Reconciling CV phonotactics and high vowel deletion in Japanese

James Whang
The MARCS Institute for Brain, Behaviour & Development

Introduction: Japanese high vowel devoicing is a productive process where /i, u/ lose their voicing between two voiceless consonants (e.g. /suki/ → [su̥ki] 'to like'). There is disagreement over the representation of Japanese high vowel devoicing, stemming from whether the process is assumed to be in the domain of phonology or phonetics. If the process is phonological, the process must apply categorically and cannot result in a completely deleted vowel as it would violate the strong CV bias in Japanese (Tsuchida 1997). If the process is phonetic, the process should be gradient and phonotactic restrictions can be disregarded, allowing devoiced vowels to range from partial devoicing to complete deletion (Beckman 1996). The available empirical evidence suggests that both views are partially correct. Devoiced vowels do variably delete resulting in surface clusters (Shaw & Kawahara 2018a), but the process is also obligatory with the loss of at least voicing, regardless of speech rate or register (Kondo 2005). This paper proposes that there is no inherent conflict between the CV phonotactics and CC clusters that result from high vowel deletion in Japanese because phonotactic repairs and high vowel devoicing occur at different phonological levels.

Proposal: Phonetically-motivated phonological processes like Japanese high vowel devoicing have been argued to apply at different phonological levels from structural processes such as syllabification and phonotactic restrictions (Hayes 1999), which means that a simple /underlying/ vs. [surface] representation is insufficient to capture the CV phonotactics and high vowel devoicing process of Japanese. This paper, therefore, adopts the multilevel phonological representations proposed in the OT learnability literature (Boersma 2011; Tesar and Smolensky, 2000) as shown in Figure 1. The innovation in this approach is that as a late process that applies at the [overt] level, high vowel devoicing is not subject to phonotactic repairs.

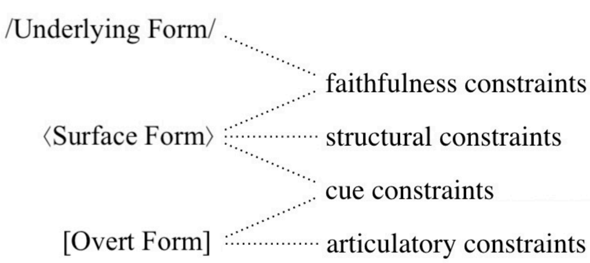

Figure 1. Phonological levels

Evidence: Phonotactic violations are repaired regardless of whether the input is an /underlying form/ (production) or [overt form] (perception). For example, many Sino-Japanese roots are underlyingly CVC. As shown in (1), the final consonant forms a geminate with the following onset if possible to satisfy CodaCondition (only allow placeless codas). Otherwise, as shown in (2) a high vowel is epenthesized to form a separate syllable to satisfy *Complex (no tautosyllabic clusters; Ito 1986).

(1) /kat+sei/ → ⟨kas.sei⟩ 'active/activated (lively+property)'
/kak+ko/ → ⟨kak.ko⟩ 'firm/determined (certain+solid)'

(2) /kat+gʲo/ → ⟨ka.t͡su.gʲo⟩ 'live fish (lively+fish)'
/kak+niN/→ ⟨ka.ku.niɴ⟩ 'confirmation (certain+acknowledge)'

The same constraints evaluate [overt] forms during perception to repair illicit consonant clusters (e.g. [ebzo] → ⟨e.bu.zo⟩), which explains why Japanese listeners have difficulty perceiving clusters accurately.

Further evidence for the separation comes from the fact that high vowel devoicing targets both underlying and epenthesized vowels, suggesting that phonotactic repairs must precede high vowel devoicing as shown in (3) below.

(3) /ku+too/ → ⟨ku.too⟩ → [ktoo] 'hard fight (difficult+battle)'
/kak+too/ → ⟨ka.ku.too⟩ → [kaktoo] 'definitive answer (certain+answer)'

Illustration: Building on observations that phonetic implementation is largely language-specific, phonetic cues that are relevant to a given language are included as part of the phonological grammar in the form of *cue constraints*, which are active in the evaluation of both surface and overt forms. Below in Table 1 is a tableau adapted from Boersma (2009) illustrating how a non-native, illegal coda is perceived by a Japanese listener. Superscript [ᵏ] indicates a [k] release burst. Because the output is a surface form, the structural constraint CODACONDITION is active as well as a number of cue constraints that help interpret acoustic cues. The constraint *⟨ ⟩ [ᵏ] penalizes a *k*-burst that is not represented in the surface form. The cue constraints *⟨o⟩ [ ] and *⟨u⟩ [ ] penalize the epenthesis of ⟨o⟩ and ⟨u⟩ in the surface form.

	[takᵏ]	CODACONDITION	*⟨ ⟩ [ᵏ]	*⟨o⟩ [ ]	*⟨u⟩ [ ]
☞	⟨taku⟩				*!
	⟨tak⟩	*!			
	⟨ta⟩		*!		
	⟨tako⟩			*!	

Table 1: Perception of illicit coda

Table 2 illustrates the high vowel devoicing process with a phonotactically conforming surface form as input and an overt form as output. Overt forms are evaluated by cue constraints and articulatory constraints. High vowel devoicing can be formalized with articulatory constraints such as *[*s.g.*][*vowel, short c.g.*][*s.g.*], which prohibits a vowel with a closed glottis gesture that is too short when between consonants with a spread glottis (voiceless) gesture. The duration of [c.g.] can also be defined, but for the current discussion, it suffices to simply point out that high vowels are most likely to devoice due to their inherent shortness. Whether the devoiced or deleted form is produced depends on the respective rankings of *⟨u⟩ [ ] and the articulatory constraint.

	⟨ku.too⟩	*⟨k⟩ [ ]	*[*s.g.*][*vowel, short c.g.*][*s.g.*]	*⟨u⟩ [ ]
☞	[kᵏu̥tᵗoo]			
☞	[kᵏtᵗoo]			*!
	[utᵗoo]	*!		
	[kᵏutᵗoo]		*!	

Table 2: Production of devoiceable high vowel

Conclusion: This paper demonstrates that the seemingly contradictory treatment of high vowels in Japanese can be reconciled by employing a more sophisticated representation of phonological levels. CV preference is due to structural constraints that repair violations at the surface level. High vowel devoicing is a late overt level process that applies after structural evaluation has occurred. The proposed theoretical account also predicts that the onset consonants that are stranded due to high vowel deletion should not resyllabify since syllabification occurs before devoicing/deletion. This prediction is supported by a recent study by Shaw and Kawahara (2018b), who showed that stranded onset consonants do not show c-center effects that would be expected in the case of resyllabification.

Selected references: Boersma (2011). A programme for bidirectional phonology and phonetics and their acquisition and evolution. *Bidirectional optimality theory.* **Ito** (1986). *Syllable Theory in Prosodic Phonology*. **Kondo** (2005). Syllable structure and its acoustic effects on vowels in devoicing. *Voicing in Japanese.* **Shaw & Kawahara** (2018a). The lingual articulation of devoiced /u/ in Tokyo Japanese. *J. Phon.* **Shaw & Kawahara** (2018b). Consequences of high vowel deletion for syllabification in Japanese. *AMP 2017.* **Tesar & Smolensky** (2000). *Learnability in Optimality Theory*. **Tsuchida** (1997). *Phonetics and phonology of Japanese vowel devoicing*.