

# Recoverability-driven coarticulation

Acoustic evidence from Japanese high vowel devoicing

---

Wojciech Stempniak

Based on Whang, J. (2018). Recoverability-driven coarticulation: Acoustic evidence from Japanese high vowel devoicing. *The Journal of the Acoustical Society of America*, 143(2), 1159-1172.

# Overview

1. Introduction
2. Previous studies
3. Study
4. Results
5. Discussion and conclusions

# Introduction

---

## Matter of the study



- High vowels = close vowels (in Japanese: /i/ and /u/)
- HV devoicing – losing fonation of V in  $C_1VC_2$  sequences where:
  - $C_1, C_2$  are voiceless obstruents (fricatives/plosives/affricates)
  - V is not accented
- High vowel devoicing is an integral feature of standard Japanese (Imai, 2010)

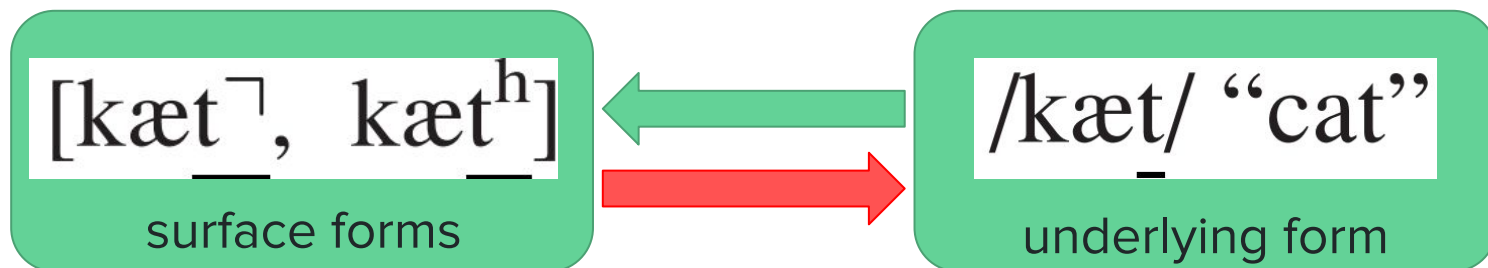
# High vowel devoicing

- results in a loss of:
  - Laryngeal adduction – **unphonated**
  - Lingual and labial gestures – **deleted** (controversial)
- is a post-lexical process: (Hirayama, 2009; Ito and Mester, 2003; Boersma, 2009; Hayes, 1999; Zsiga, 2000)
  - applies after lexical and structural processes (eg. syllabification and phonotactic evaluation)
  - devoiced sequences are not constrained by phonotactic rules

(1a) |ki + tai| → /ki.tai/ → [k̚tai]  
“expectation (time period + wait),”

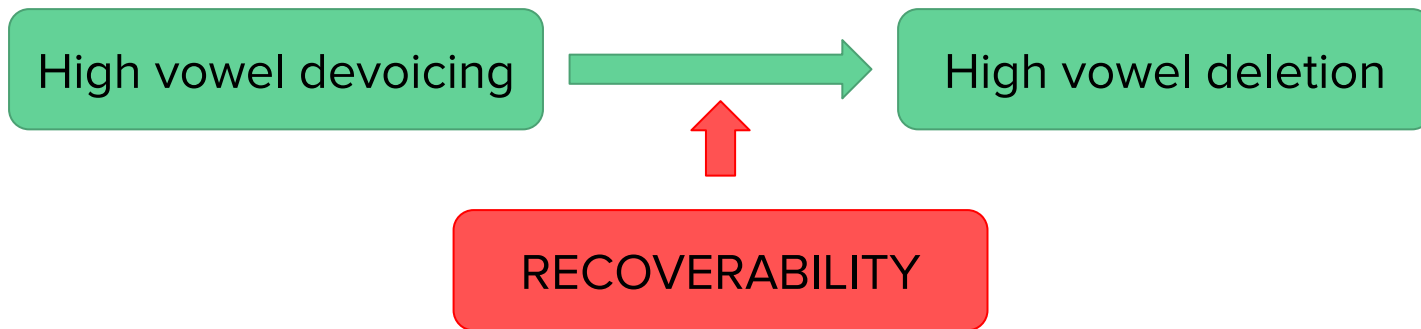
(1b) |tek + tai| → /te.ki.tai/ → [tek̚tai]  
“hostility (enemy + toward).”

