



# Acoustic distance effect on the perception of sibilants mergers between retroflexes and alveolars in Taiwan Mandarin (5aSC)

Baichen Du

Department of Linguistics, University of Hong Kong | baichen@connect.hku.hk



## Overview

- A production and perception experiment investigated aspects of the Mandarin sibilant merger and found:
  - Lower center of gravity (CoG) in production of all sibilants than previously reported
  - Merger-in-progress of retroflexes and alveolars, conditioned by social factors
  - A link between production and perception: Lower center of gravity in production than previous reports.
- Implications on the cause of the merger and sound change:
  - Listener-based sound change: Effects of environment on perception.
  - Social effects of gender and region.
- Future research may study the cause and exact structure of the merger, i.e., category shifting or expanding.

## Background

### Sibilants in Mandarin

- A three-way contrast among sibilant sounds in Mandarin, distinguished by place and manner of articulation, and aspiration<sup>1</sup>.
  - Alveolars: /ts/, /ts<sup>h</sup>/, /s/
  - Retroflexes: /tʂ/, /tʂ<sup>h</sup>/, /ʂ/
  - Palatoalveolars: /tʃ/, /tʃ<sup>h</sup>/, /ʃ/
- Components of sibilants:
  - Affricates: closure, burst, **frication**, (aspiration)
  - Fricatives: **frication**
- Canonical production: higher frequency frication for alveolars and lower for retroflexes<sup>2,3,4</sup>.
  - CoG for Alveolars: 9000-10000Hz
  - CoG for Retroflexes: 4000-5000Hz
- Merger-in progress: from retroflexes to alveolars<sup>1,4,5</sup>.
  - Loss of retroflexion: Fronted place of articulation.
  - Acoustic correlate: Increased CoG for retroflex.
- Asymmetry & Unidirectionality

### Speech production and perception link

- Bidirectional influence between production repertoire and perceptual space.
- Explained by the Exemplar Model<sup>6</sup>:
- Ample evidence: speech accommodation/imitation<sup>7,8,9</sup> and second language acquisition<sup>10</sup>.
- Complexity of the link: "Filtering" of tokens entering production repertoire, causing a mismatch between production and perception<sup>7</sup>.

## Objectives

- The current study investigates if the link applies to the merger.
- Research Question: Is there a production and perception link among speakers in the sibilant merger? Specifically, do merged speakers also shift their perceptual boundary between retroflexes and alveolars?**
- Hypothesis: Merged speakers have a perceptual boundary shifted towards higher frequency, i.e., the merger is relatively symmetric between production and perception.
- Null hypothesis: Merged speakers do not shift the boundary, suggesting an asymmetric and complex relationship between production and perception.

## The Experiment

### Phase I: Production

- Production of minimal pairs
  - 23 mono- or disyllabic Mandarin Tone 1 words
  - Initial target syllable position (initial stress)
  - Onsets: /ts/, /ts<sup>h</sup>/, /s/ and /tʂ/, /tʂ<sup>h</sup>/, /ʂ/
  - Vowels: /i/, /ə/, /a/
  - 10 additional fillers
  - Three repetitions in random order

### Phase II: Perception

- Speech synthesis
  - Synthesis of frication using a Praat<sup>11</sup> script<sup>12</sup>
  - Three pieces of white noise were high- and low-pass filtered at different frequencies, by adjustable amplitude and slopes.
  - Parameters gained from natural production of Taiwan Mandarin, where sibilants are almost merged.
  - Splicing of frications onto retroflexes' closure, burst, aspirations and vowels.
- Two-alternatives forced choice.
  - The same 23 minimal pairs as two ends of continuums.

