Assessing phonological control of parasagittal tongue shape in Japanese sibilants

Michael C. Stern[a], Jason A. Shaw[a], Shigeto Kawahara[b]

[a]Yale University

[b]Keio University

“Phonological control”

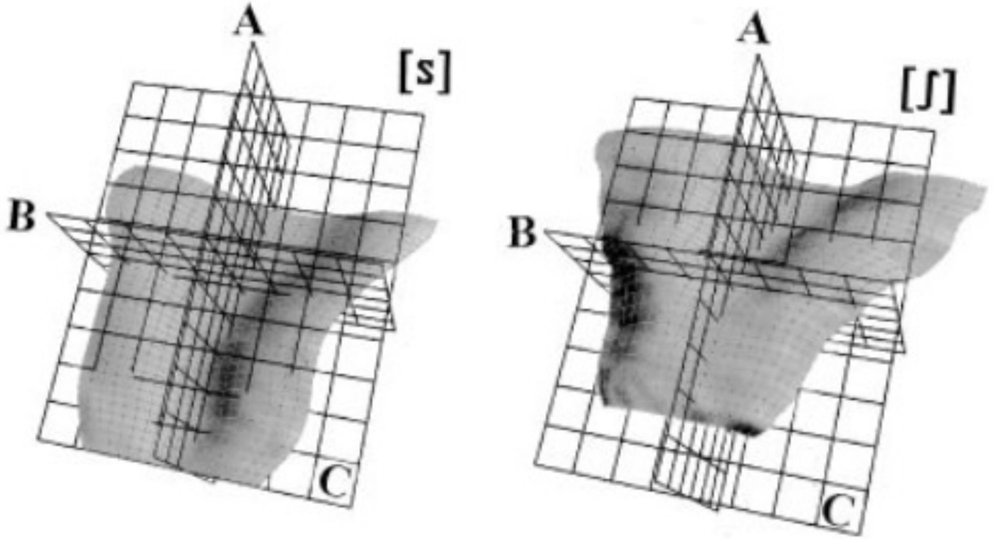
- Abstract phonological primitives (e.g., features or gestures) correspond in some way to physical dimensions, whether articulatory, auditory/acoustic, or a combination thereof
- Other dimensions—those not under phonological control—may passively covary with controlled dimensions
- For example: f_0 during a vowel can be raised as a passive consequence of actively closing the vocal folds for [ʔ] (e.g., Hombert et al., 1979)

This talk

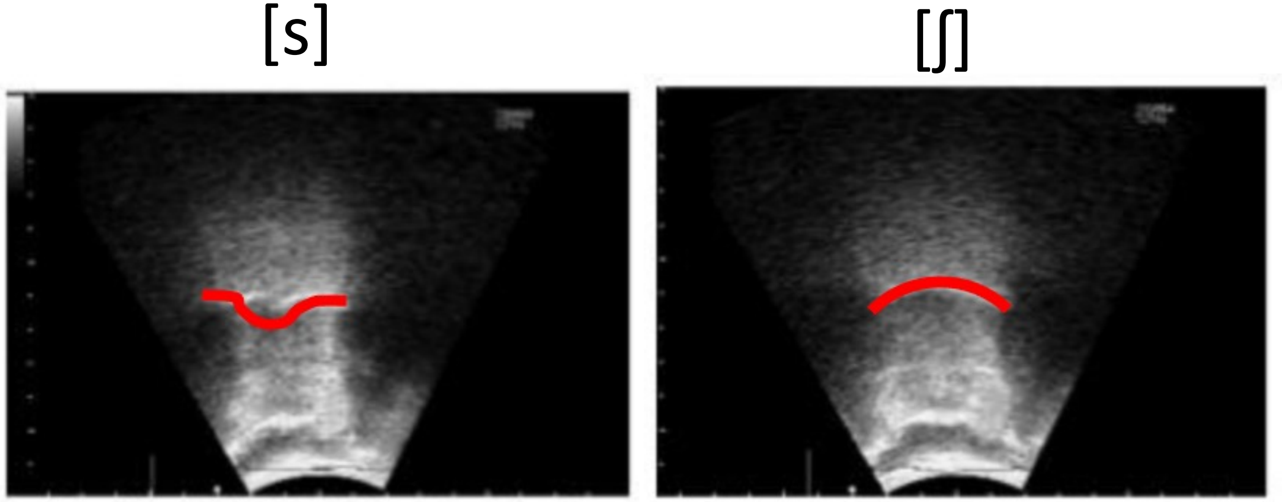
- Which dimensions are under phonological control during Japanese sibilant production?

- Japanese has two sibilants which are similar to English sibilants
    - **Japanese**: anterior [s] vs. posterior [ɕ]; **English**: anterior [s] vs. posterior [ʃ]
    - Precise phonetic difference between Japanese [ɕ] and English [ʃ] is somewhat unclear
    - There are also by-language differences in phonological patterning and acquisition error patterns

English [s] and [ʃ] differ in parasagittal tongue shape: Ultrasound data



Stone & Lundberg (1996)



Whalen et al. (2011)

