Speech Errors in Fast Speech of Alveolar and Alveo-palatal Consonants: A Preliminary Study Using Ultrasound

Misato Shimizu, Takayuki Arai (Sophia University),
Ai Mizoguchi (Maebashi Institute of Technology), Maho Morimoto (Chuo University),
Weiyu Li (Sophia University), Noriko Yamane (Hiroshima University)

November 19, 2024 187th Meeting of the Acoustical Society of America 1pSCd, Student Choose-Your-Own Adventure

Introduction

- Speech errors occur even in typical speakers during fast speech.

- Speech errors are often gradual (Slis and van Lieshout, 2016; Dawson, 2020)
  ➞ difficult to identify based on auditory perception alone.

- We focused on alveolar and alveo-palatal consonants, /sa ɕa/ and /ta tɕa/,
  and observed the spectral Center of Gravity (CoG) and tongue movements
  - CoG corresponds to the length of the front cavity (Stevens, 1998)

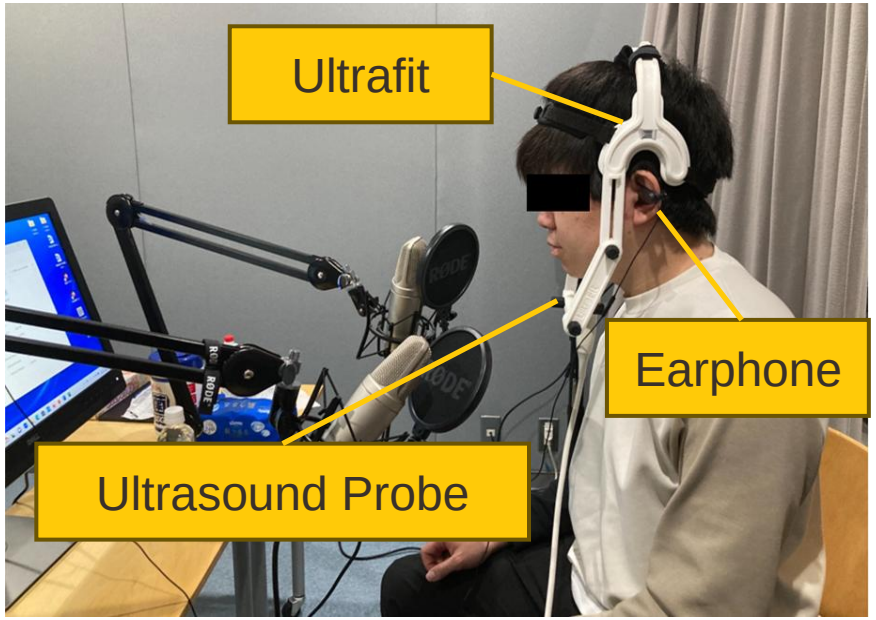
【Hypothesis】

As the tongue body elevates, the CoG would also increase.

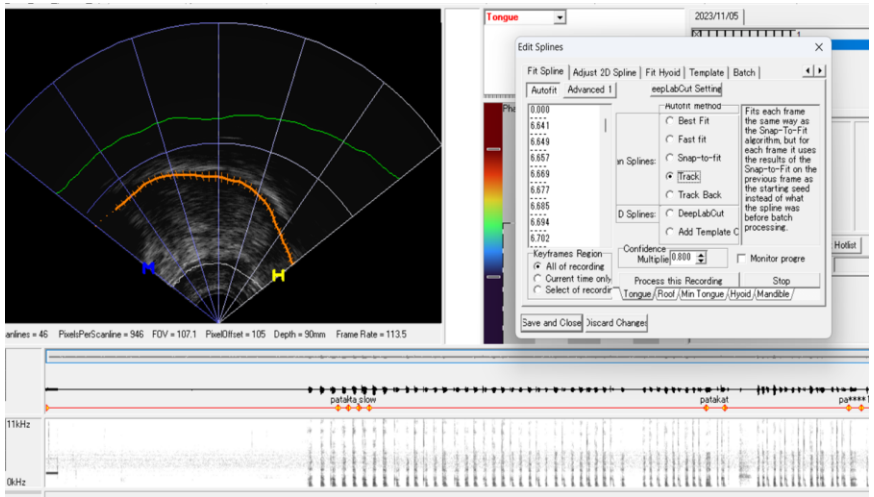
Methods

Participants: Native Japanese speakers
(M:2, F:1)