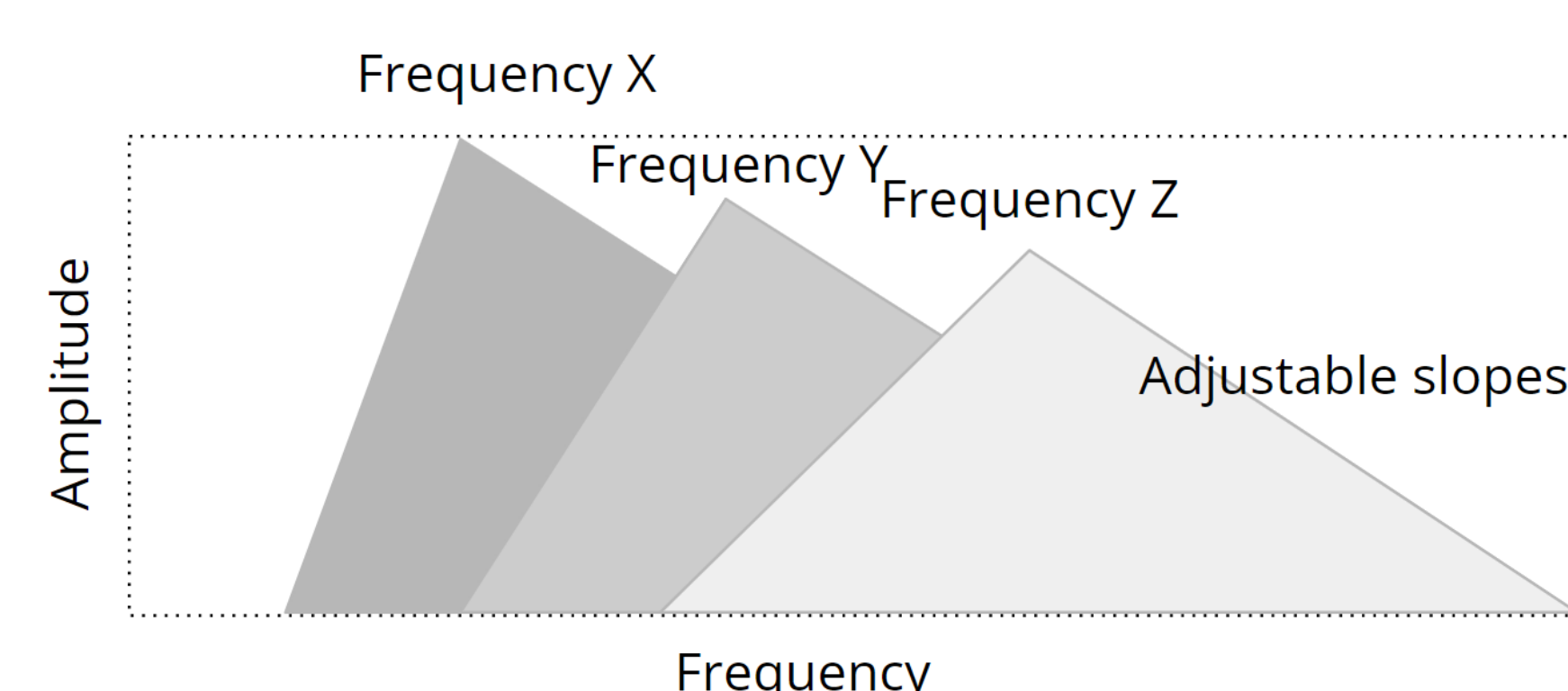


Figure 1. Illustration of frication synthesis.

### Phase III: Demographic questionnaire

- Age, gender, language and dialect background, living history

### Participants

- 23 native speakers of Mandarin at the University of Hong Kong.
- Born and raised in China until at least 18 years old.
- Bilinguals or trilinguals of Mandarin, English, and a Chinese dialect.

### Procedure

- Experiment conducted in a sound-treated booth.
- Three phases were presented in PsychoPy<sup>13</sup> and Qualtrics.
- Recording using Praat.
- Production preceded perception to avoid phonetic accommodation.

## Analysis

- Annotation and segmentation of recordings and stimuli in Praat.
- Measurement of four spectral moments of frications:
  - center of gravity (CoG)**
  - standard deviation
  - skewness
  - kurtosis.
- Kolmogorov-Smirnov test on CoG was performed to classify subjects into unmerged and merged speakers.
- Perception results and questionnaire were analyzed using R.
- Tests on the degree of merger predicted by social factors and perceptual boundary predicted by merger, step, and onsets.