## Recoverability



- The ease of accessing the **underlying form** from a given **surface form**
- Sources: (Mattingly, 1981; McCarthy, 1999; Chitoran et al., 2002)
  - perceptibility of articulatory cues present in the acoustic signal
  - predictability based on linguistic knowledge (eg. phonotactics)

## The aim of study



### “Recoverability-driven coarticulation”

**Hypothesis:** The choice between deletion and unphonating is dependent on the vowel’s recoverability

# Previous studies

---

# High vowel devoicing in Japanese

- Theories on ways of manifestation:
  - lengthening the burst/frication noise of  $C_1$  (Han, 1994)
  - unphonating the vowel and *coloring* the  $C_1$  burst/frication noise (Beckman and Shoji, 1984)
  - deleting the vowel altogether (Vance, 2008)
- Problems:
  - All three ways are contradicted by data (eg. Kondo, 1997)
  - Differences in the methodologies and stimuli among the studies

## Measuring predictability

- Two measures from Information Theory (Shannon, 1948)
- Surprisal – amount of information/effort necessary to predict a vowel
- Entropy – here: the weighted average of surprisal in a given context

$$\text{Surprisal} : -\log_2 \Pr(v|C_{1\_})$$

$$\text{Entropy (H)} : \sum \Pr(v|C_{1\_}) * (-\log_2 \Pr(v|C_{1\_}))$$

# Phonotactic predictability in Japanese

	<i>IPA</i>	<i>Entropy</i>	<i>Surprisal /i/</i>	<i>Surprisal /u/</i>
low predictability	k	$9.998 \times 10^{-1}$	0.979	1.021
actually /ç/ →	ɟ	0.555	0.199	2.955
high predictability	ϕ	0.123	5.903	0.024
	s	0.042	7.762	0.007
actually /tç/ →	tʃ	0.013	0.002	9.768
	ç	0.008	0.001	10.653

# Possible effects of predictability on coarticulation

- HV devoicing and deletion is either:
  - blind to predictability and driven by phonotactics (Kubozono, 2015)
  - a not systematic consequence of lexicalization of the vowel (Ogasawara and Warner, 2009)
  - fully constrained by recoverability → The claim of the study
- Question: is there a difference in lengthening or spectral changes of  $C_1$  based on the **predictability** of the vowel?

## Demographic differences

- Men tend to devoice more than woman (Okamoto, 1995)
- Especially younger men devoice more than younger woman (Varden and Sato, 1996; Imai, 2010)
- The authors believe that HV devoicing might be a feature of gendered speech and therefore construct a proper diverse sample

Study

---

# Participants and materials

- Participants
  - 22 Japanese speakers (12 woman + 10 man) aged 18-24, undergraduate students born and raised in the greater Tokyo area
- Materials
  - 60 native Japanese and Sino-Japanese words
  - controlled medium frequency based on a corpus (Sharoff, 2008)
  - none of the stimuli had a target vowel in an accented syllable

# Materials

- $C_1VC_2$  target sequence,  $V = \{i, u\}$
- $C_1$  – divided into **low predictability** (/k/, /ʃ/) and **high predictability** (/ϕ/, /s/, /tʃ/, /ç/) groups
- $C_2$  – divided into **devoicing** and **voicing** contexts
  - HV devoicing typically requires the target vowel to be flanked by two voiceless obstruents
  - $C_2$  is a **voiced** or a **devoiced stop**

## Design and procedure

- All tokens were placed in the context of a meaningful carrier sentences
- The sentences were a part of a story
- Participants recorded each sentence

(2) manatsu-no figaisen-ni-wa ki-o-tsu-ke-majo:  
midsummer's UV rays DAT-TOP be careful.VOL  
“Let’s be careful of midsummer’s ultraviolet rays”