Experiment:
Pronounced repeatedly in time with the metronome
(BPM = 480, 6s)
Stimuli ➞ /sa ɕa/ and /ta tɕa/



Analysis:
1. Auditorily judgements of each token
   as correct or error
2. Measurements of each consonant's spectral CoG
3. Analysis of ultrasound image in AAA
(Articulate Instruments Ltd., 2023)



Results 1. /sa ɕa/

Error rate:

SP01 = 59%, SP02 = 24%, SP03 = 25%

(updated from the values listed in the abstract)

Mean CoG for correct utterances:

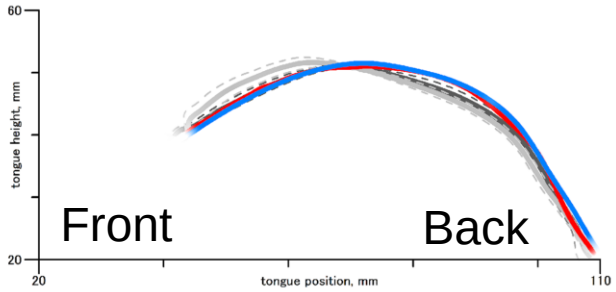
SP01→**/s/** was higher than /ɕ/

SP02 & SP03→ **/ɕ/** was higher than /s/

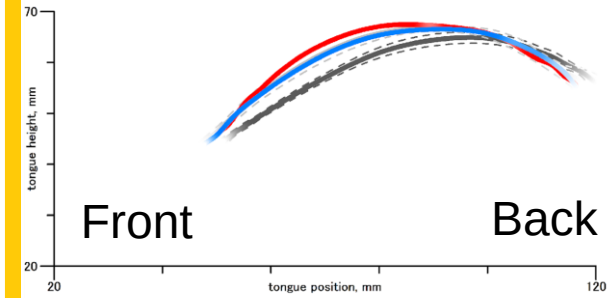
Four patterns of error pronunciation➡

adapted from Shimizu et al. (2024)

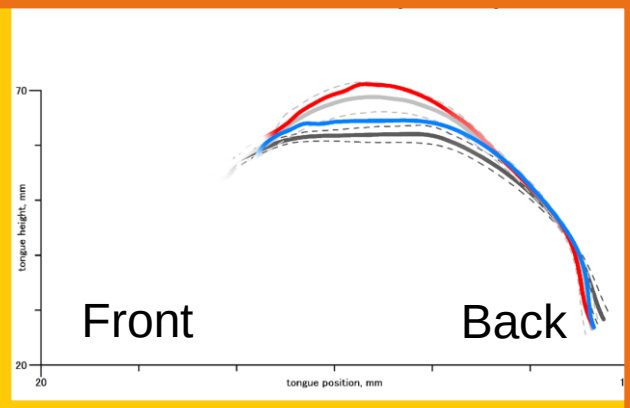
1. Intended /ɕ/ resembled tongue shape of /s/ (SP01)



2. Intended /s/ resembled tongue shape of /ɕ/ (SP02)



3. Metathesis (SP03)



4. Tongue contours are intermediate between /s/ and /ɕ/ (SP03)

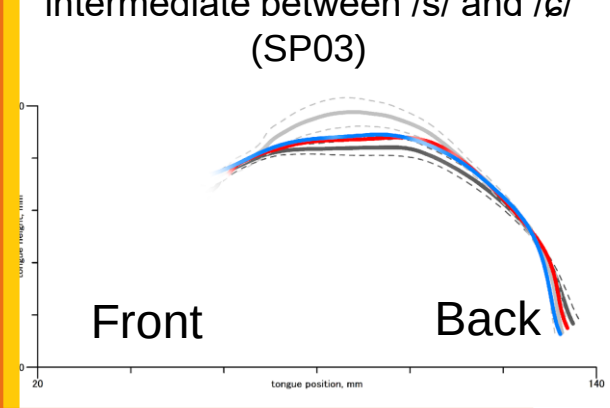


Fig.1 Average tongue contours of the correct pronunciation (**dark gray is /s/,** light gray is /ɕ/) and the tokens judged to be errors (**red is intended /s/**, **blue is intended /ɕ/**). The left side is tongue tip.