## Production results

- CoG of sibilants is lower than previous reports found.

Pair	Retroflexes	Alveolars	Difference	N
/s/-/s/	3156	6657	3501	420
/ʂ/-/ʂ/	2649	5998	3349	540
/tʂh/-/tʂh/	2413	5232	2724	418

Table 1. Production results.

- Female, northeastern and southeastern speakers lead the merger.

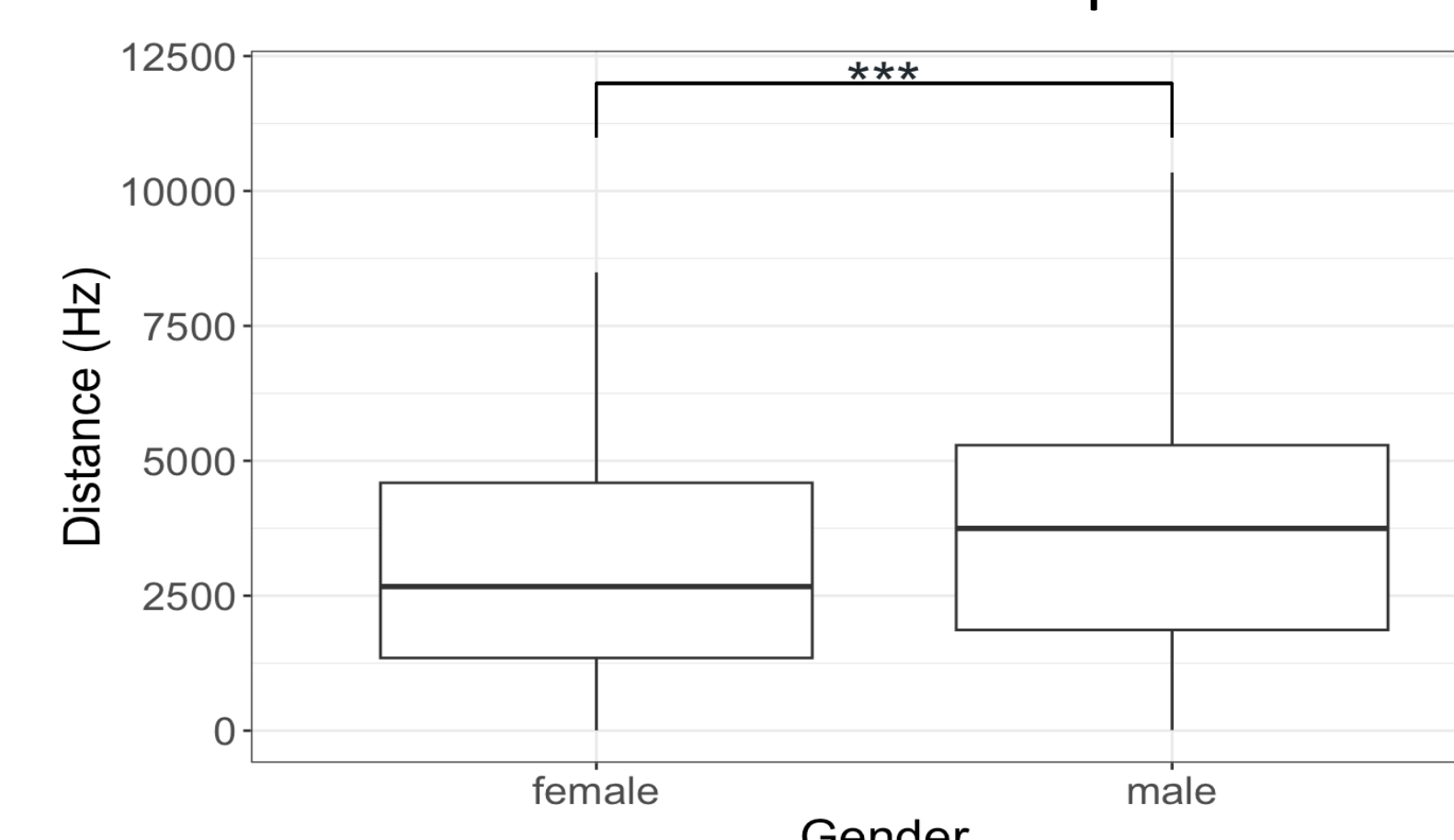


Figure 2. Gender effect on the merger.