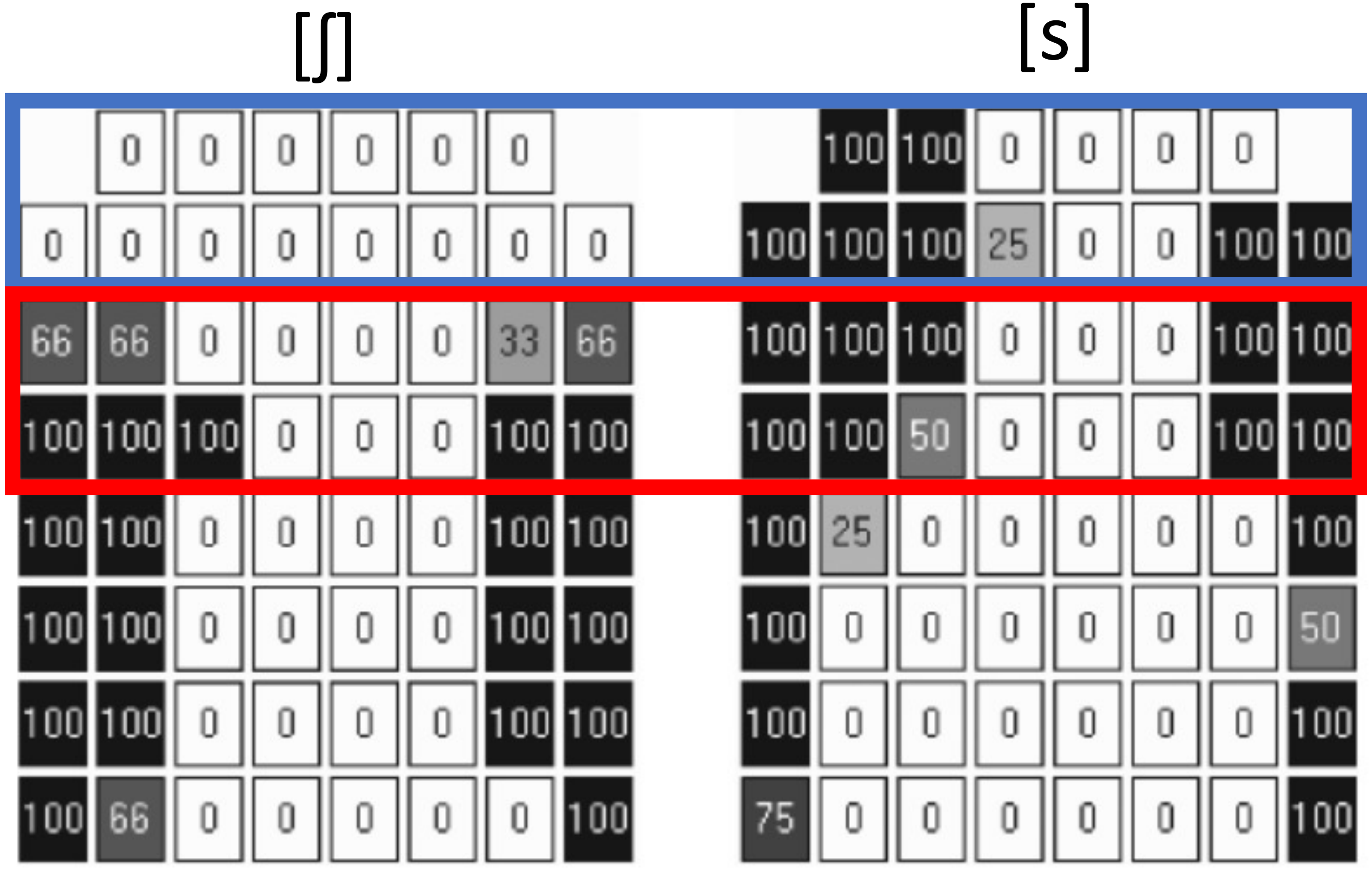
Deeper and narrower groove for [s] than [ʃ]

English sibilants: Electropalatography (EPG)

More **alveolar contact** for [s] (right) than [ʃ] (left)

Also more **post-alveolar contact** for [s] than [ʃ]

Consistent with deeper and narrower groove for [s] than [ʃ]



Pouplier et al. (2011)

Japanese sibilants: EPG data

Like English, more **alveolar contact** for [s] than [ɕ]

Unlike English, *less* **post-alveolar contact** for [s] than [ɕ]

By-language difference in parasagittal control?

[ɕ]

	0	0	0	0	0	0	
78	61	0	0	0	0	0	26
100	100	26	0	0	0	59	76
100	100	100	29	0	0	58	100
100	100	100	39	0	0	100	100
100	100	88	19	0	0	100	100
100	100	76	13	0	0	39	100
100	59	0	0	0	0	11	100

[s]

	39	22	0	0	0	47	
100	100	69	0	0	0	23	80
100	61	26	0	0	0	0	100
100	52	0	0	0	0	0	100
100	19	0	0	0	0	0	100
100	19	0	0	0	0	0	100
100	13	0	0	0	0	0	93
100	93	0	0	0	0	0	31

Matsui (2017)

Differences in phonological patterning

- The English sibilant contrast is less susceptible to influence from surrounding vowels:
- English sibilants contrast before all vowels (but neutralized before certain clusters, e.g. [strit] ~ [ʃtrit])
- The Japanese sibilant contrast is more limited
    - Complementary distribution in native words (Yamato lexical stratum): [ɕ] occurs before [i], and [s] occurs before all other vowels (/si/ → [ɕi])
    - In Sino-Japanese and recent loans, [s] and [ɕ] contrast before non-front vowels ([ɕa, ɕu, ɕo]), but rarely before [i]
- Perhaps this is due to a difference in phonological control

Differences in acquisition error patterns

Li et al. (2009):

- English-learning children tend to replace /ʃ/ with [s]
- Japanese-learning children tend to replace /s/ with [ɕ]

Perhaps English-learning and Japanese-learning children learn different dimensions of phonological control

Hypothesis

- These phonetic, phonological, and acquisition facts follow from a by-language difference in phonological control:

- English sibilant production involves active parasagittal control
  - [s] = deep, narrow groove
  - [ʃ] = wider groove or doming

- Japanese sibilant production does not
  - [s] ~ [ɕ] contrast is maintained by midsagittal constriction location

This study

Investigate parasagittal control during Japanese sibilant production using 3D Electromagnetic Articulography (EMA)

Participants: Three adult native Japanese speakers

- S01: Male, 30s, Tokyo
- S02: Male, 30s, Osaka
- S03: Female, 30s, Tokyo

Materials & Procedure

24 real Japanese words beginning with either [s] or [ɕ] followed by either [u] or [i]

Carrier phrase*: okee* ___ *to itte* ‘okay say ___ again’

