Persistence of prosody

Shigeto Kawahara & Jason A. Shaw
Keio University & Yale University

Preamble

In October 2016, at a workshop held at the National Institute for Japanese Language and Linguistics (NINJAL),[1] Junko and Armin presented a talk in which they argued against Kubozono's (1999; 2003) proposal that VVN sequences in Japanese are syllabified as two separate syllables (V.VN) (Ito & Mester, 2016a). One of their arguments involved the consequences for VNC sequences (e.g. /beruriŋko/ 'people from Berlin'); in particular the proposal would require positing syllables headed by a nasal (i.e. V.NC, or [be.ru.ri.ŋk.ko]). They argue that syllables headed by a nasal segment are "questionable syllable types", at least in the context of Japanese phonology. We are happy to dedicate this paper to Junko and Armin, in which we argue that Japanese has syllables headed by a fricative, and possibly those headed by an affricate.

1 Introduction

Segments or prosody, which comes first? This question has been an important topic in phonetic and phonological theories. A classic view in generative phonology is that input segments are given first, and syllables and higher prosodic structures are built over segments according to universal and language-specific algorithms (Clements & Keyser 1983; Ito 1986; Kahn 1976; Steriade 1982 and subsequent research). An almost standard assumption in this line of research that syllabification does not exist in the underlying representation (Blevins, 1995; Clements, 1986; Hayes, 1989), and this assumption reflects the view that segments come before prosody.[2] However, there are also proposals to the effect that prosodic templates are given first, and segments are "filled in" later; such is the case for patterns of prosodic morphology, such as reduplication and truncation (Ito, 1990; Levin, 1985; Marantz, 1982; McCarthy, 1981; McCarthy & Prince, 1986, 1990; Mester, 1990). Compensatory lengthening, in which segments are lengthened to fill "already-existing" prosodic positions (Hayes, 1989; Kavitskaya, 2002; Wetzels & Sezer, 1986), also instantiates a case in which prosody comes first. Thus, the question of which comes first—segments or prosody—does not seem to have a simple answer in phonological theorization.

Optimality Theory (Prince & Smolensky, 1993/2004) provided a third possibility—segments and prosodic structures are built simultaneously, and some explicit arguments are made for "parallel" evaluation

---

[1] "Syllables and Prosody" which Shigeto had proudly co-organized with Junko.

[2] Key evidence is the observation that no languages seem to use different syllabification patterns to signal lexical contrasts (though see Elfner 2006 for a potential counterexample). In Optimality Theory (Prince & Smolensky, 1993/2004), this lack of contrast can be accounted for by postulating that there are no faithfulness constraints that protect underlying syllabification (Kirchner, 1997; McCarthy, 2003). Given the Richness of the Base (Prince & Smolensky, 1993/2004; Smolensky, 1996), inputs should not be prohibited from having syllable structure, so this assumption about the lack of syllabification in underlying representations is much weakened, if not entirely abandoned, in Optimality Theory.

of segments and prosodic structures (Adler & Zymet, 2017; Anttila & Shapiro, 2017; Rosenthall, 1997; Prince & Smolensky, 1993/2004). Generally, due to parallel evaluation of output wellformedness, Optimality Theory rendered moot the question of "which comes first." The question does not even arise because everything happens all at once. However, recent proposals to incorporate derivation back into Optimality Theory (e.g. McCarthy 2007, 2010) brought this question back on the table—see for example a debate between Pruitt (2010) and Hyde (2012) about whether footing should occur derivationally or in parallel. McCarthy (2008) for example argues that footing needs to precede syncope in some languages, and that this analysis is possible only in a derivational version of Optimality Theory (though cf. Kager 1997). Thus the question of the derivational relationship between segments and prosodic structure—including the very general question of whether derivation exists at all in the phonological component of grammar—is still a matter of debate in phonological theory.

A similar question has been addressed in the context of speech production. There is a large body of literature suggesting that prosodic information is planned prior to phonetic specification of segments. In tip-of-the-tongue phenomena, for example, there are cases in which speakers can recall the stress patterns— hence the prosodic structures—of the words in question, even when the segments cannot be recalled (Brown & MacNeill, 1966). Cutler (1980) analyzes a corpus of speech errors in English and points out that "omission or addition of a syllable can be caused by an initial error involving the misplacement of stress" (p. 68), and consequently suggests that "lexical stress errors arise at a fairly early level in the production process" (p. 71).