# Devoicing analysis – devoicing context

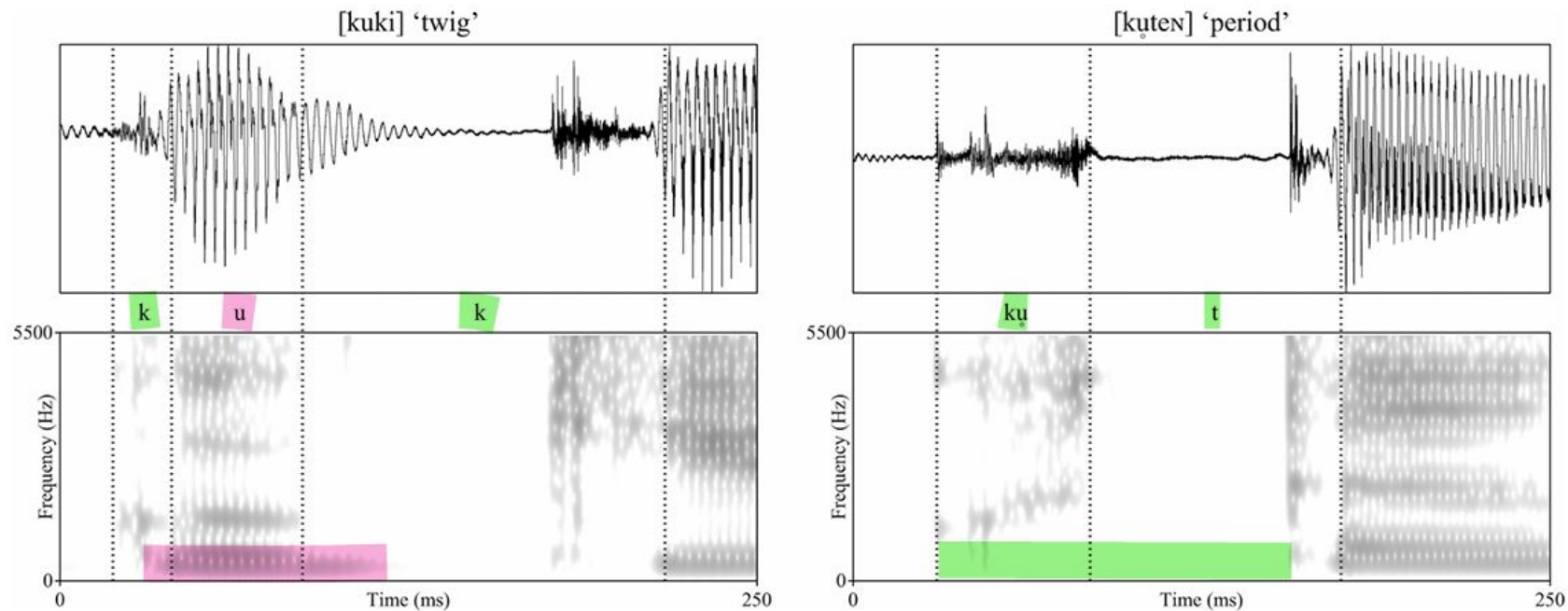


FIG. 1. Waveform and spectrogram of **voiced (left)** and **devoiced (right)** vowels in **devoicing environments**, showing landmarks for  $C_1$ , vowel, and  $C_2$  duration.