Stimuli were presented on screen in Japanese orthography

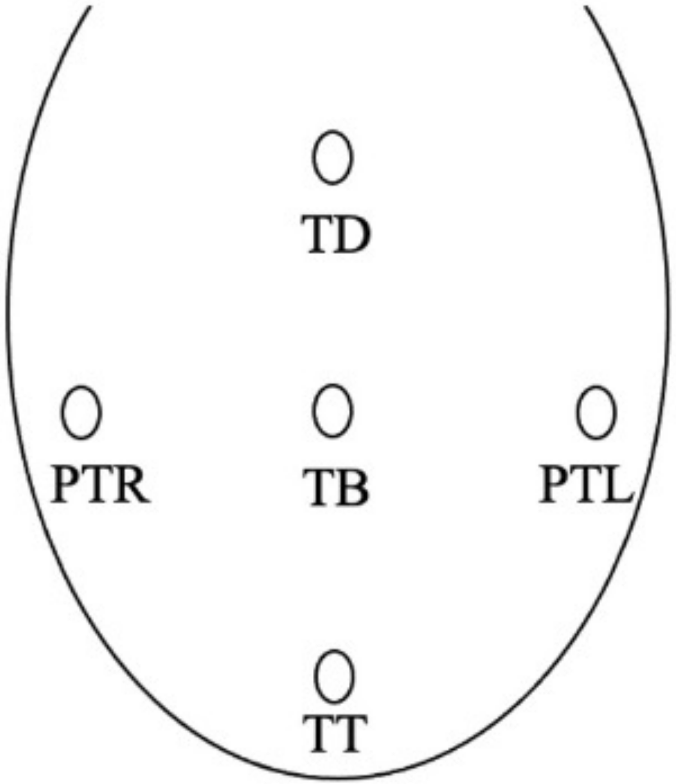
Each item was presented in random order within a block (15 blocks)

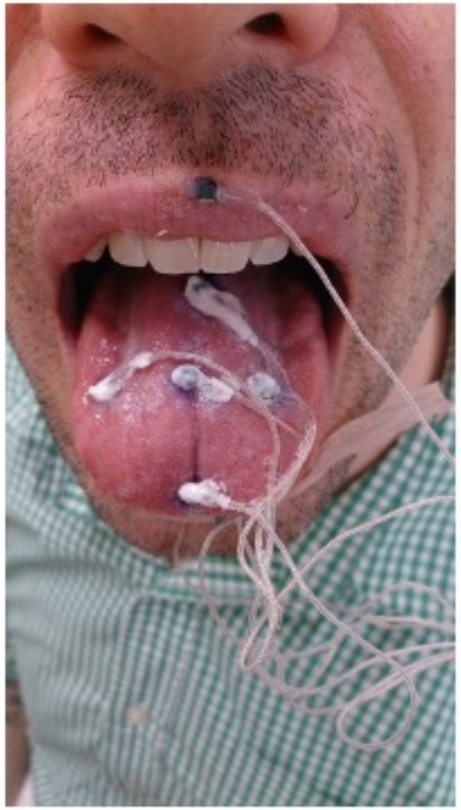
Total number of tokens included in analysis = 942

Data collection

NDI Wave EMA system
sampling at 100 Hz

Lingual sensors:
Tongue dorsum (TD)
Tongue blade (TB)
Tongue tip (TT)
Parasagittal tongue left (PTL)
Parasagittal tongue right (PTR)

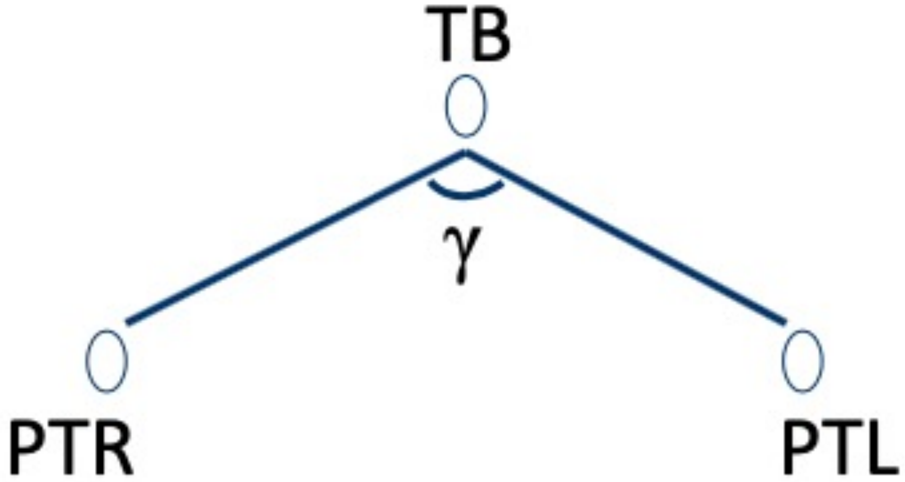




Primary parasagittal measure

Angle under the tongue (γ) in degrees calculated using the law of cosines

(Howson et al., 2015)



