

Effects of recoverability on perception of illusory vowels

LSA 2016
January 10, 2016
James Whang
NYU

1

Recoverability

- Gestural overlap controlled to preserve recoverability (Chitoran et al. 2002, Silverman 1995, Whang 2014).
- Predictability: Recovery from context (phonotactics).
- Interpretability: Recovery from phonetic cues.
- Predictability and interpretability interaction in perception? (Varden 2010)

2

Japanese high vowel reduction

- Japanese listeners sensitive to high vowel coarticulation (Beckman & Shoji 1984), but not Brazilian Portuguese speakers (Dupoux et al. 2011) or German speakers (Zimmerer et al. 2013).
- High vowel reduction in Japanese.

3

Japanese high vowel reduction

- Short /i, u/ reduce in C_1VC_2 when C_1 and C_2 both voiceless.
 - /kita/ → [kᵢta] 'north'
 - /ki:ta/ → *[kᵢ:ta] 'hear (past tense)'
 - /gita:/ → *[gᵢta:] 'guitar'
 - /kata/ → *[kᵛta:] 'model/type'

4

Predictability in high vowel reduction

Summary of possible $C_1V_{[high]}$ in Japanese

		i	u
High Predictability	\widehat{ts}	—	✓
	$tʃ$	✓	u:
	ϕ	—	✓
	s	—	✓
	ζ	✓	u:
Low Predictability	p	✓	✓
	k	✓	✓
	$f(\epsilon)$	✓	✓

5

Japanese high vowel reduction

- Devoicing - only voicing is lost (Beckman & Shoji 1984).
- Deletion - vowel completely lost (Ogasawara 2013, Vance 2008).
- Devoicing vs. deletion depends on predictability (Varden 2010, Whang 2014).
- Low predictability = high interpretability (devoicing)
- High predictability = low interpretability (deletion)

6

Main Questions

- Perception affected by recoverability?
 - Predictability & interpretability inversely related?
- Illusory vowel epenthesis related to high vowel reduction in Japanese?
 - Phonotactic repair? (Dupoux et al. 1999, 2011)
 - Sensitivity to phonetic cues not controlled for.
 - Only high vowels recoverable?
 - Less CVC = CC confusion for non-reducing contexts?

7

Experiment details

Stimuli for perception experiment

<i>No-Reduce</i>	eb_ko	ez_po	eg_to	ob_ke	oz_pe	og_te
<i>Lo-Predict</i>	ep_ko	eʃ_po	ek_to	op_ke	oʃ_pe	ok_te
<i>Hi-Predict</i>	eφ_ko	es_po	eç_to	oφ_ke	os_pe	oç_te

- Participants: 29 monolingual Japanese (16 women, 13 men)
- Medial vowels (V_M): /i, u, a/
- Manipulate C_1 and V_M duration by halves.
- Identification task: <ekto>, <ekato>, <ekito>, <ekuto>

8



[ek(u)to]