In the modular feed-forward model of speech production planning developed in Roelofs (1997) and Levelt et al. (1999), prosodic templates including syllable counts and lexical stress position are stored in lexical entries independently from the segments of a lexical item. This aspect of the speech production model makes it theoretically possible to retrieve word prosody without segmental content. The architecture of the model is motivated as well by the implicit form priming paradigm, in which shared prosody across words, including stress position, facilitates lexical retrieval (Roelofs & Meyers, 1998). However, in this model, segmental and prosodic templates of words are merged at the level of the prosodic word, a stage of planning followed by phonetic encoding and finally construction of higher level prosodic structure. The stages of the model have been criticized for not specifying enough prosodic structure prior to phonetic encoding (Keating, 2003; Keating & Shattuck-Hufnagel, 2002; Shattuck-Hufnagel, 2006). The basic argument is that prosody must be available early in the speech planning process so that it can condition phonetic form. In other words, in speech production, it is prosody first (Shattuck-Hufnagel, 2006). More recent work has identified a possible neural basis for dissociation between segments and prosodic organization. Long et al. (2016) found that perturbation of normal brain function (through focal cooling) could selectively influence speech timing or segmental content, depending on the brain region targeted. To the extent that segments dictate articulatory goals and prosody conditions timing, this result provides another converging line of evidence for, at the least, a dissociation between segmental and prosodic planning.

1.1 *Consequences of vowel deletion for syllabification* In this paper we would like to address the general issue of the relationship between segments and prosody by examining the consequences of vowel deletion for syllabification. Our main empirical focus is Japanese, but before we present our analysis of Japanese, we start with a brief cross-linguistic examination to put our analysis in a broader perspective. Given a $C_1V_1C_2V_2$ sequence, when V_1 deletes, we can conceive of two outcomes regarding how C_1 is syllabified: (1) C_1 is resyllabified with a surrounding vowel, or (2) C_1 maintains its syllabicity. Both patterns have been claimed to be attested in the previous literature, as summarized in (1)-(9). Forms on the left are those with vowels (vowel present); forms on the right are those without vowels (vowel absent). Syllable and foot boundaries are shown only where relevant.

(1)   Resyllabification: Latvian (Karins, 1995: 19)

	vowel present	vowel absent	
a.	splig.t**a**s au.ras	spligt.sau.ras	'dazzling auras'
b.	spilg.t**i** a:.buo.li	spilg.ta:.buo.li	'dazzling apples'

(2)   Resyllabification: Leti (Hume, 1997)

	vowel present	vowel absent
a.	lop**u** do	lop.dʷo

(3)   Syllabicity maintenance: English (Kaisse & Shaw, 1985: 6)

	vowel present	vowel absent
a.	sə.pʰort	s.pʰort
b.	pə.tʰɛɪ.rɔʊ	p.tʰɛɪrɔʊ
c.	tə.li.dɔʊ	t.lidɔʊ

(4)   Syllabicity maintenance: French (Rialland, 1986)

	vowel present	vowel absent	
a.	də rôle	d.rôle	'some role'
b.	baʁə tʁuve	ba.ʁ.tʁuve	'stocking found again'

(5)   Syllabicity maintenance: Lushootseed (Urbanczyk, 1996: 119)

	vowel present	vowel absent	
a.	qəsi	q.siʔ	'uncle'
b.	c'ək'apaʔ	c'.kapaʔ	'rose hips'

(6)   Syllabicity maintenance: Triqui (p.c. Christian DiCanio)

	vowel present	vowel absent	
a.	n**i**^{1}jah^3	j^1.jah^3	'when'
b.	n**a**^{2}nah^2	n^2.nah^2	'slowly'

(7)   Syllabicity maintenance: Carib (Kager 1997 based on Abbott 1991)

	vowel present	vowel absent	
a.	peʔm**a**ra	(pèʔ)(m.rá:)	'free'
b.	seepɔrɔ	(sèe)(p.rɔ́:)	'alone'