$$\gamma = \arccos((LB^2 + RB^2 - LR^2) / (2 * LB * RB)) * (180 / \pi)$$

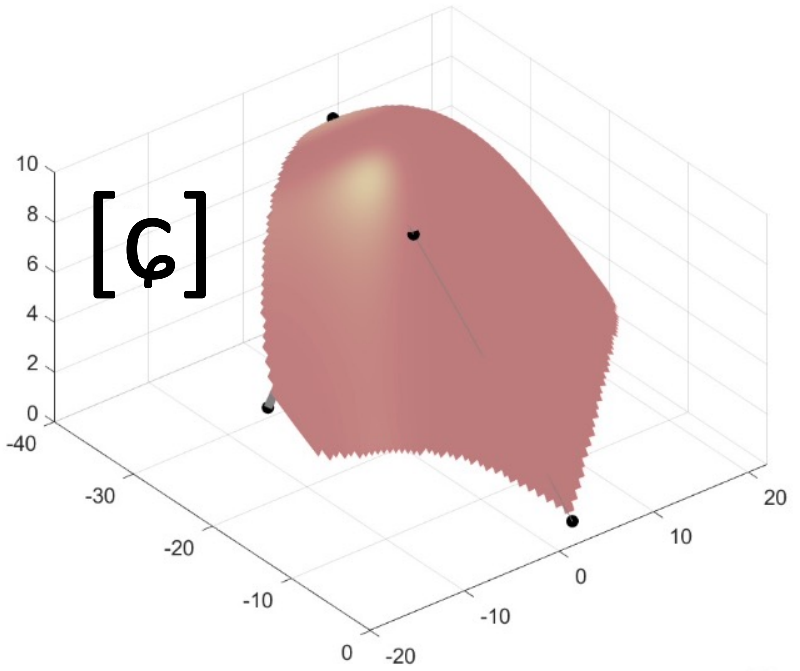
LB is the Euclidean distance between TB and PTL; RB is the Euclidean distance between TB and PTR; and LR is the Euclidean distance between PTL and PTR

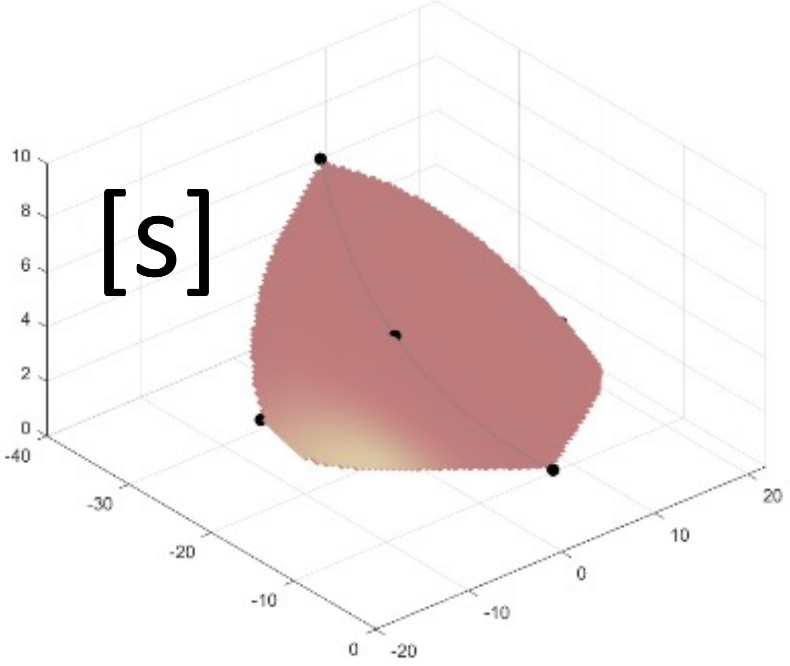
Qualitative results

(1) [ɕ] tends to be more domed than [s]

(2) [ɕ] tends to have a higher TB than [s]

(3) [ɕ] tends to have a lower TT than [s]





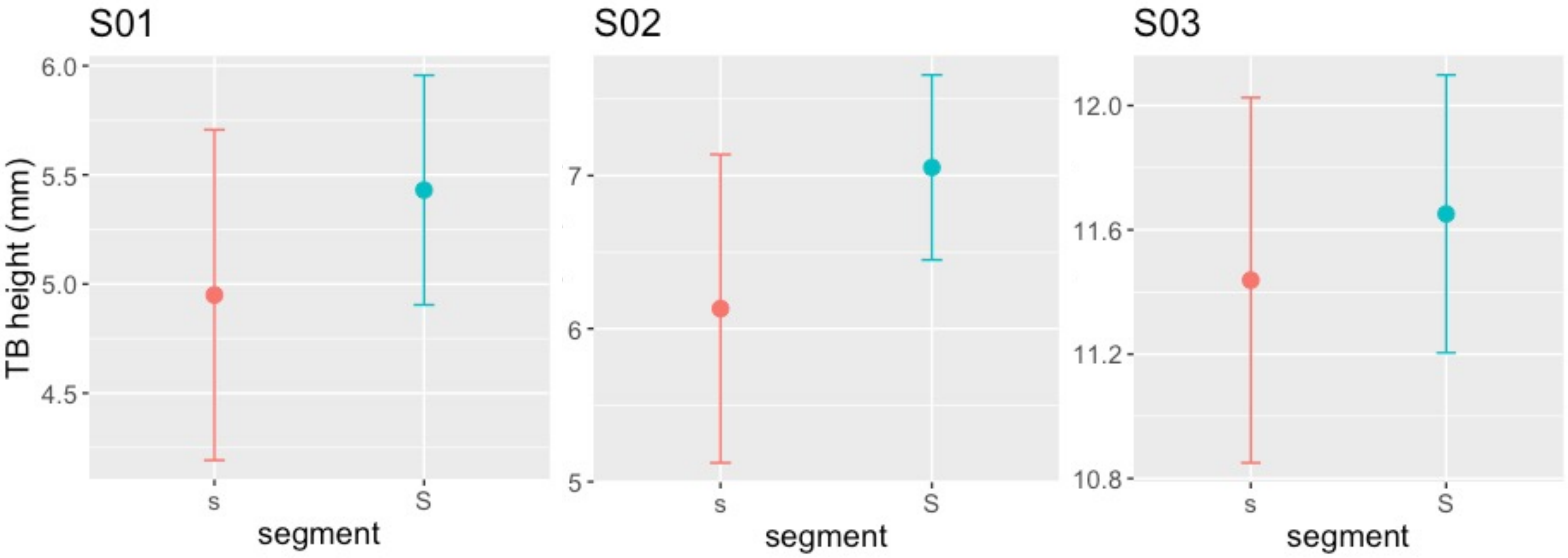
Hypotheses from qualitative patterns

- [ɕ] is articulated primarily with the TB (TB under phonological control)
- [s] is articulated primarily with the TT (TT under phonological control)
- Domed shape of [ɕ] is a passive consequence of raising the TB
- **Prediction**: Negative relationship between TB height and angle under the tongue (γ), regardless of segment

