3aSC21. Effects of predictability on vowel reduction

James D. Whang, Department of Linguistics, New York University, james.whang@nyu.edu

High vowel devoicing or deletion?

In standard Tokyo Japanese, unaccented high vowels /i, u/ reduce when both C1 (consonant preceding the vowel) and C2 (consonant following) are voiceless. There is disagreement regarding how reduced vowels are manifested acoustically due to different definitions and limited stimuli types.

C1 compensatory lengthening – presence of vowel signaled by extra long C1 burst/frication.

- Duration of /s, k, t/ longer when following vowel is reduced (Han 1994, Kondo 2005), but contrary results also reported (Beckman 1982, Faber & Vance 2000).

Devoicing – high vowels lose voicing but retain oral gestures, which color C1 burst/frication.

- Visible formant structures apparent in [k] burst (Varden 2010); significant and perceptible difference in mean frequency between [ʃi̥] and [ʃu̥] (Beckman & Shoji 1984).

Deletion – high vowels lose both voicing and oral gestures.

- No visible formant structures apparent in /k, t/ burst/frication (Ogasawara 2013); s-initial sequences that have undergone vowel reduction are acoustically identical to similar clusters in English (Vance 2000).

Question: **How are reduced vowels manifested acoustically and what factors affect them?**

- Is vowel reduction signaled by lengthening the preceding consonant (C1)?
- Do reduced vowels devoice (retain their oral gestures and color the preceding consonant) or delete (lose both voicing and oral gestures)?
- *Predictability* – Testing the idea proposed by Varden (2010) that when high vowels trigger allophonic variation on the preceding consonant, the allophone C1 makes the underlying vowel predictable. The vowel can thus be deleted.

Methodology

- **Stimuli**: 40 lexical tokens (20 target, 20 control) embedded in unique and meaningful sentences, read by 4 female and 4 male native speakers of Tokyo Japanese in a soundproof booth. Stimuli controlled for medium frequency (Sharoff 2008) and never occurred immediately before or after a major phrase break.
  - 10 stop-stop tokens each for target and control tokens.
  - 10 fricative-stop tokens each for target and control tokens.
- Predictability is determined by whether one or both of the high vowels /i, u/ can follow a given C1 (Varden 2010).

		i	u
Predictable	ɸ	-	✓
	s	-	✓
	ç	✓	-
Unpredictable	ʃ	✓	✓
	k	✓	✓

- C2 for target = [t, k]; for control = [b, d, g]

Acoustic measurements:

C1 duration

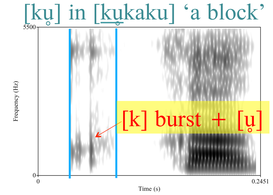
- *Reduced* tokens = no evidence of vowel voicing and formant structure after C1 burst/frication.
  - C1 measured from onset of aperiodic burst/frication energy to C2 stop closure. Since boundary between C1 and reduced vowel is unclear, measurement is technically C1+reduced vowel.
- *Unreduced* tokens = evidence of voicing after C1 burst/frication and before C2 stop closure. For control tokens, additional requirement of formant structure.
  - C1 measured from onset of aperiodic burst/frication energy to onset of voicing and/or formant structure.

Vowel coloration on C1 measured by center of gravity (COG), an amplitude weighted mean of frequencies in a signal.

- First half (COG1) and second half (COG2) of C1 measured separately to see change in degree of overlap with vowel gesture.

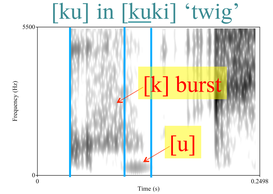
Reduced

[ku] in [kukaku] 'a block'

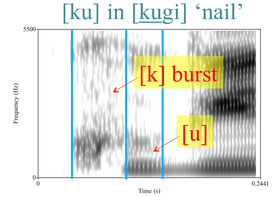


Unreduced

[ku] in [kuki] 'twig'



[ku] in [kugi] 'nail'



Results and Discussion

Mixed effects linear regression model fit to data. Results significant if $|t| > 2$ (Gelman and Hill 2007). Random intercepts and random slope for reduction by subject and word.