Results 2. /ta tɕa/

Error rate:

SP01 = 50%, SP02 = 50%, SP03 = 37%

(updated from the values listed in the abstract)

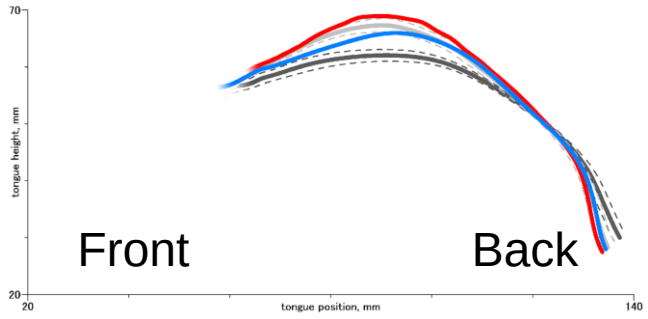
Mean CoG for correct utterances:

/tɕ/ was higher than /t/

Tongue contours:

- The same four error patterns observed in “sa sha”.
- In some cases, the tongue elevation changed without corresponding changes in CoG.

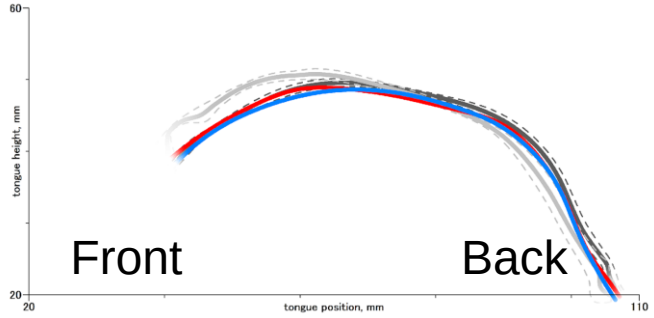
SP03



Tongue shape
Intended /t/ was higher than **intended /tɕ/.**

CoG
Intended /t/ (2521.3 Hz)
was higher than **intended /tɕ/** (929.7 Hz).

SP01



Tongue shape
Intended /t/ and **intended /tɕ/** resembled /t/.

CoG
Intended /tɕ/ (3268.1 Hz)
was higher than **intended /t/** (1971.7 Hz).

Fig.2 The overlay of the average tongue contour (**dark gray is /t/,** light gray is /tɕ/) of the correct pronunciation and the tokens judged to be incorrect (**red is intended /t/**, **blue is intended /tɕ/**). The left side is tongue tip.

Summary

- Some speakers found /sa ɕa/ more difficult, while others struggled more with /ta tɕa/.

- Errors were categorized into four types.
  1. intended /s, t/ resembles tongue shape of /ɕ, tɕ/
  2. intended /ɕ, tɕ/ resembles tongue shape of /s, t/
  3. metathesis
  4. tongue contours are intermediate between two targets

- CoG does not always reflect tongue elevation during errors.

References

Dawson, K. M. (2020) “Speech Motor Control in Younger and Older Adults: The Effect of Age and Individual Differences on Speech Error Production Age and Individual Differences on Speech Error Production.” City University of New York.

Slis, A. W., & Van Lieshout, P. H. H. M. (2016) “The effect of phonetic context on the dynamics of intrusions and reductions.” *Journal of Phonetics*, *57*, 1–20.

Stevens, K. (1998). *Acoustic phonetics*. Cambridge, MA: MIT Press.

Articulate Instruments Ltd. (2023). *Articulate Assistant Advanced (AAA)*. Version 221.2.0.

Boersma, P., Weenink, D. (2023) *Praat: Doing phonetics by computer.* Version 6.3.15.