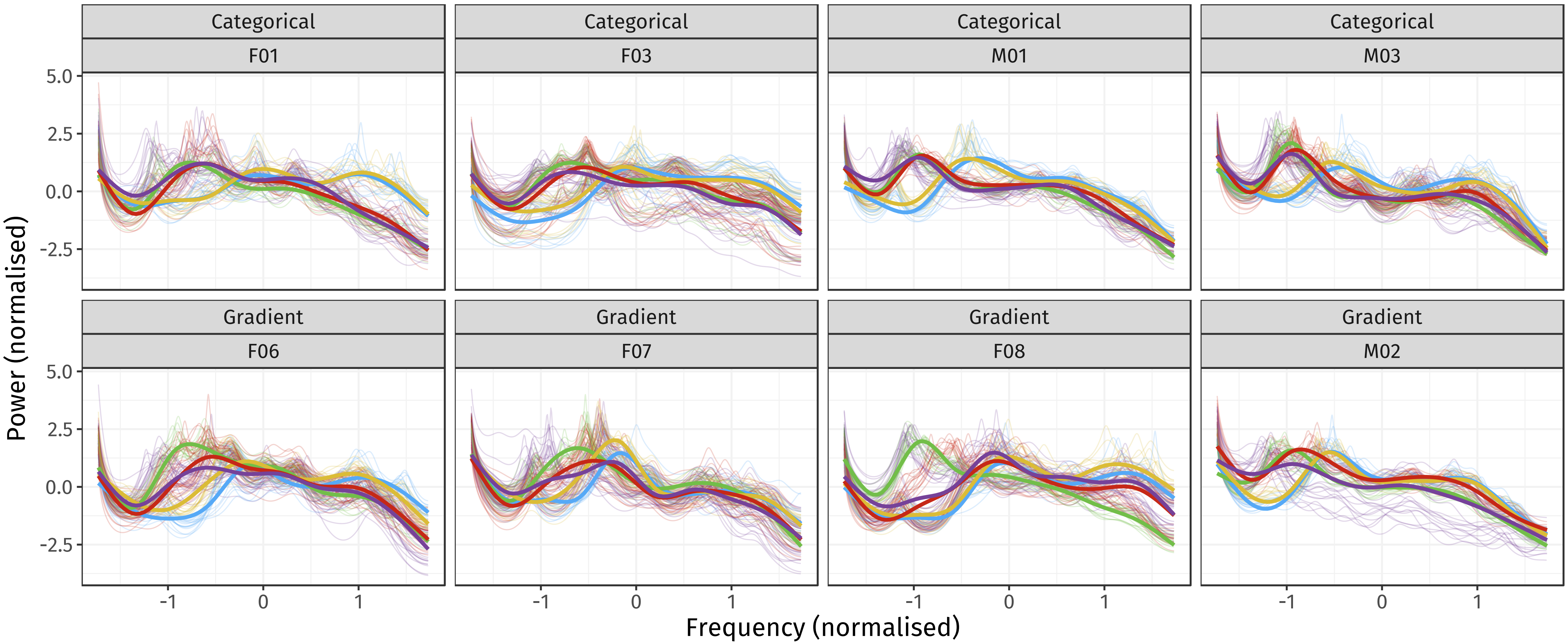
APPENDICES

Functional principle components analysis (FPCA)

- Single spectral moments (e.g. CoG, skew, kurtosis) often used to distinguish sibilants (Haley et al. 2010:548-9)
- But this is an over-simplification of a complex acoustic signal
- We also analyse the entire curve:
  1. LPC smoothing of spectral slice
  2. Use FPCA to reduce dimensionality and describe curve shapes using two or three principle components (PCs)