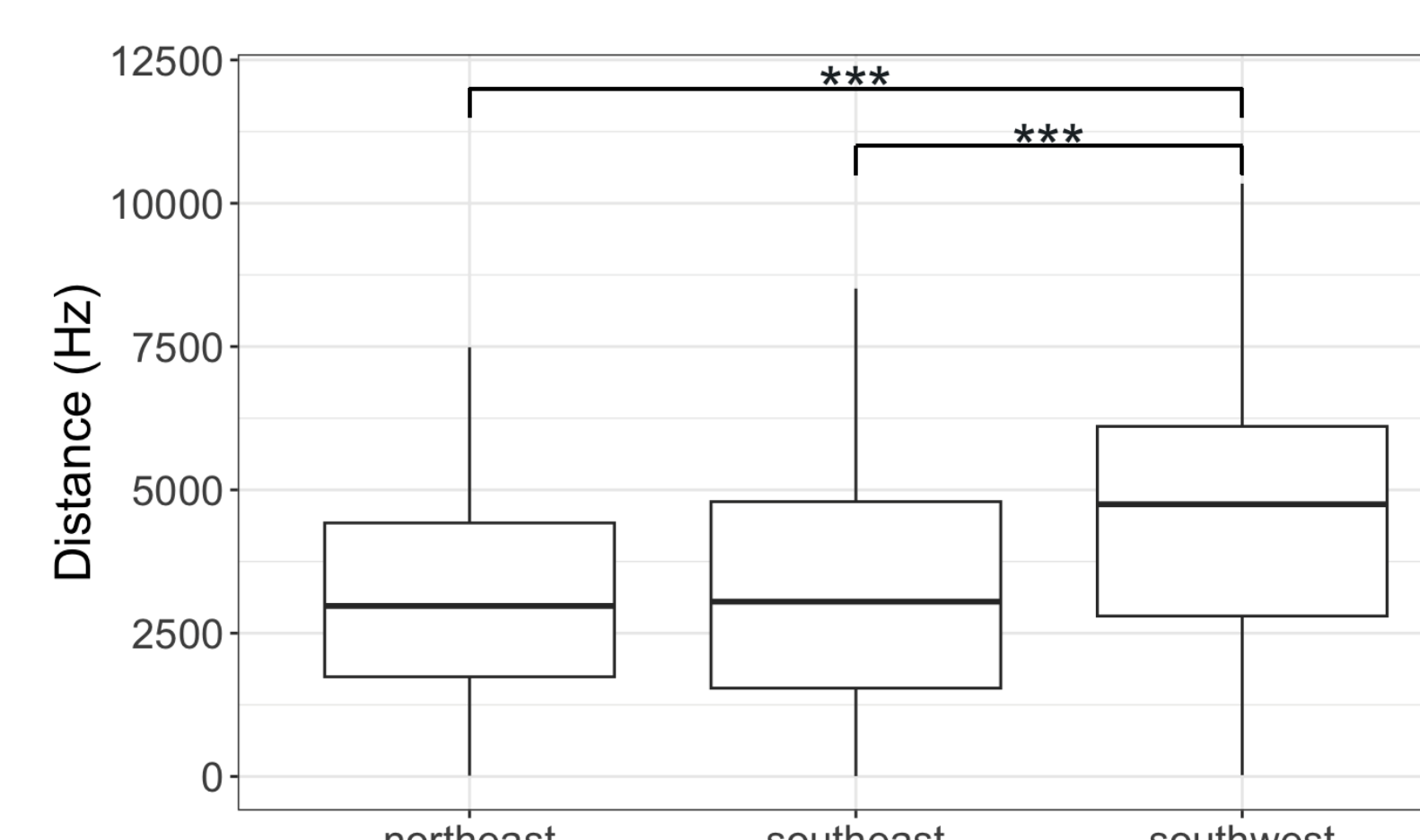


Figure 3. Geographic effect on the merger.