ekato

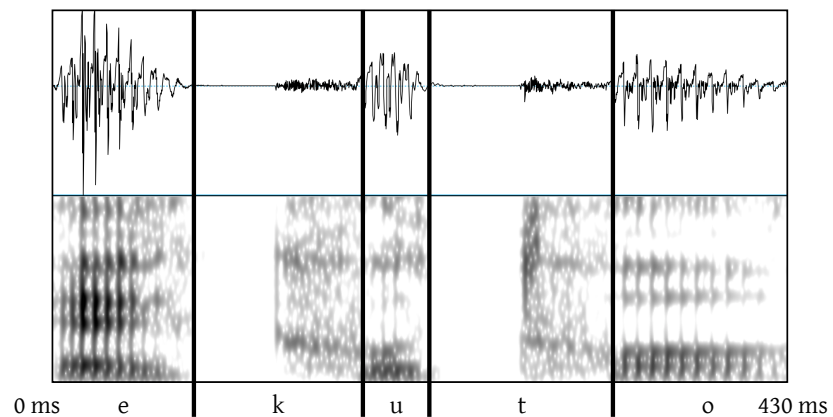
ekito

ekuto

ekto

9

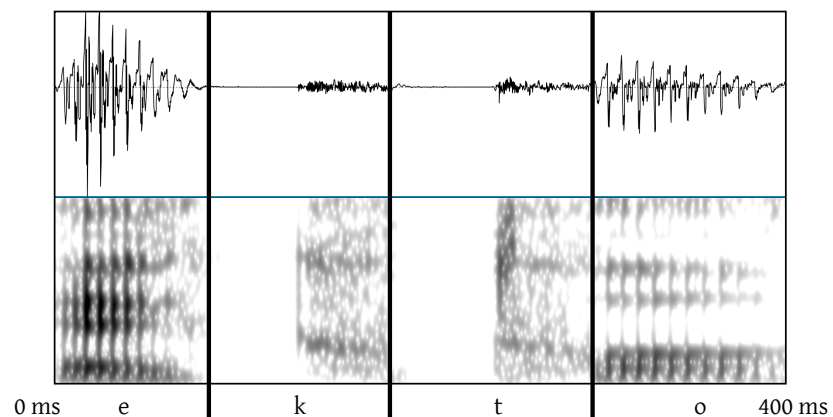
Full C_1 and V_M



Spectrogram and waveform of [ekuto] with full C_1 and V .

10

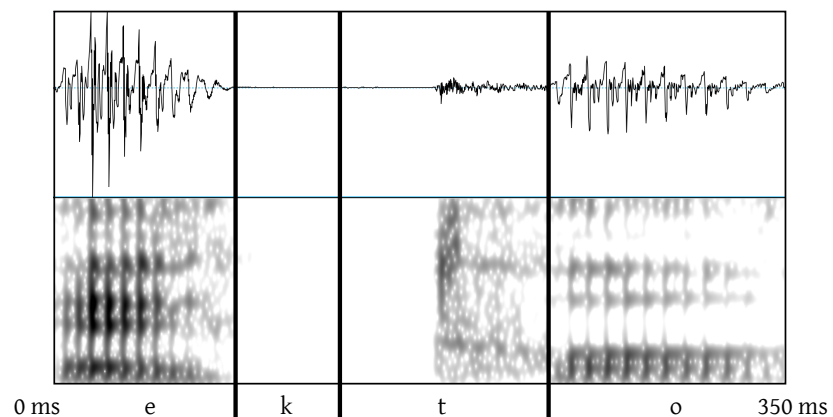
Full C_1 and no V_M



Spectrogram and waveform of [ek(u)to] with full C_1 and no V .

11

No C_1 and V_M

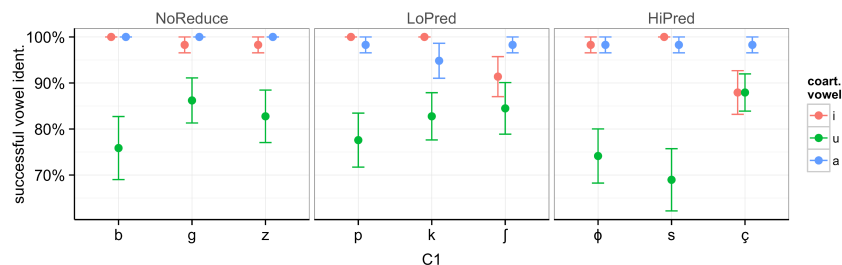


Spectrogram and waveform of [ek(u)to] with no C_1 and no V .

12

Full medial vowel

$VC_1V_MC_2V$

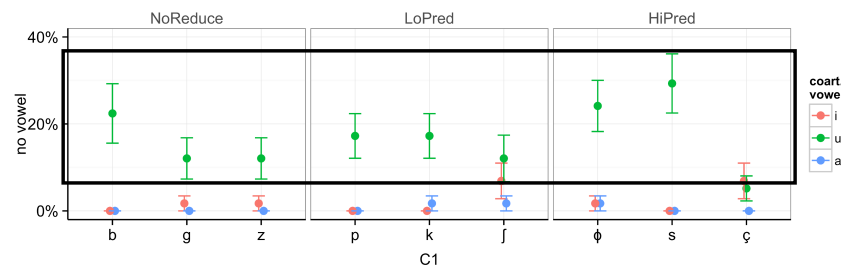


- Overall performance not different across contexts.
- [u] identification surprisingly low for all C1.
- [i] identification also low for [f, chi].

13

Full medial vowel

$VC_1V_MC_2V$

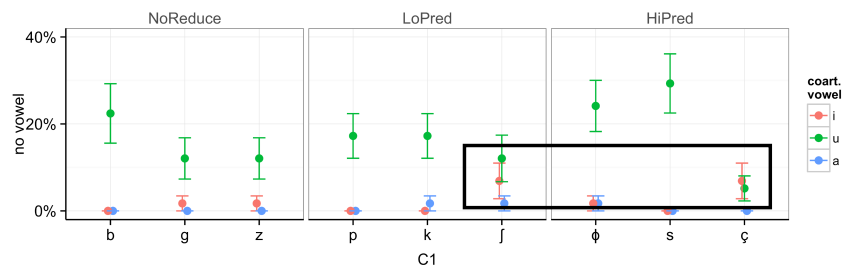


- 'no vowel' responses for [u] > 30 ms (17%).
- some [i] confusion after [f, chi] (7%).
- no [a] confusion (< 1%).