Duration

- **Reduction has no significant effect on C1 duration** for all consonants ($|t| < 2$), but opposite patterns emerge for [k] vs. fricatives.
  - [k] is generally *longer* when vowel is reduced (Figure 1), and significantly so in women ($t = 3.186$*) but not in men ($t = 0.572$).
  - In contrast, fricatives are consistently *shorter* when target vowel is reduced (Table 1).
  - C1 fricatives possibly forming clusters with C2 stops.

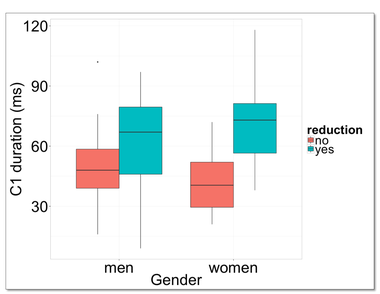


Figure 1: Duration of [k] by gender and reduction status.

	Unreduced	Reduced	*t* value
ʃ	92 (21)	86 (18)	-0.458
ɸ	60 (19)	58 (16)	0.000
s	99 (15)	82 (21)	-1.296
ç	80 (26)	64 (16)	-0.534

Table 1: Mean duration and standard deviation (in parentheses) of all fricatives in ms.

Center of Gravity

When vowel is predictable, the vowel deletes (i.e., after [ɸ, s, ç]).

COG1 is not significantly different between un/reduced tokens (Table 2).

- COG *drops* for unreduced tokens (COG2 – COG1 = negative), suggesting *increased* vowel gestural overlap in second half of C1.
- COG *rises* for reduced tokens (COG2 – COG1 = positive), suggesting *decreased* or lack of vowel gestural overlap in second half of C1.

➢ Vowel gesture unnecessary to aid recovery of reduced vowel.

When vowel is unpredictable, the vowel devoices (i.e., after [ʃ, k]).

COG significantly lower for /u/ than /i/ regardless of reduction; no significant effect of reduction within vowel type (Figures 2 & 3).

- /ʃ/: unlike fricatives in predictable cases, COG drops regardless of reduction, suggesting vowel gesture is actively retained.
- /k/: COG1 essentially same regardless of reduction, suggesting early overlap with following vowel.

➢ Retention of vowel gesture aids recovery of reduced vowel.

❖ COG results provide evidence that overlap of gestures are coordinated in order to preserve recoverability (Silverman 1995, Chitoran et al. 2002). No vowel gesture if recoverability is not an issue (i.e., vowel is predictable). If vowel is unpredictable, retain and overlap vowel gesture with preceding consonant as much as possible to aid recovery.

		COG1	COG2	ΔCOG	*t* (COG1)
ɸ	unreduced	1445 (1508)	1055 (1024)	-379 (543)	1.336
	reduced	2233 (742)	3326 (994)	1098 (1193)	
s	unreduced	7091 (1098)	5967 (1404)	-1124 (1027)	-0.745
	reduced	6310 (1521)	6733 (1004)	422 (1138)	
ç	unreduced	3081 (1195)	2830 (1404)	-251 (1021)	1.421
	reduced	3686 (692)	4363 (611)	677 (1006)	

Table 2: Mean COG and standard deviation (in parentheses) of all fricatives in Hz.

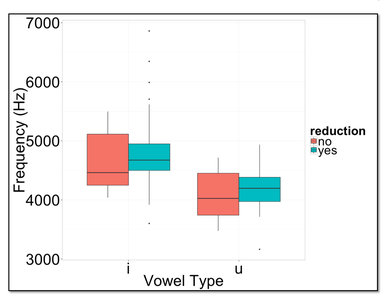


Figure 2: COG2 of [ʃ] by vowel type (t = -2.550*) and reduction (t = 0.665).

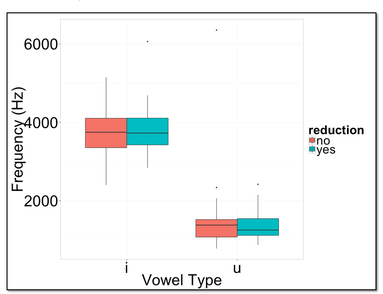


Figure 3: COG1 of [k] by vowel type (t = -10.387*) and reduction. (t = 0.215).