## Perception results

- General perceptual boundary is at 6000 to 7000Hz (Step 4 or 5).

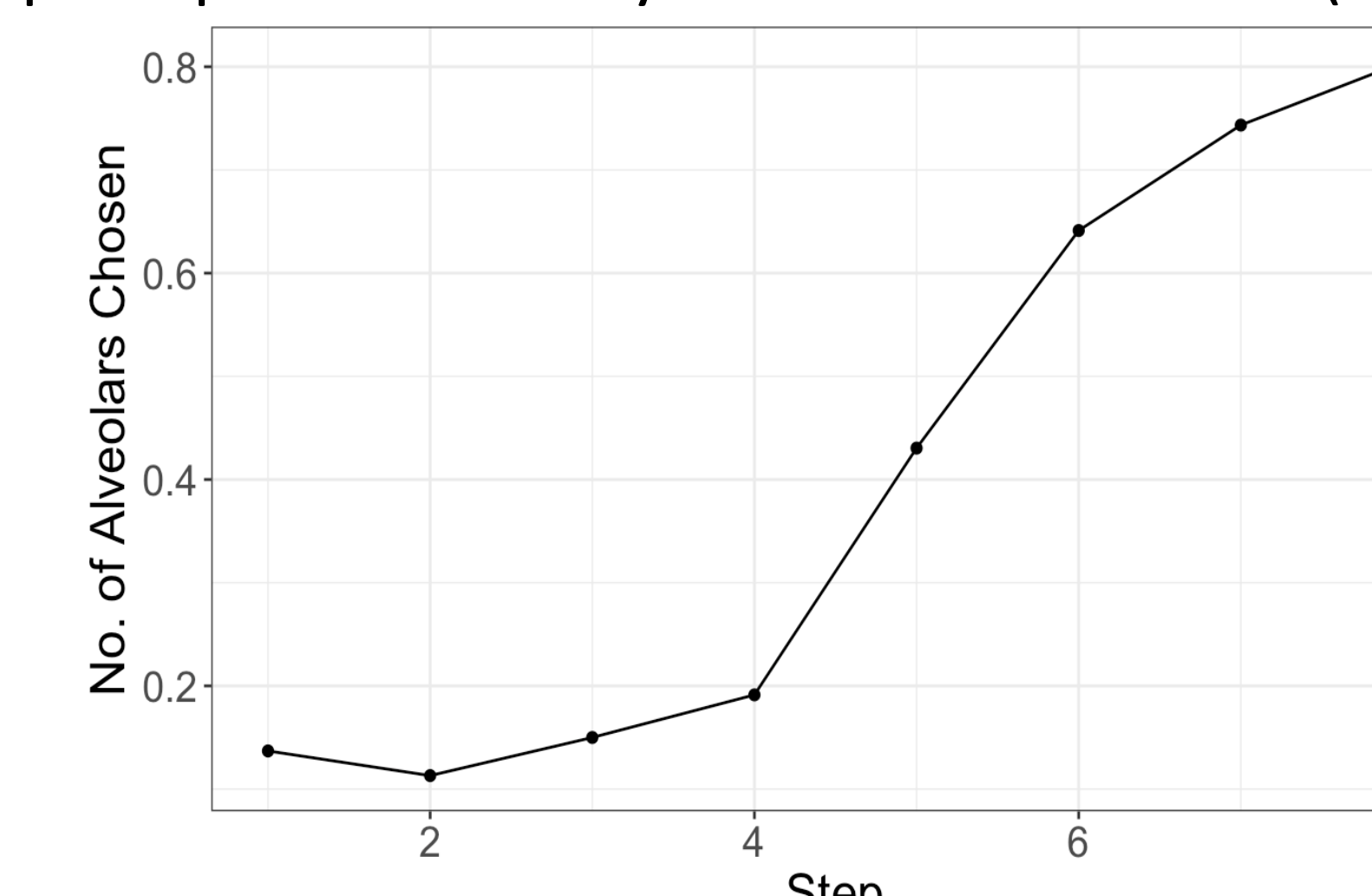


Figure 4. General perceptual curve of all participants.

## Results

- The boundary might also be conditioned by onset (Figure 5), but this may be confounded by how we synthesized frication (Figure 6).

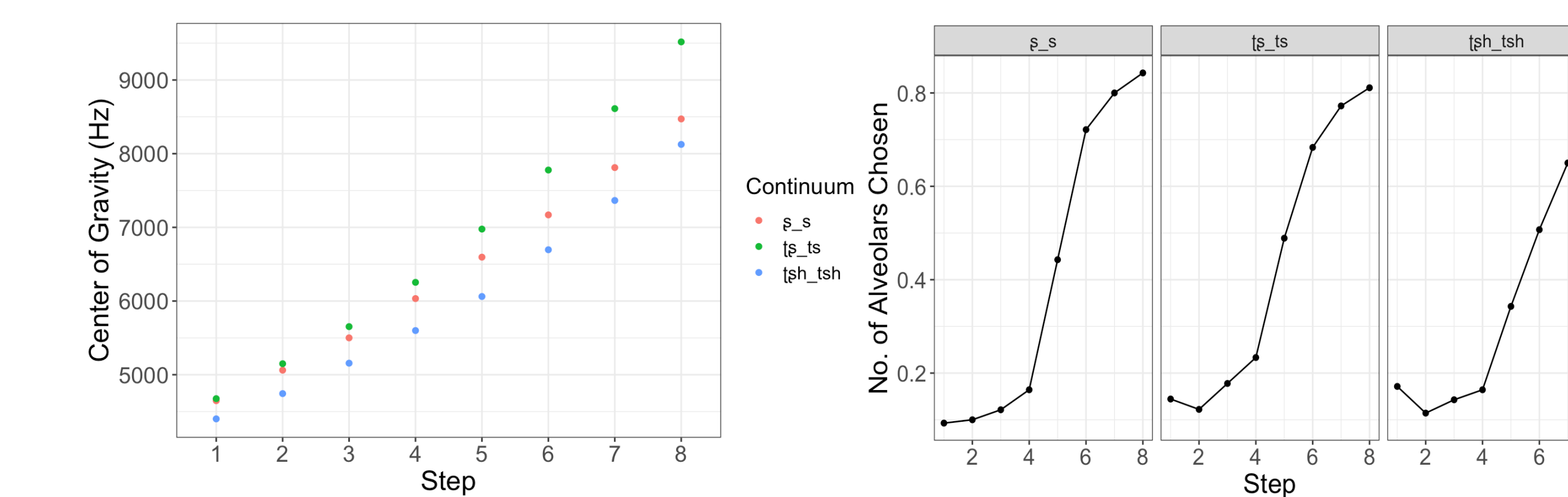


Figure 5. CoG of synthesized stimuli.

Figure 6. By-onset boundary.

### Production-perception link

- Compared to unmerged speakers, merged speakers are more likely to tolerate high-frequency retroflexes. It is also possible that they are more sensitive to cues other than CoG, such as F2 offset on vowels, due to the contexts surrounding synthesized frications.