14

Full medial vowel

$VC_1V_MC_2V$

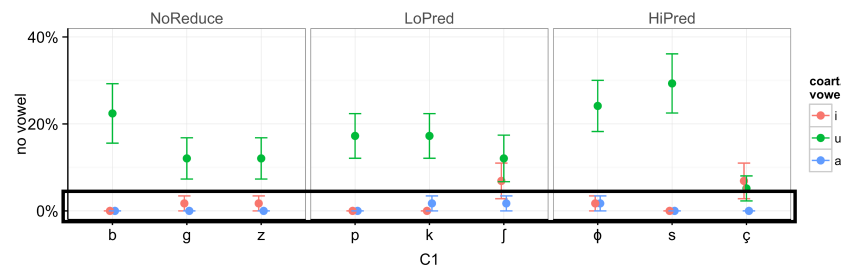


- 'no vowel' responses for [u] > 30 ms (17%).
- some [i] confusion after [f, chi] (7%).
- no [a] confusion (< 1%).

15

Full medial vowel

$VC_1V_MC_2V$



- 'no vowel' responses for [u] > 30 ms (17%).
- some [i] confusion after [f, chi] (7%).
- no [a] confusion (< 1%).

16

Naturally produced

VC₁C₂V

	NoReduce			LoPred			HiPred		
	ebko	egto	ezpo	epko	ekto	efpo	eφko	espo	eçto
a	0.14	0.02	0.03	0.10	0.02	0.00	0.00	0.00	0.00
i	0.10	0.05	0.09	0.24	0.02	0.55	0.07	0.07	0.76
u	0.34	0.43	0.50	0.29	0.59	0.26	0.60	0.60	0.14
∅	0.41	0.50	0.38	0.36	0.38	0.19	0.33	0.33	0.10

- Most 'no vowel' response in non-reducing contexts.
- Mostly report hearing a vowel.

17

Naturally produced

VC₁C₂V

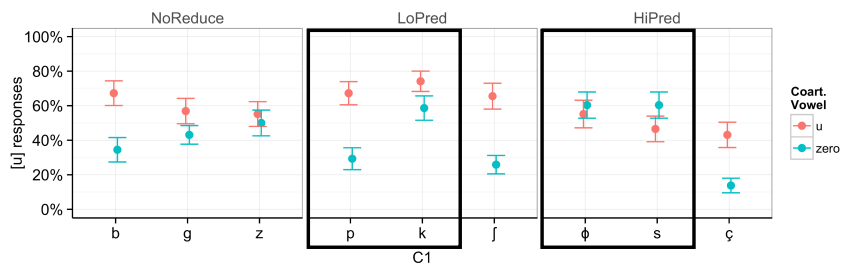
	NoReduce			LoPred			HiPred		
	ebko	egto	ezpo	epko	ekto	efpo	eφko	espo	eçto
a	0.14	0.02	0.03	0.10	0.02	0.00	0.00	0.00	0.00
i	0.10	0.05	0.09	0.24	0.02	0.55	0.07	0.07	0.76
u	0.34	0.43	0.50	0.29	0.59	0.26	0.60	0.60	0.14
∅	0.41	0.50	0.38	0.36	0.38	0.19	0.33	0.33	0.10

- Among vowel responses, [u] most common.
- Default vowel after [ʃ, ç] is [i].

18

Spliced medial vowel

VC(u)CV

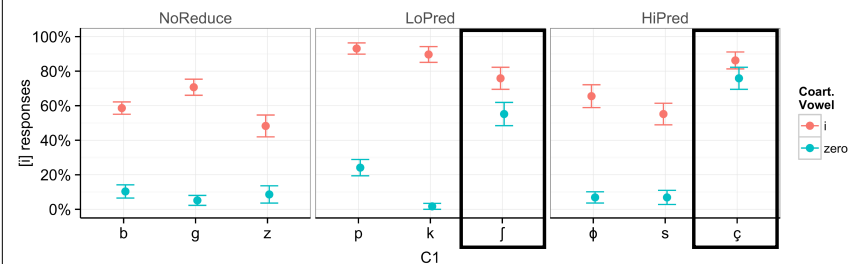


- [u] coarticulation drives up [u] responses.
- Effect most noticeable in LoPred contexts.
- [u] response essentially same in HiPred.