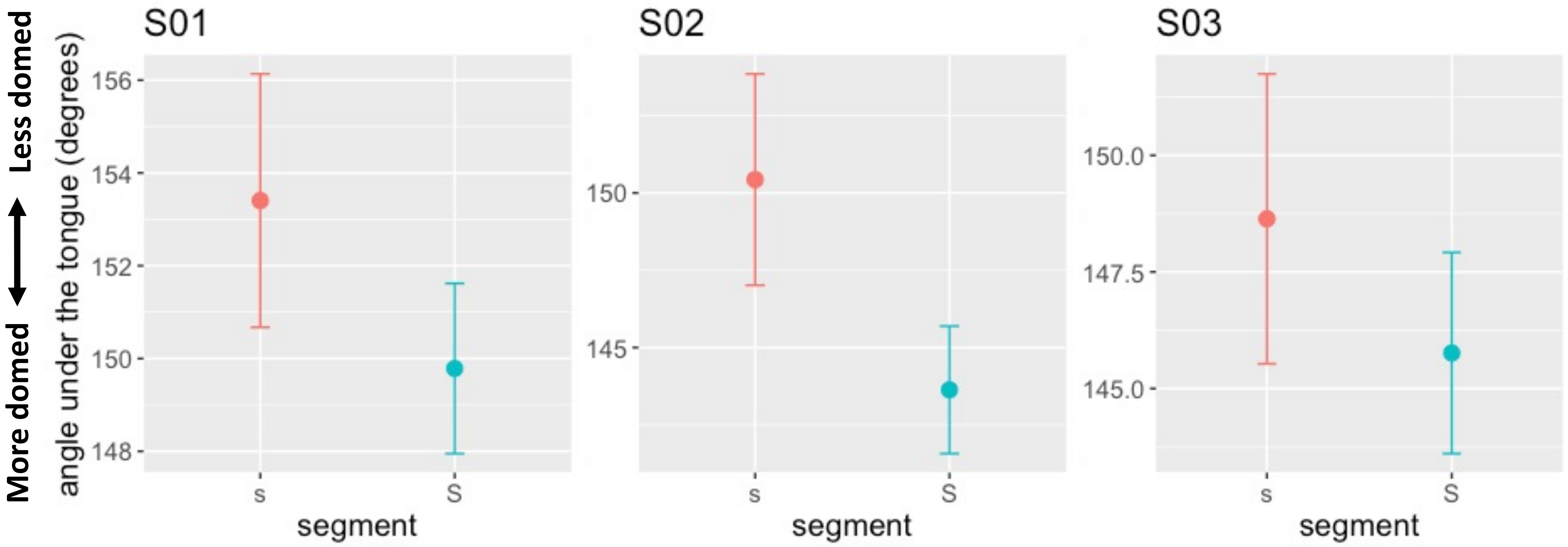
Quantitative analysis

- Acoustic data was force aligned using WebMAUS (Kisler et al., 2017)
- Gamma (doming) and TB height were calculated at the temporal midpoint of each sibilant token
- Examine average TB height and gamma by segment, as well as relationship between TB height and gamma
- Linear mixed effects models