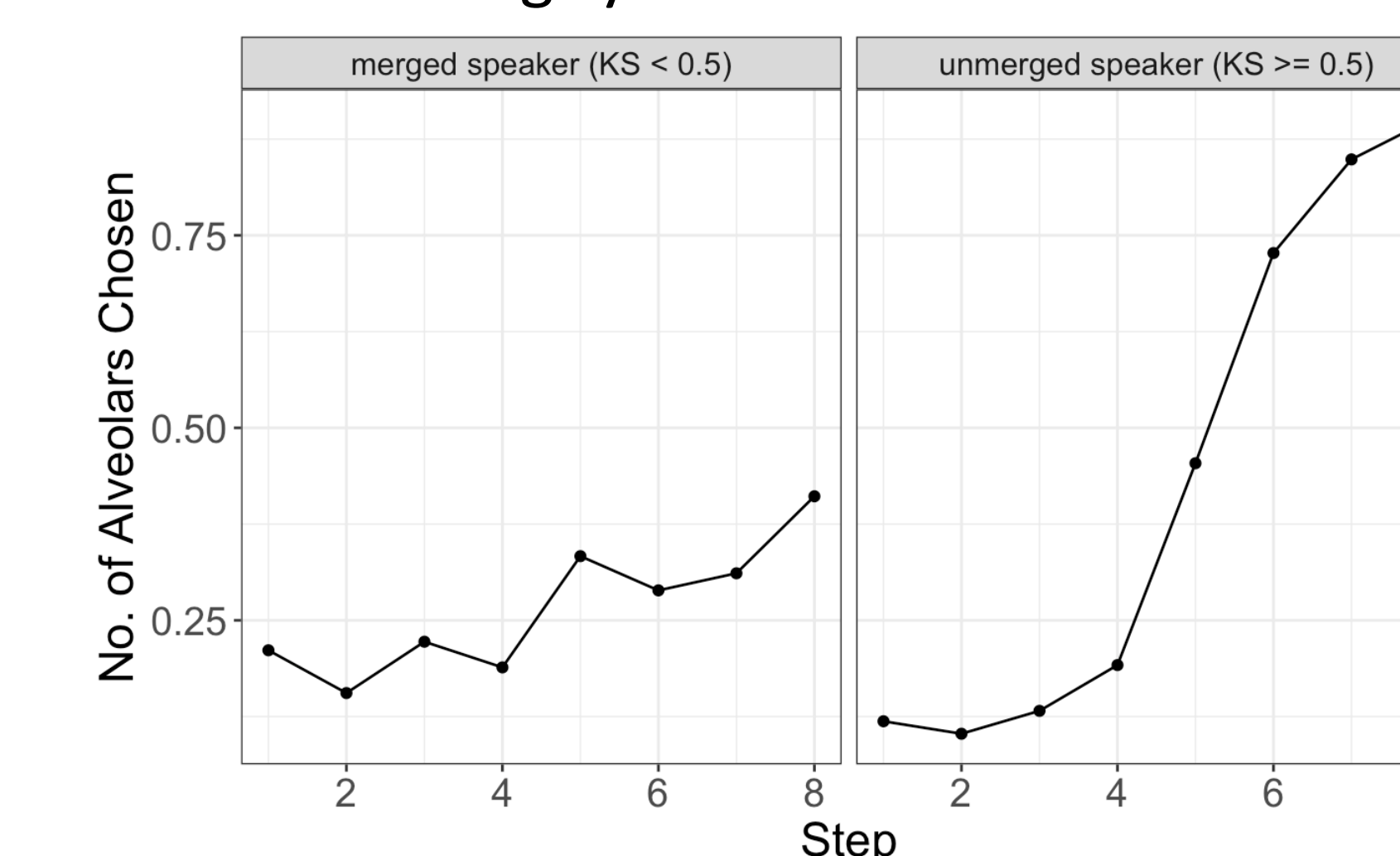


Figure 7. Perceptual boundary of merged and unmerged speakers.

- Logistic Regression revealed significant effects of step, onsets, and speaker (merged/unmerged). Mixed-effects model failed to converge. Bayesian Mixed-effects regression showed similar results.