19

Spliced medial vowel

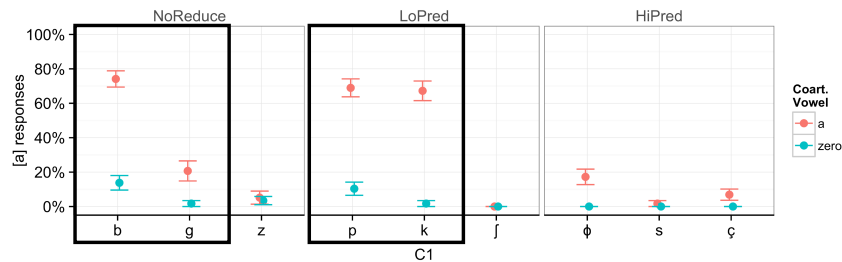
VC(i)CV



- [i] responses higher for [i]-coarticulated tokens.
- Highest accuracy in LoPred contexts.
- [i] response essentially same for [ç] but not [ʃ].

20

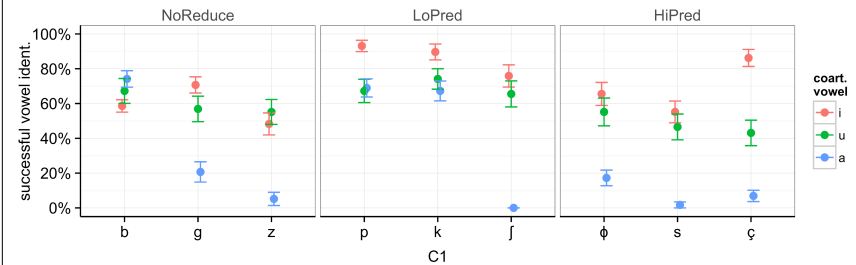
Spliced medial vowel VC(a)CV



- [a] coarticulation drives up [a] responses as well.
- Really only in stops.
- Affricates needed for comparison in HiPred.

21

Spliced tokens



- Successful identification by context:
 - LoPred (67%)
 - NoReduce (51%)
 - HiPred (42%)
- Successful identification by vowel:
 - [i] (71%)
 - [u] (59%)
 - [a] (25%)

22

Summary

- High vowels confused with 'no vowel' but not [a].
- Vowel coarticulation affects perceived vowel (least in HiPred).
- Vowel coarticulation effect most noticeable in LoPred.
- Low vowel coarticulation also seems recoverable but only in stops (perhaps due to gestural overlap difference).

23

Conclusions

- Sensitivity to phonetic cues highest in LoPred contexts.
- Bidirectional $C_1C_2 = C_1V_{[+high]}C_2$ confusion, suggests effect of high vowel reduction.
- Illusory vowel epenthesis in Japanese speakers is due in large part to hyper-sensitivity to vowel-like content in acoustic signal.
- Phonotactic violation is not the sole factor (contra. Dupoux et al. 1999, 2011).

24

Thank you!

References

- Beckman, Mary, and A. Shoji. 1984. Spectral and perceptual evidence for CV coarticulation in devoiced /si/ and /syu/ in Japanese. *Phonetica* 41:61–71.
- Dupoux, Emmanuel, Kazuhiko Kakehi, Yuki Hirose, Christophe Pallier, and Jacques Mehler. 1999. Epenthetic vowels in Japanese: a perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25:1568–1578 .
- Dupoux, Emmanuel, Erika Parlato, Sonia Frota, Yuki Hirose, and Sharon Peperkamp. 2011. Where do illusory vowels come from? *Journal of Memory and Language* 64:199–210.
- Ogasawara, N. 2013. Lexical representation of Japanese high vowel devoicing. *Language and Speech*, 56:5–22.
- Vance, T.J. 2008. *The sounds of Japanese*. New York: Cambridge University Press.
- Varden, J.K. 2010. Acoustic correlates of devoiced Japanese vowels: Velar context. *J. Eng. Am. Lit. Ling.* 125, 35–49.
- Whang, J. Under revision. Recoverability-driven gestural coordination: Acoustic evidence from Japanese high vowel reduction. *Journal of the Acoustical Society of America*.
- Zimmerer, Frank, Rei Yasuda, and Henning Reetz. 2013. Architekt or Archtekt? Perception of devoiced vowels produced by Japanese speakers of German. In *Proceedings of Interspeech*, 417–420. Lyon.