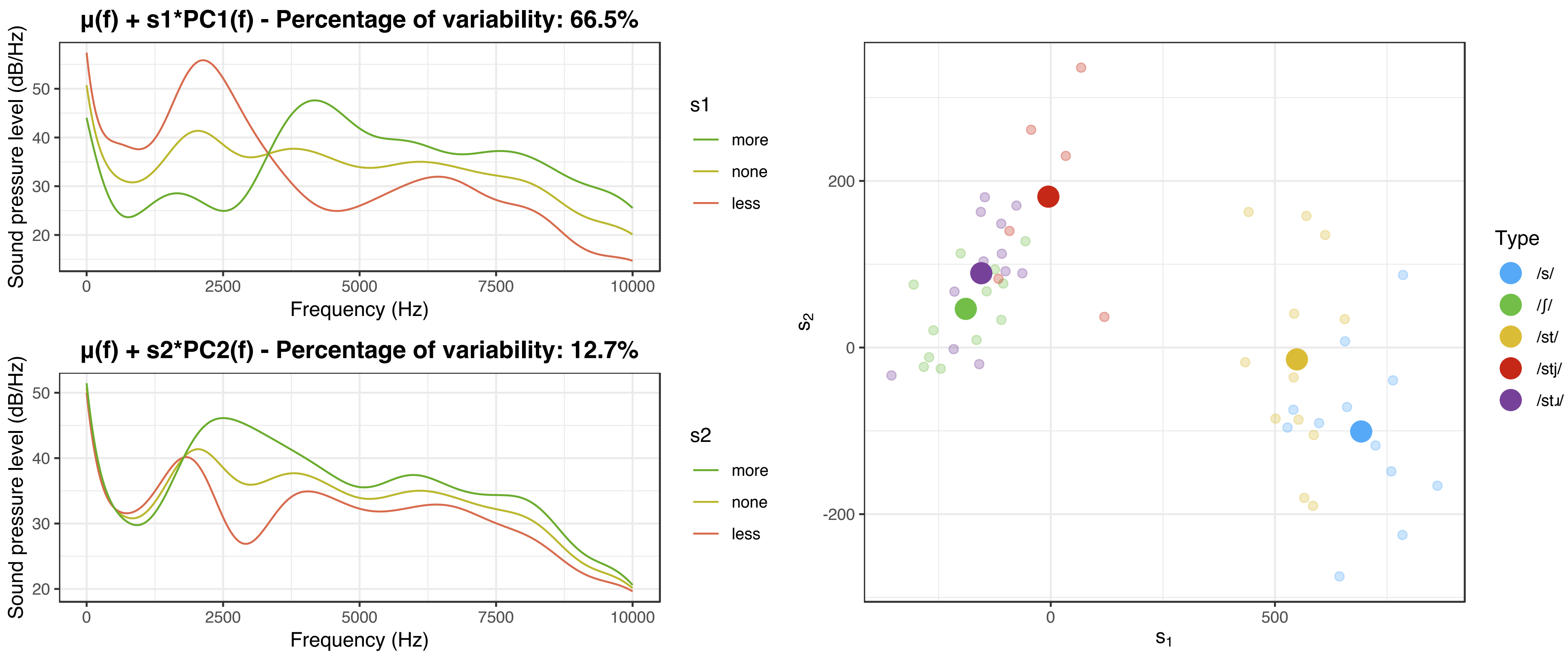


LPC-SMOOTHED SPECTRAL SLICES



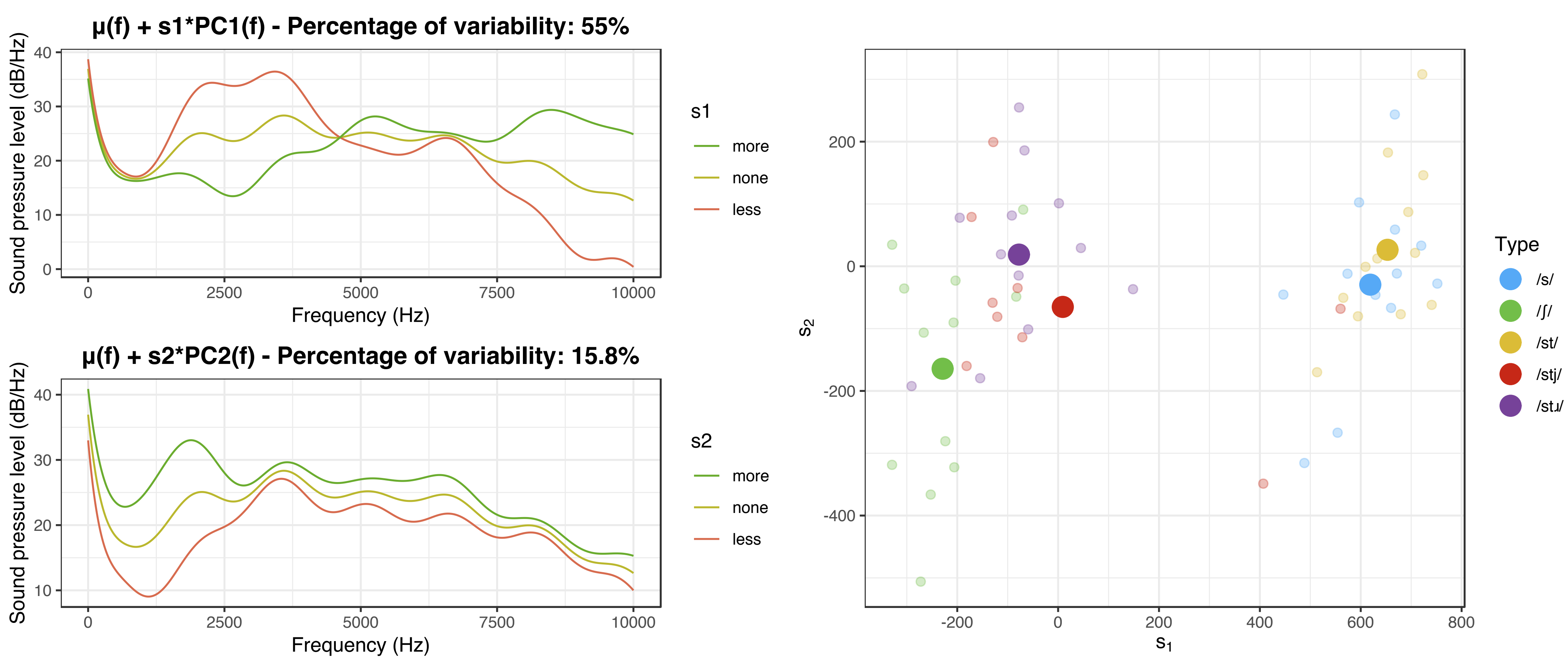
- Looking at the entire spectral profile, the same two patterns emerge as with CoG:
  - “Categorical” speakers, where /stɹ/ and /stj/ patterns with /ʃ/
  - “Gradient” speakers, where /stɹ/ and /stj/ are intermediate between /s/ and /ʃ/