# Devoicing analysis – voicing context

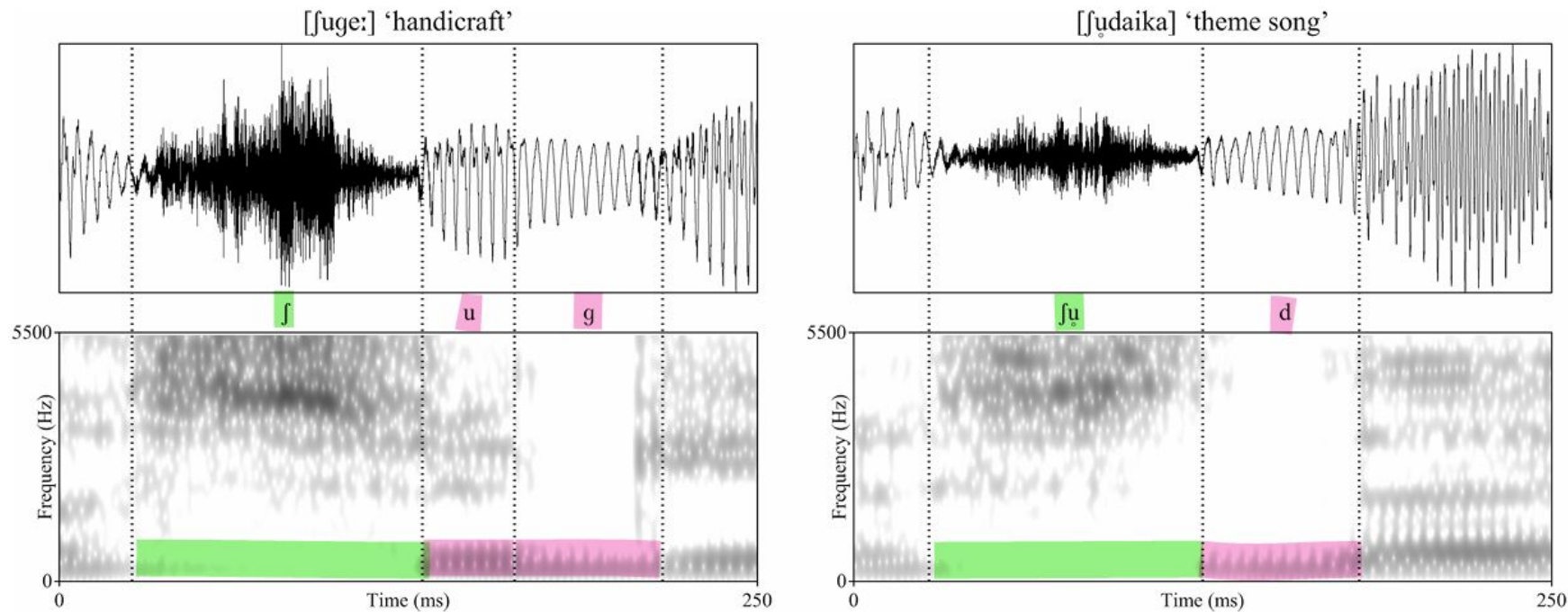


FIG. 2. Waveform and spectrogram of **voiced (left)** and **devoiced (right)** vowels in **voicing environments**, showing landmarks for C<sub>1</sub>, vowel, and C<sub>2</sub> duration.

# Duration and COG analysis

- Duration
  - [k] and [tʃ] – duration measurements excluded the silence from closure
- COG – the amplitude weighted mean of frequencies present in the signal in a 20 ms window:
  - COG1 – starting 10 ms after the beginning of C<sub>1</sub> burst/frication
  - COG2 – ending 10 ms before the end of C<sub>1</sub> burst/frication