	Estimate	Std. Error	Z value	Pr (> z )
(Intercept)	-2.4063	0.2397	-10.038	< 2e-16 ***
step 2	-0.2221	0.2015	-1.102	0.27028
step 3	0.1078	0.1897	0.568	0.56976
step 4	0.4062	0.1817	2.236	0.02536 *
step 5	1.6042	0.1672	9.594	< 2e-16 ***
step 6	2.5092	0.1699	14.772	< 2e-16 ***
step 7	3.019	0.1763	17.126	< 2e-16 ***
step 8	3.3866	0.1836	18.442	< 2e-16 ***
tʂ_ts	-0.294	0.2668	-1.102	0.27055
tʂh_tʂh	-0.6429	0.2482	-2.59	0.00959 **
unmerged	0.664	0.214	3.103	0.00192 **
tʂ_ts:unmerged	0.4726	0.2874	1.644	0.10008
tʂh_tʂh:unmerged	0.5461	0.2754	1.983	0.04738 *

Table 2. Logistic Regression on step, onsets, speaker status, and interactions.

## Conclusion & Future Directions

- There is indeed a merger in the participants we tested, led by female speakers, consistent with previous sound change studies.
- Compared to unmerged speakers, merged speakers are more likely to perceive more retroflexes along the same continua.
- The merger might be conditioned by onsets or vowels, or cue weighting differences in perception among the population.
- Future research should investigate if the merger is articulatorily or perceptually motivated, if there are any cue weighting differences among the population, if the merger is caused by category shifting or expanding, and if the two categories collapsed in production or perception first.

## Selected References

- Chiu, Chenhao, Po-Chun Wei, Masaki Noguchi, and Noriko Yamane. "Sibilant fricative merging in Taiwan Mandarin: An investigation of tongue postures using ultrasound imaging." *Language and speech* 63, no. 4 (2020): 877-897.
- Svantesson, Jan-Olof. "Acoustic analysis of Chinese fricatives and affricates." *Journal of Chinese Linguistics* 14, (1986): 177-189.
- Lee, Chao-Yang, Yu Zhang, and Ximing Li. "Acoustic characteristics of voiceless fricatives in Mandarin Chinese." *Journal of Chinese Linguistics* 42, no. 1 (2014): 150-171.
- Chang, Yang-Hsiang. "Variability in cross-dialectal production and perception of contrasting phonemes: The case of the alveolar-retroflex contrast in Beijing and Taiwan Mandarin." *Journal of Phonetics* 40, no. 1 (2012): 177-189.
- Chuang, Yu-Ying, and Janice Fon. "The effect of prosodic prominence on the realizations of voiceless dental and retroflex sibilants in Taiwan Mandarin spontaneous speech." In *Speech Prosody 2010-Fifth International Conference*. 2010.
- Garrett, Andrew, and Keith Johnson. "Phonetic bias in sound change." *Origins of sound change: Approaches to phonologization* 1 (2013): 51-97.
- Babel, Molly Elizabeth. *Phonetic and social selectivity in speech accommodation*. University of California, Berkeley, 2009.
- Babel, Molly. "Evidence for phonetic and social selectivity in spontaneous phonetic imitation." *Journal of Phonetics* 40, no. 1 (2012): 177-189.
- Babel, Molly, Michael McAuliffe, and Graham Haber. "Can mergers-in-progress be unmerged in speech accommodation?" *Frontiers in psychology* 4 (2013): 653.
- Munro, Murray J., Tracey M. Derwing, and James E. Flege. "Canadians in Alabama: A perceptual study of dialect acquisition in adults." *Journal of Phonetics* 27, no. 4 (1999): 385-403.
- Boersma, Paul. "Praat: Doing Phonetics by Computer." Praat: doing phonetics by computer, 2023. <https://www.fon.hum.uva.nl/praat/>.
- Winn, Matthew. "Make fricative continuum." 2022. <https://www.matthewwinn.com/praat.html#fricativeContinuum/>.
- Peirce, Jonathan, Jeremy R. Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman, and Jonas Kristoffer Lindelöv. "PsychoPy2: Experiments in Behavior Made Easy." *Behavior Research Methods* 51, no. 1 (February 1, 2019): 195-203.