(8)   Syllabicity maintenance: Odawa (Bowers 2015 based on Rhodes 1985)

	vowel present	vowel absent	
a.	mʌ kɪzɪn	(m.ˈkɪ)(ˈzɪn)	'shoe'

(9)   Syllabicity maintenance: Québec French (Garcia et al., 2016)

	vowel present	vowel absent	
a.	al**i**mãtasjɔ̃	a(l.mã)(ta.sjɔ̃)	'nourishment'
b.	kɔbine	kɔ(b.ne)	'to combine'

From the perspective of cross-linguistic markedness, the cases of resyllabification, as in Latvian and Leti, seem more natural; syllables headed by a vowel are less marked than syllables headed by a consonant. We are thus more interested in alleged cases in which consonants maintain their syllabicity after vowel deletion, so let us examine each case in greater detail. Especially, since syllables are often thought of as being built around high sonority segments (Dell & Elmedlaoui, 1985; Selkirk, 1982, 1984; Steriade, 1982), it is worth considering the strength of the evidence for each analysis positing consonantal syllables.

For English, it seems reasonable to posit a syllable boundary between the two word-initial consonants after the schwa is deleted; as for (a, b), the second consonants are aspirated, a hallmark of syllable-initial consonants in English; for (c), English does not allow [tl] clusters syllable-initially (Kahn, 1976; Massaro & Cohen, 1983; Moreton, 2002). We thus seem to have good evidence to consider that resyllabification does not occur after schwa deletion in English. However, Davidson (2006) points out that it is possible—and even likely—that schwa "deletion" in English does not involve phonological deletion, but instead that it is phonetic reduction. In that sense, these schwas in English are not deleted phonologically, and therefore, it may not be necessary to posit consonantal syllables in English.[3]

For French, Barnes & Kavitskaya (2003) summarize Rialland's (1986) argument as follows: "[she] observed a curious fact concerning certain instances of deletion of French schwa. Specifically, she noted that the preceding consonant, in non-postpausal contexts ostensibly resyllabified as a coda, nonetheless appears in spectrograms to retain much of the phonetic character of its corresponding onset variant, and not to lengthen the preceding vowel, as it would be expected to do were it in fact in the coda (p.41)." To account for this observation, Rialland (1986) posits an empty vocalic timing slot after schwa deletion in French, effectively arguing for consonantal syllables in French. Like Rialland (1986), Fougeron & Steriade (1997) found that consonant clusters created via schwa deletion (e.g. d'rôle 'some role') and underlying consonant clusters (drôle 'funny') behave differently. Fougeron & Steriade (1997) and Steriade (2000), however, argue against Rialland's interpretation, based on the observation that vowel deletion in French does reduce the number of syllable counts in poetry reading, suggesting instead that [d] in d'rôle keeps its underlying articulatory specification as a prevocalic consonant via phonetic analogy; crucially, however, they argue that these consonants are nevertheless resyllabified. In addition, there is a debate about whether schwas in French are entirely deleted or merely reduced (see Bürki et al. 2007, 2011 for evidence that supports the deletion view). Overall, the existence of consonantal syllables is debatable in French.

Su Urbanczyk (p.c.) informed us that the primary reason to posit "consonantal syllables" in Lushootseed is because "there is no evidence for obstruent-obstruent complex onsets in the language, so it is unlikely that [the consonant cluster] would form a complex onset." This type of argument is recurrent when examining patterns like those in (1)-(9), and came up for the analysis of English above, and will become relevant for the case of Japanese that we will discuss in detail below. For the case of Lushootseed, there remains a question of whether vowels are entirely deleted, or whether they are merely devoiced. Unfortunately, phonetic data which would allow us to address this issue in Lushootseed is currently unavailable.