# Results and discussion

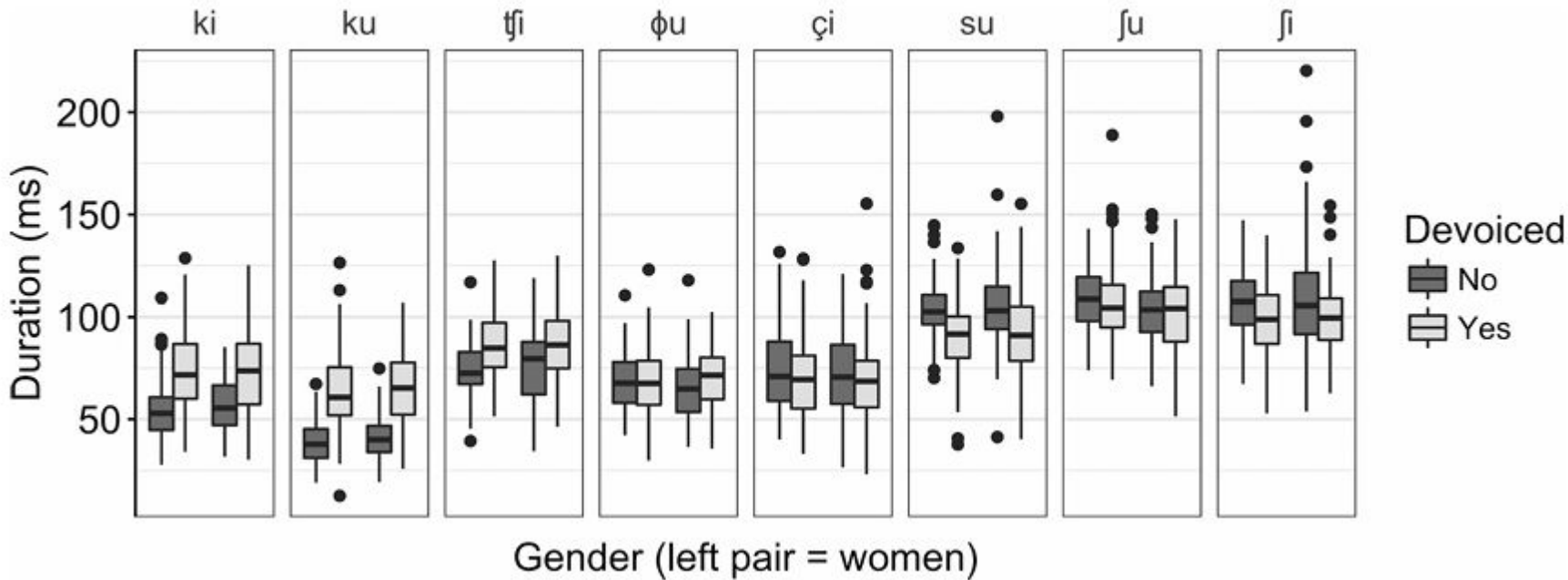
---

# Devoicing

- The difference in devoicing rates between devoicing and voicing contexts was significant ( $p < 0.001$ )
- As expected – the loss of phonation in these contexts is phonological
- Men are more likely to devoice than women ( $p = 0.018$ )

<i>Stimulus type</i>	<i>C</i> <sub>1</sub>	<i>V</i>	<i>devoicing</i>	<i>voicing</i>
low predictability	k	i	1.000	0.077
		u	0.959	0.032
	ʃ	i	1.000	0.086
		u	0.973	0.073
high predictability	ʧ	i	1.000	0.191
	ç	i	1.000	0.015
	φ	u	1.000	0.042
	s	u	1.000	0.214
<i>overall</i>			0.992	0.091