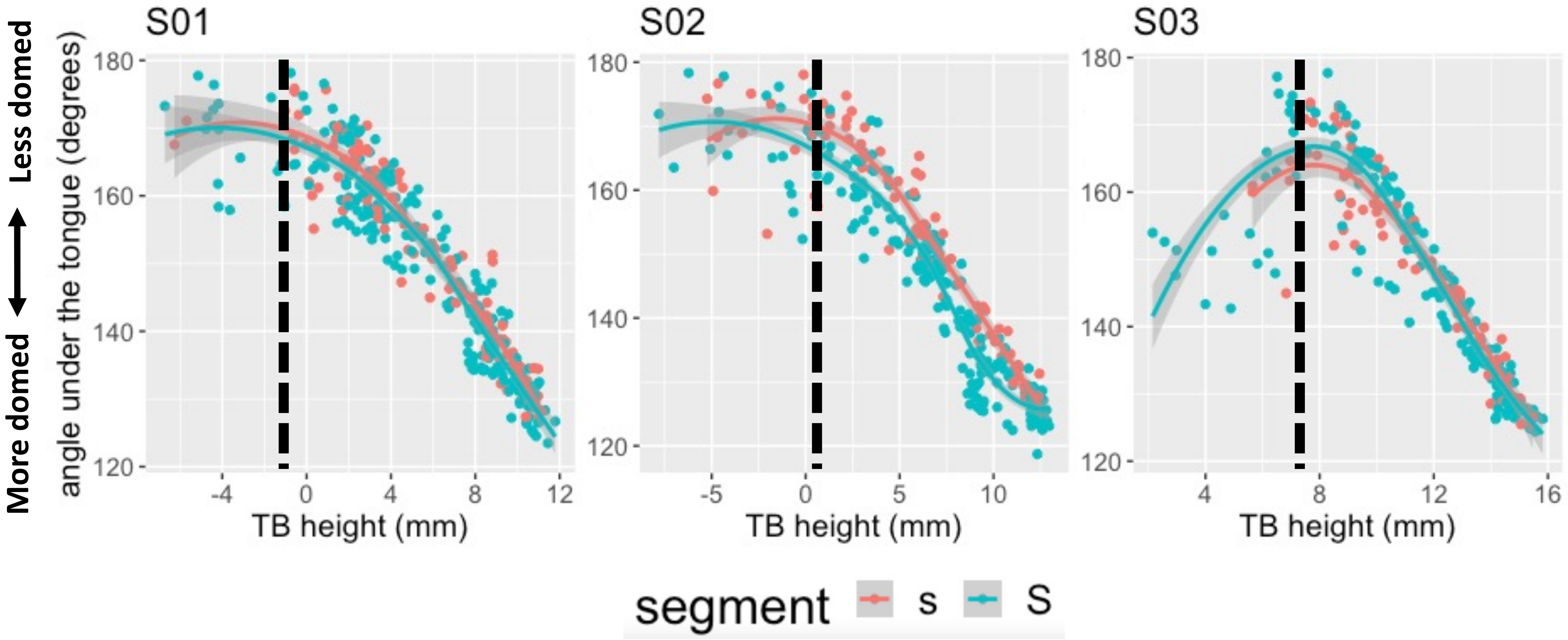
TB height by segment



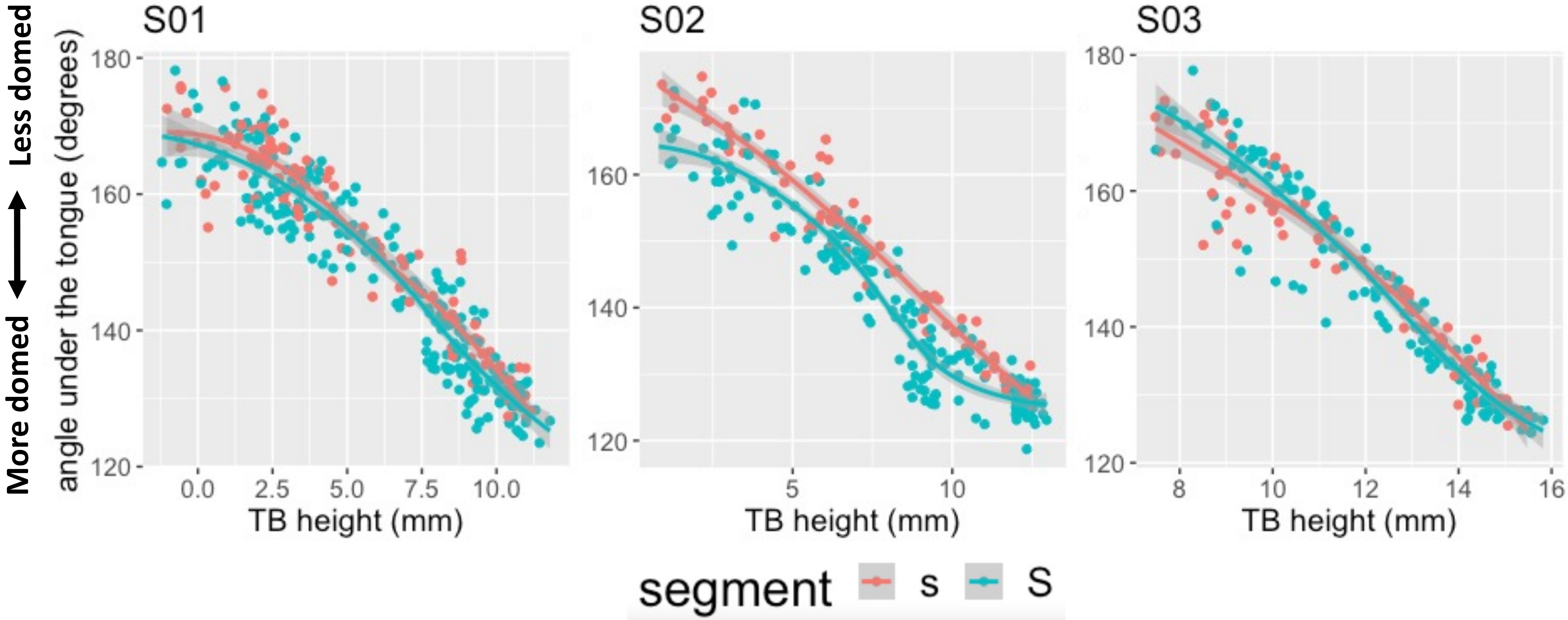
Angle under the tongue (gamma) by segment



Relationship between TB height and gamma



Trimmed dataset (> 1.5 SD below the mean)



Interim summary

- [ɕ] has higher TB than [s]
- [ɕ] is more domed than [s]
- For both segments, higher TB = more doming
- Can the by-segment difference in doming be entirely explained by the by-segment difference in TB height?
  - Or is segment type independently predictive of doming?

Linear mixed effects models

- To test this, we fit nested linear mixed effects models to gamma
- Both gamma and TB height were z-scored
- Segment identity was sum-coded: [s] = 1, [ɕ] = -1

Model structure

Baseline =	gamma ~ (TB_height + segment | subject) + (TB_height | item)
+ TB_height =	gamma ~ TB_height + (TB_height + segment | subject) + (TB_height | item)
+ segment =	gamma ~ TB_height + segment + (TB_height + segment | subject) + (TB_height | item)
+ interaction =	gamma ~ TB_height * segment + (TB_height + segment | subject) + (TB_height | item)

Model comparison

	npar	AIC	BIC	deviance	Chi-Sq	df	p value
baseline	11	548.57	600.84	526.57			
+ TB_height	**12**	**535.66**	**592.69**	**511.66**	**14.907**	**1**	**< .001**
+ segment	13	535.81	597.59	509.81	1.849	1	0.174
+ interaction	14	535.85	602.38	507.85	1.962	1	0.161

Summary of best-fitting model: fixed effects

	Estimate	Std. Error	df	t value
(Intercept)	0.254	0.056	2.494	4.567
TB_height	-1.282	0.072	2.171	-17.764

By-subject random effects

	Estimates		
Subject	(Intercept)	TB_height	segment=s
S01	0.110	-1.082	0.051
S02	0.146	-1.132	0.145
S03	0.329	-1.386	-0.016

All subjects show strong consistent effects of TB height

S01 and S02 show a small effect of segment type in the expected direction, but S03 shows a small effect in the opposite direction