Functional principle components analysis (FPCA)



M01

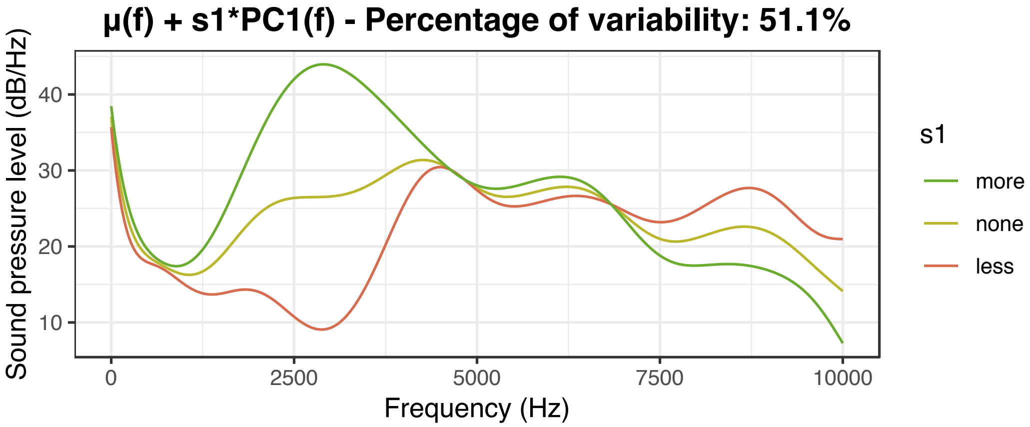
categorical

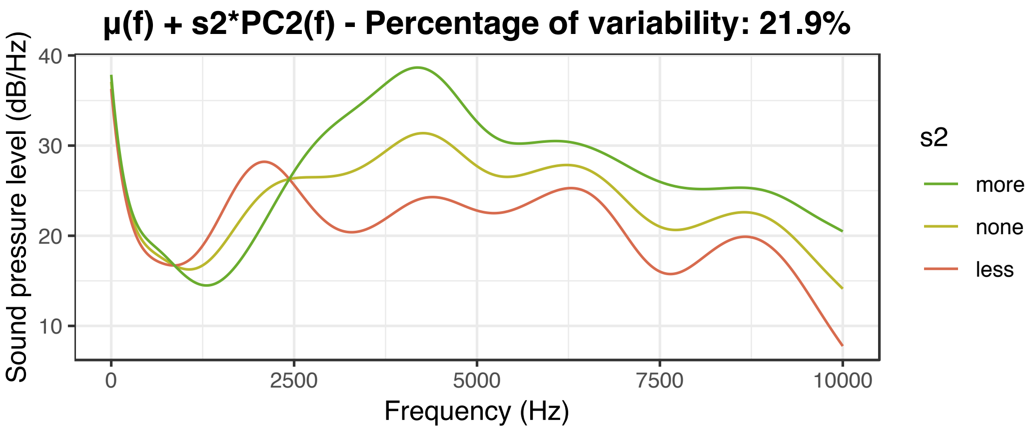


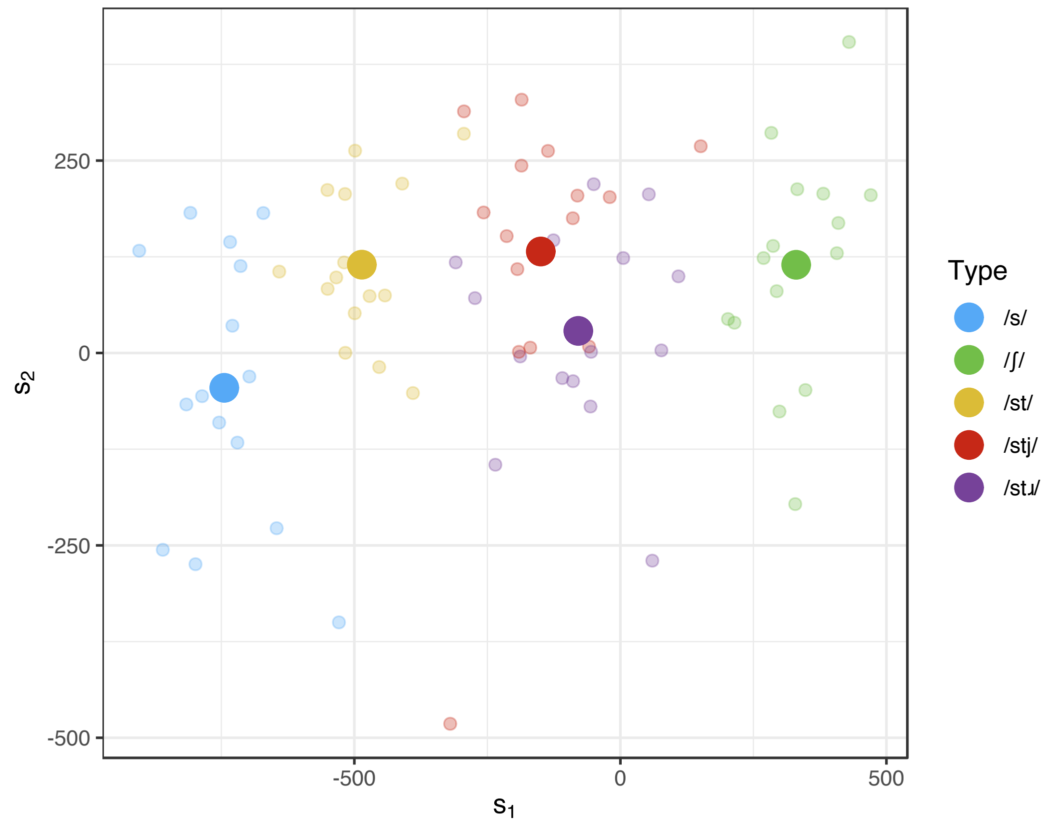
F01

categorical

Functional principle components analysis (FPCA)