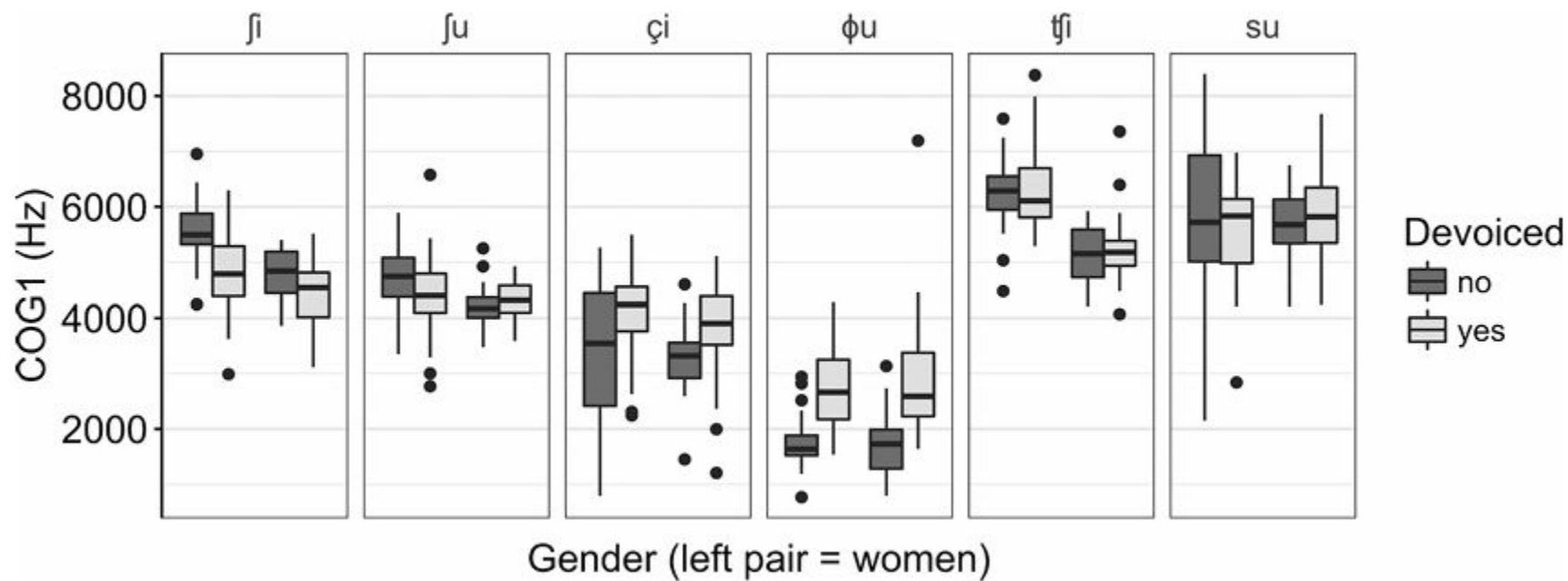
# Duration



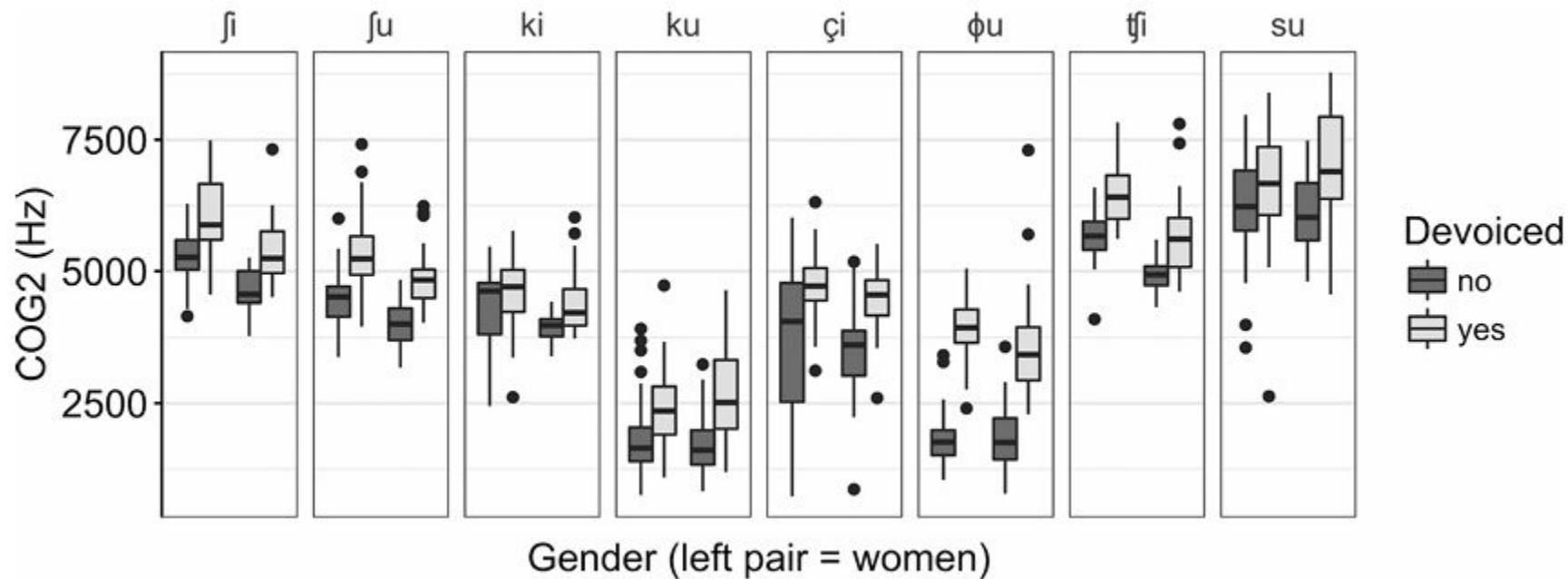
## Duration – insights

- Overall C<sub>1</sub> burst/frication durations are not different between women and men
- Devoicing has:
  - a lengthening effect on non-fricative obstruents (/ki/, /ku/, /tʃi/)
  - no effect on /ɸu/
  - a shortening effect on /çɪ/, /su/, /ʃu/, /ʃi/