Summary of model results

- Most variance in gamma is explained by TB height

- Effect of segment type on gamma not significant across subjects

Discussion

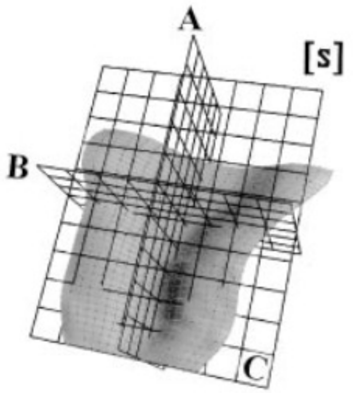
- Results consistent with the hypothesis that Japanese sibilants are produced without active parasagittal control
- Rather, parasagittal tongue shape during Japanese sibilant production may be a passive consequence of TB height control
- Consistent with cross-linguistic variation in dimensions of phonological control, even for very similar sounds

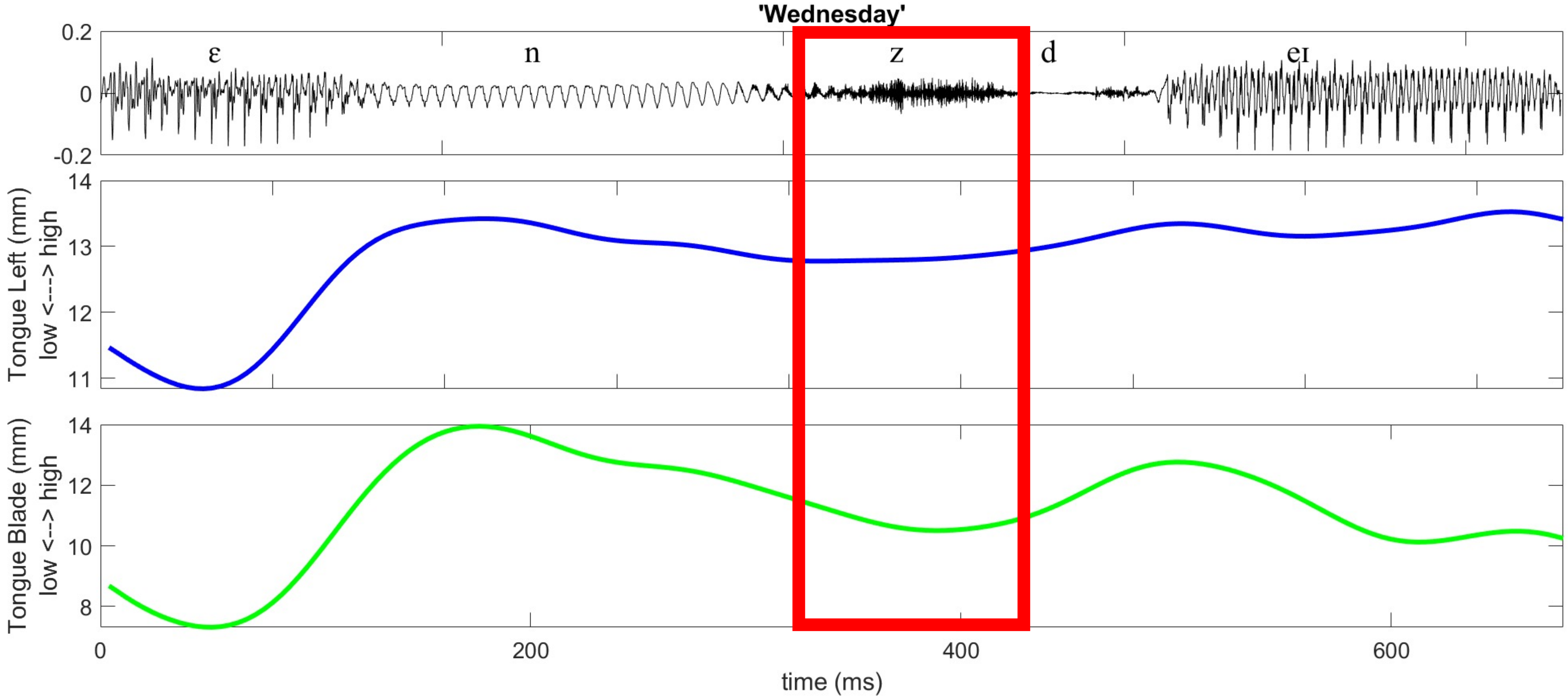
Next step: English EMA data

- English is predicted to show a different relationship between segment identity, TB height, and parasagittal tongue shape

- If English sibilants involve active parasagittal control, we would expect a stronger, more consistent effect of segment identity on parasagittal tongue shape

Preview: [z] grooving in ‘Wednesday’





(data from Ji et al., 2014)

Implications

- Is parasagittal control related to phonotactics?
  - e.g., deeper grooving sustains sibilant in English consonant clusters?

- Does English tense/lax distinction involve parasagittal control? (Stone & Lundberg, 1996)

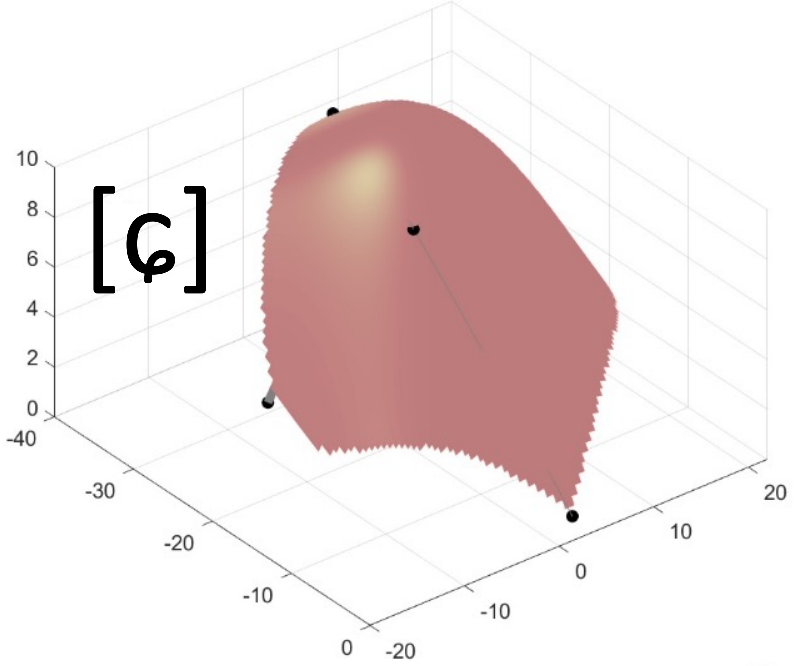
- Does relative lack of parasagittal control in Japanese underlie difficulty of Japanese speakers learning English rhotic ~ lateral contrast, which likely involves parasagittal control? (Ying et al., 2021)