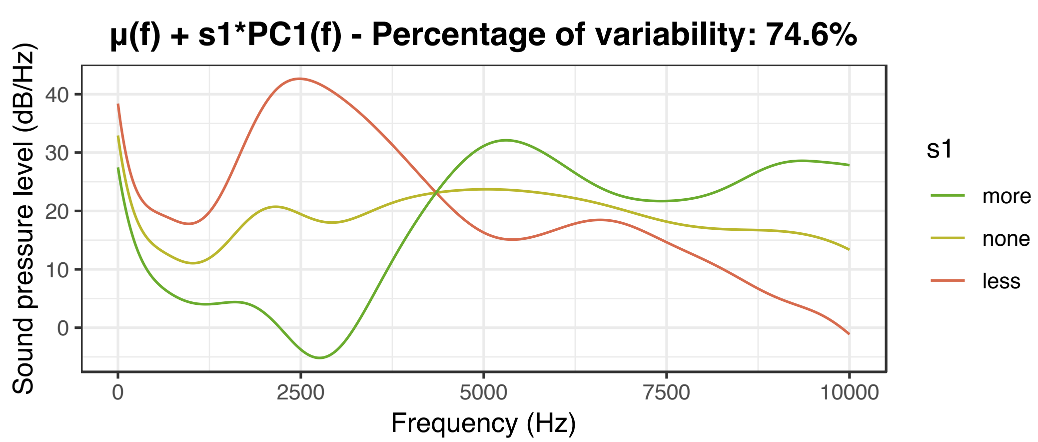


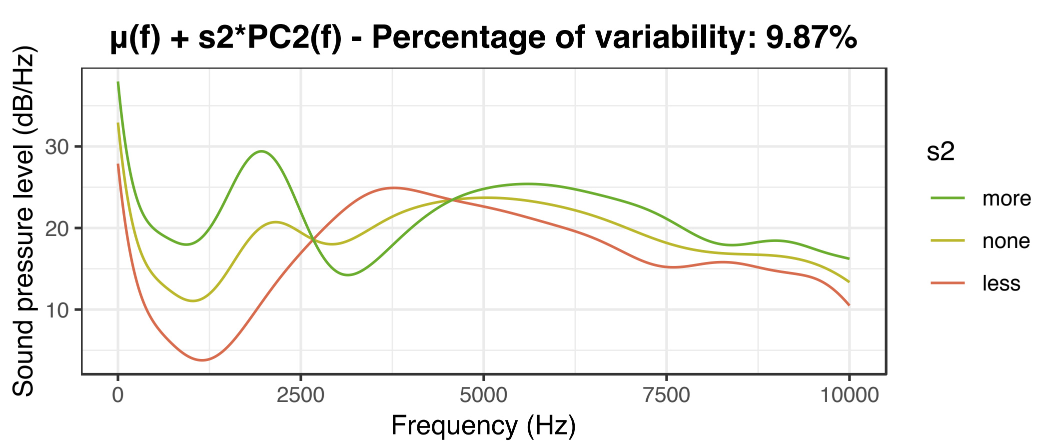


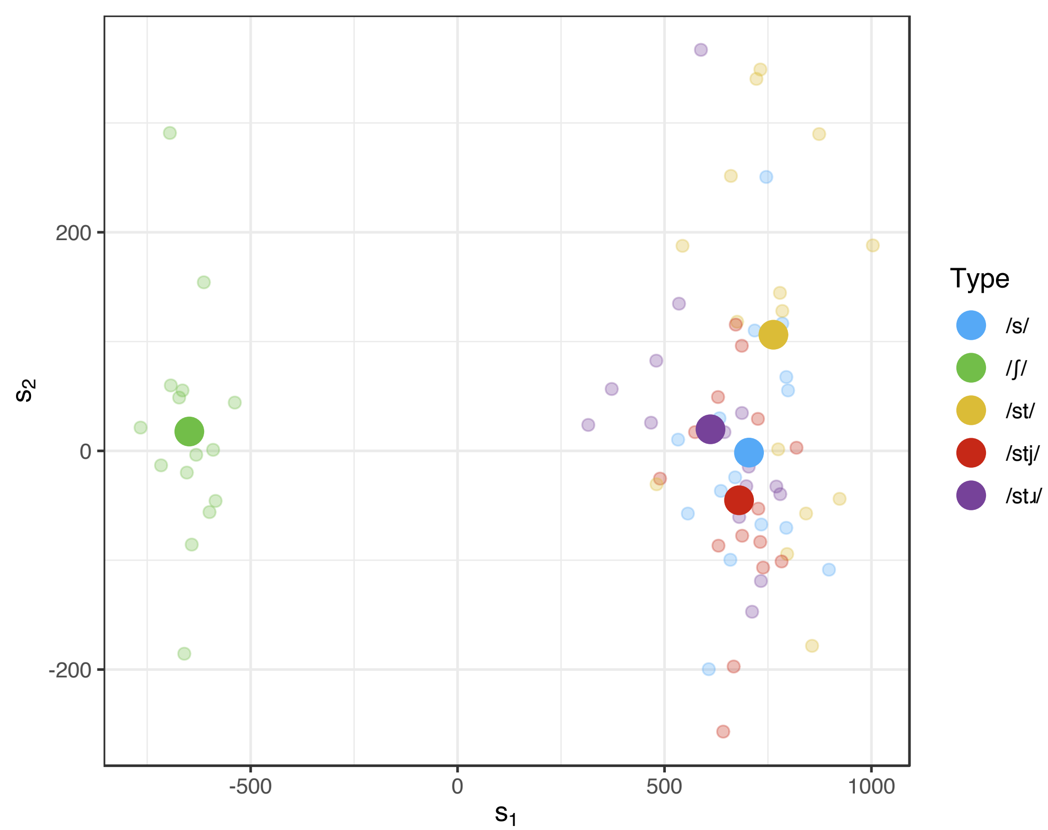


F06

gradient



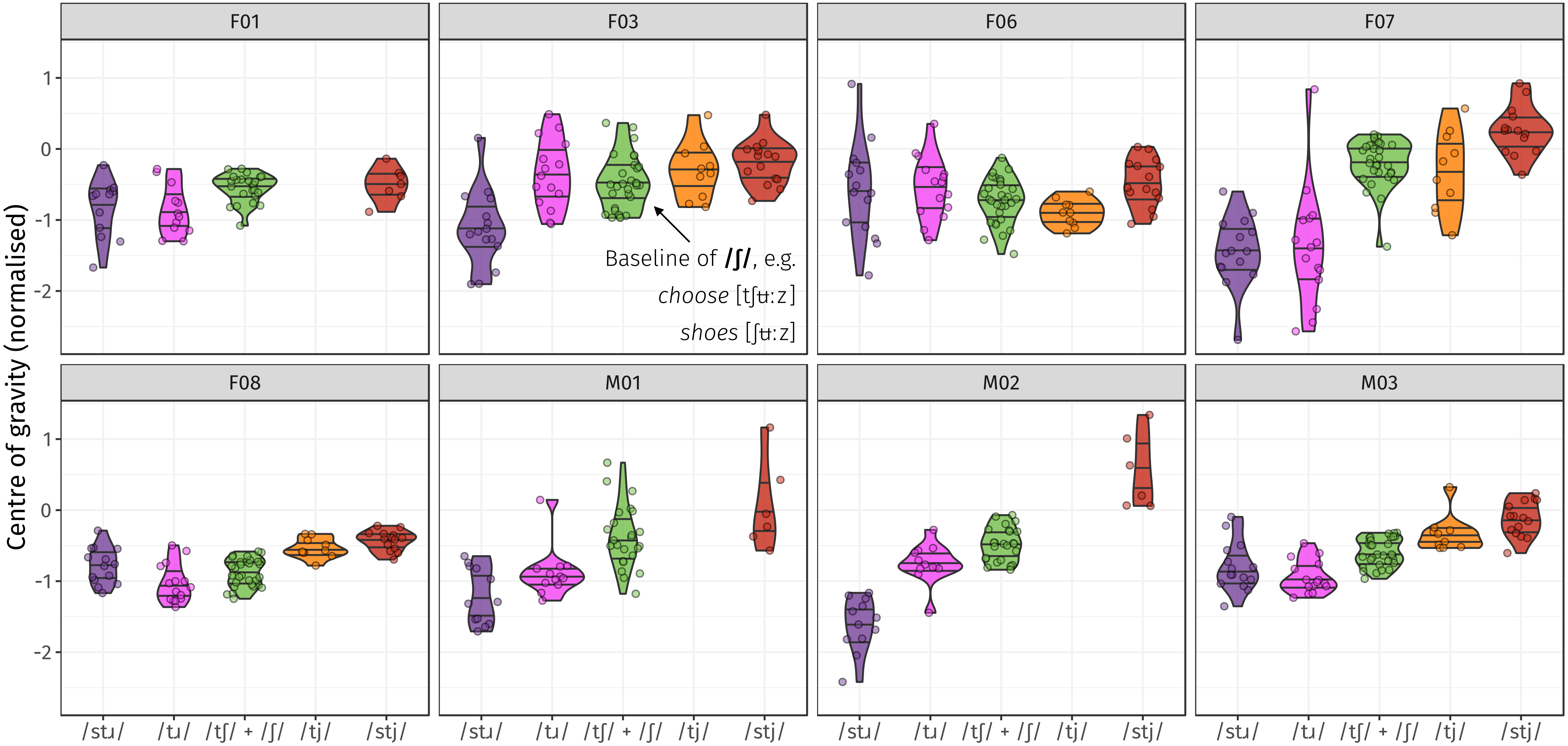




F08

gradient (none)

AFFRICATION?



- For most speakers, the fricated portions of pre-**/ɹ/** affrication and **/tj/**-coalescence are identical both to each other and to underlying **/tʃ/**
- But **some** speakers do differentiate the affricated **/t/** depending on whether it is followed by **/j/** or **/ɹ/** (see F07, M01, M02)