The Triqui pattern was brought to our attention by Christian DiCanio (see also DiCanio 2012, 2014). The data exemplify a process of pre-tonic (i.e. penultimate) vowel syncope (DiCanio, 2012), which results in word-initial geminates. When the resulting geminates are sonorant, DiCanio suggests that they are separated by a syllable boundary, as there is tone movement across the sonorant geminates. He is not confident, however, that vowel deletion is complete in Triqui; apparent "deletion" may alternatively involve (heavy) reduction. He also informed us that when C_2 is a stop and the resulting geminate is a stop geminate (e.g. /ni^3.tah^2/ → [ttah^{23}] 'NEG.exist'). Evidence that the form resulting from vowel deletion remains disyllabic, for the case of initial geminate stops, is not currently available.

Kager (1997) makes an interesting set of arguments for the proposal that vowel deletion in Carib preserves syllable structure in the output, and maintains that Carib does not show evidence for syllabification after vowel deletion. Kager (1997) states that "[b]oth lengthening and vowel reduction are cross-linguistically common processes in iambic languages, increasing the durational differences which are inherent to the iamb: *a quantitatively unbalanced rhythm unit* (Hayes, 1995) of a light plus a heavy syllable. From a typological perspective some analysis is preferrable that expresses this connection between foot type and reduction. But then vowel reduction must crucially preserve the weak syllable in the iamb as a degenerate syllable, containing a nucleus that is void of vocalic features (emphasis in the original)." If Kager is correct, then this is a case of "persistence of prosody"—a vowel is deleted, but its rhythmic structure is maintained. While our analysis of Japanese developed below in detail is in a very similar spirit with that of Kager, Kager (1997) also

---

[3] Kawahara (2002) made a general observation that marked structures that are otherwise not tolerated in the language can be produced as a result of an optional phonological process, like vowel deletion, the observation which he dubbed "the emergence of the marked". Kawahara (2002) did not examine the issue of whether these optional processes are indeed phonological, and hence it is important to address, for example, whether vowels are indeed deleted phonologically, rather than phonetically reduced. See below for more on this issue of establishing whether vowels are deleted phonologically (i.e. categorically).

emphasizes that vowel deletion is *optional* and *gradient*. A question thus remains whether vowel deletion in Carib can be considered as complete deletion, or whether it is merely heavy reduction.

For Odawa, there is evidence for vowel deletion, at least diachronically, as recent surveys have shown that speakers have lost vowel alternations once conditioned by rhythmic syncope in favor non-alternating stems that exclude the vowel (Bowers, 2018). In this case, as well, it seems like higher level prosody has been preserved. Despite other fairly dramatic restructuring of stems and morphological inflection, the pattern of stressed vowels indicative of higher level prosodic structure persists. Indirect evidence against the resyllabification of consonants comes from phonotactic evidence elsewhere in the synchronic grammar for active avoidance of complex onsets (Bowers, 2015).

Finally, there are many cases of rhythmic syncope in Québec French by Garcia et al. (2016), who, building on Verluyten (1982) show that this deletion is conditioned by a rhythmic iambic requirement, just like in Carib and Odawa.

What is emerging from our brief cross-linguistic survey is that in order to establish the existence of consonantal syllables, two things need to be shown: (1) vowels are entirely deleted, not merely reduced or devoiced (2) consonants are not resyllabified. In this paper, we intend to establish both of these types of evidence for Japanese.

1.2 *The case of Japanese* Japanese is well-known as a language without consonant clusters, allowing only homorganic nasal-consonant clusters and geminates (Ito, 1986, 1989). In fact, not only does Japanese have no words with non-homorganic consonant clusters, Japanese speakers resort to epenthesis when they borrow words with consonant clusters from other languages; for example, the English word *strike* is pronounced as [s**u**to**raik**u] when borrowed into Japanese, in which the original, monosyllabic word becomes a four-syllable word with three epenthetic vowels (Katayama, 1998). The German last name *Wurmbrand* is borrowed as [ur**umubu**rand**o**], which can be argued to be different from the original pronunciation. This phonotactic restriction is claimed to condition perceptual epenthesis as well—Japanese listeners report hearing vowels between non-homorganic consonant clusters (Dupoux et al., 1999, 2011; Dehaene-Lambertz et al., 2000). Moreover, the Japanese orthographic system is organized in such a way that each letter represents a combination of a consonant and a vowel; i.e., there is no character that exclusively represents an onset consonant. All of these observations lead to the oft-stated characterization that "Japanese is a strict CV-language".