# COG1



# COG2



## COG – insights

- COG1 – devoicing:
  - /ɸ/, /ç/ – has a significant raising effect for both men and women
  - /s/ – does not have a significant effect
  - /tʃ/, /ʃ/ – does not have a significant effect, lower for male speakers
- COG2 – results generally mirror COG1

# Conclusions

---

# Conclusions

- Japanese high vowel devoicing can result in complete deletion of the vowel
- The choice between deletion and unphonating is dependent on the vowel's recoverability
- Vowels are **less** likely to be deleted completely when they are **unpredictable** (after /ʃ/ and /k/)
- Speakers retain the perceptibility of a devoiced vowel even if the vowel is deleted completely (because it is highly predictable from the phonotactics)

# Bibliography

- Beckman, M., and Shoji, A. (1984). "Spectral and perceptual evidence for CVcoarticulation in devoiced /si/ and /syu/ in Japanese," *Phonetica* 41(2), 61–71.
- Boersma, P. (2009). " Cue constraints and their interactions in phonological perception and production," in *Phonology in Perception*, edited by P. Boersma and S. Hamann (Mouton de Gruyter, Berlin), pp. 55–110.
- Chitoran, I., Goldstein, L., and Byrd, D. (2002). "Gestural overlap and recoverability: Articulatory evidence from Georgian," in *Papers in Laboratory Phonology VII*, edited by N. Warner and C. Gussenhoven (Mouton de Gruyter, Berlin).
- Han, M. S. (1994). "Acoustic manifestations of mora timing in Japanese," *J. Acoust. Soc. Am.* 96(1), 73–82.
- Hayes, B. (1999). "Phonetically driven phonology: The role of Optimality Theory and inductive grounding," in *Functionalism and Formalism in Linguistics*, edited by M. Darnell, E. Moravcsik, M. Noonan, F. J. Newmeyer, and K. M. Wheatley (John Benjamins, Amsterdam), pp. 243–285.
- Hirayama, M. (2009). "Postlexical prosodic structure and vowel devoicing in Japanese," Doctoral dissertation, University of Toronto.
- Imai, T. (2010). "An emerging gender difference in Japanese vowel devoicing," in *A Reader in Sociolinguistics*, edited by D. Richard Preston and N. A. Niedzielski (Walter de Gruyter, Berlin), Vol. 219, Chap. 6, pp. 177–187.
- Ito, J., and Mester, A. (2015). "Sino-Japanese phonology," in *Handbook of Japanese Phonetics and Phonology*, edited by H. Kubozono (Mouton de Gruyter, Berlin), Chap. 7.
- Kondo, M. (1997). "Mechanisms of vowel devoicing in Japanese," Doctoral dissertation, University of Edinburgh.
- Mattingly, I. G. (1981). "Phonetic representation and speech synthesis by rule," in *The Cognitive Representation of Speech*, edited by J. Myers, J. Laver, and J. Anderson (North-Holland Publishing Company, Amsterdam), pp. 415–420.
- McCarthy, J. J. (1999). "Sympathy and phonological opacity," *Phonology* 16(3), 331–399.
- Okamoto, S. (1995). "'Tasteless' Japanese: Less 'feminine' speech among young Japanese women," in *Gender Articulated: Language and the Socially Constructed Self*, edited by K. Hall and M. Bucholtz (Routledge, NewYork), pp. 297–325.
- Shannon, C. E. (1948). "A mathematical theory of communication," *Bell System Tech. J.* 27, 379–423.
- Vance, T. J. (2008). *The Sounds of Japanese* (Cambridge University Press, NewYork).
- Varden, J. K., and Sato, T. (1996). "Devoicing of Japanese vowels by Taiwanese learners of Japanese," *Proc. Int. Conf. Spoken Lang. Process* 96(2), 618–621.
- Zsiga, E. (2000). "Phonetic alignment constraints: Consonant overlap in English and Russian," *J. Phon.* 28(1), 69–102.