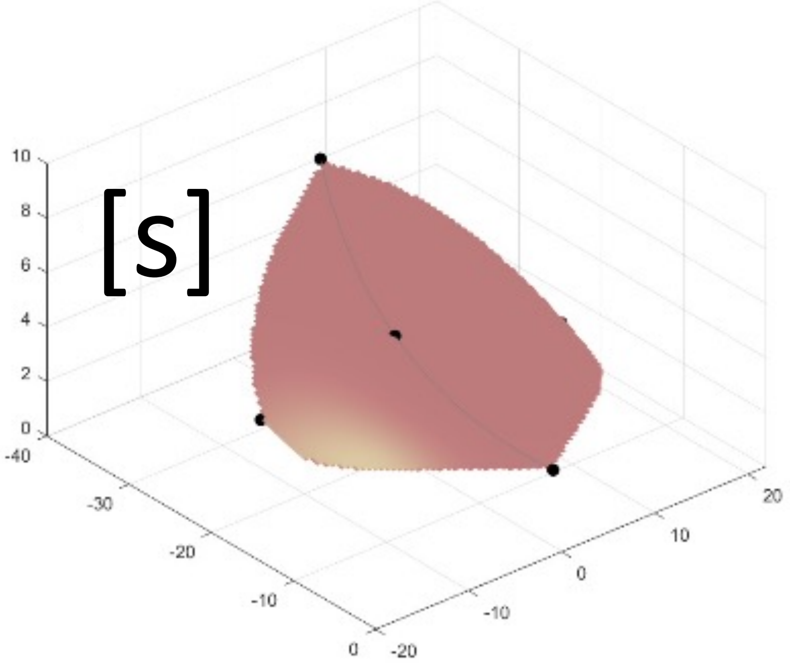
Other next steps

- More Japanese data (different speakers, different EMA sensor arrangements)
  - How robust is this pattern in Japanese?
- Biomechanical modeling in Artisynth (Stavness et al, 2014)
  - What underlies the relationship between TB height and parasagittal tongue shape?
- Investigate the nonlinearity: why the qualitatively different pattern at lower TB heights?

Thank you!

- To the experiment participants
- To the Yale Phonologroup
- To AMP 2021 participants





References

Hombert, J.-M., Ohala, J. J., & Ewan, W. G. (1979). Phonetic explanations for the development of tones. *Language*, *55*(1), 37–58.

Howson, P., Kochetov, A., & van Lieshout, P. (2015). Examination of the grooving patterns of the Czech trill-fricative. *Journal of Phonetics*, *49*, 117–129.

Ji, A., Berry, J. J., & Johnson, M. T. (2014). The electromagnetic articulography Mandarin accented English (EMA-MAE) corpus of acoustic and 3D articulatory kinematic data. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 7719–7723.

Kisler, T., Reichel, U. D., & Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech & Language*, *45*, 326–347.

Li, F., Edwards, J., & Beckman, M. E. (2009). Contrast and covert contrast: The phonetic development of voiceless sibilant fricatives in English and Japanese toddlers. *Journal of Phonetics*, *37*(1), 111–124.

Matsui, M. F. (2017). On the Input Information of the C/D Model for Vowel Devoicing in Japanese. *Journal of the Phonetic Society of Japan*, *21*(1), 127–140.

Pouplier, M., Hoole, P., & Scobbie, J. M. (2011). Investigating the asymmetry of English sibilant assimilation: Acoustic and EPG data. *Laboratory Phonology*, *2*(1), 1–33.

Stavness, I., Nazari, M. A., Flynn, C., Perrier, P., Payan, Y., Lloyd, J. E., & Fels, S. (2014). Coupled Biomechanical Modeling of the Face, Jaw, Skull, Tongue, and Hyoid Bone. In N. Magnenat-Thalmann, O. Ratib, & H. F. Choi (Eds.), *3D Multiscale Physiological Human* (pp. 253–274). London: Springer London.

Stone, M., & Lundberg, A. (1996). Three-dimensional tongue surface shapes of English consonants and vowels. *Journal of the Acoustical Society of America*, *99*(6), 3728–3737.

Whalen, D. H., Shaw, P., Noiray, A., & Antony, R. (2011). Analogs of Tahltan Consonant Harmony in English CVC Syllables. *International Congress of Phonetic Sciences (ICPhS)*, 2129–2132.

Ying, J., Shaw, J. A., Carignan, C., Proctor, M., Derrick, D., & Best, C. T. (2021). Evidence for active control of tongue lateralization in Australian English /l/. *Journal of Phonetics*, *86*.

Stimuli

弛緩性	ɕikansei	志願制	ɕigansei
しこり	ɕikori	仕事	ɕigoto
死闘	ɕitoo	指導	ɕidoo
死体	ɕitai	私大	ɕidai
主観	ɕukan	主眼	ɕugan
主体性	ɕutaisei	主題歌	ɕudaika
趣向	ɕukoo	酒豪	ɕugoo
酒盗	ɕutoo	手動	ɕudoo
ストライク	sutoraiku	すどく	sudoku
スタンプ	sutampu	すだち	sudachi
少し	sukoɕi	菅野	sugano
スカイ	sukai	すごい	sugoi