However, Japanese is also known to devoice high vowels between two voiceless obstruents and after a voiceless consonant word-finally, which results in apparent consonant clusters and word-final consonants (e.g. [ɸu̥soku̥] or [ɸsok] 'shortage'). Some researchers argue that these high vowels are simply devoiced—not deleted—and therefore, Japanese does not have consonant clusters after all (Faber & Vance, 2010; Jun & Beckman, 1993; Kawahara, 2015a; Sawashima, 1971). Other researchers argue that acoustically, there is no evidence for the presence of vowels at all; they therefore conclude that these vowels are entirely deleted (Beckman, 1982, 1996; Beckman & Shoji, 1984).[4] Beckman & Shoji (1984), for example, state that "[w]hen the waveform of a devoiced syllable is examined...neither its spectral nor its temporal structure indicates the presence of a voiceless vowel (p.63)." Beckman (1982) states that "devoicing" is a better term psychologically, but physically "[the term] 'deletion' is more correct, since there is generally no spectral evidence for a voiceless vowel" (p. 118). These studies are often limited in the sense that they rely on acoustic information to infer whether there remains an articulatory target for voiceless vowels or not— we independently know, however, that inferring articulation from acoustics is not always straightforward, especially when it comes to detecting the presence of a vowel (e.g. Davidson & Stone 2004; Davidson 2005; Shaw & Kawahara 2018a). The acoustic consequences of vocalic gestures can be rendered inaudible due to gestural overlap of surrounding voiceless consonants (Jun & Beckman, 1993; Jun et al., 1998), and conversely, vowel-like acoustics can be observed, even without intended vocalic gestures, when consonantal gestures are not sufficiently not overlapped (Davidson & Stone, 2004; Davidson, 2005; Hall, 2006).

To address the issue of whether high devoiced vowels in Japanese are deleted or not in a way that is more direct than inference from acoustic data, a recent articulatory study by Shaw & Kawahara (2018c) used ElectroMagnetic Articulography (EMA) to address this issue—mere devoicing vs. wholesale deletion—by

---

[4] Yet others argue that vowels are merely devoiced in some environments and deleted altogether in other environments (Kawakami, 1977; Maekawa, 1989; Whang, 2017, 2018).

examining whether the devoiced vowels retain their lingual articulation.[5] They found that at least some devoiced tokens lack vowel height targets altogether, suggesting that these high vowels are not merely devoiced but entirely deleted (see Figure 6; see also Figure 2 for relevant EPG data). They also found that those tokens that lack vowel height targets show patterns of temporal variation consistent with consonant-to-consonant (C-C) coordination. That is, the flanking consonants appear to be timed directly to each other instead of to an intervening vowel, i.e., consonant-to-vowel (C-V) coordination, providing further evidence that there is no vowel in the surface representation of these tokens. These results mean that Japanese, as a consequence of high vowel deletion, has consonant clusters (e.g. [ɸsoku]), contrary to the "CV-language" characterization often given to Japanese.

Based on this recent result reported in Shaw & Kawahara (2018c), this paper addresses how such consonant clusters, arising from high vowel deletion, are syllabified. We compare two specific hypotheses regarding this question, (1) the resyllabification hypothesis and (2) the consonantal syllable hypothesis, as anticipated in section 1.1, and present evidence for the consonantal syllable hypothesis. Our argumentation is based on two kinds of evidence. The first is a phonological consideration (section 3); we show that phonological processes that are sensitive to syllable structure, such as prosodic truncation and pitch accent placement, are unaltered by high vowel deletion. The other one is a phonetic consideration (section 4); patterns of temporal stability in speech production are inconsistent with the resyllabification hypothesis. In addition to addressing a specific question in Japanese phonology, our results bear on more general theoretical issues, including how different syllable structures manifest themselves in articulatory timing patterns (Browman & Goldstein, 1988; Byrd, 1995; Hermes et al., 2013, 2017; Marin, 2013; Marin & Pouplier, 2010; Shaw & Gafos, 2015; Shaw et al., 2009), and the independence of prosodic and segmental levels of representation, as reviewed at the beginning of this paper. The convergence of the phonetic and phonological evidence also bolsters the claim that syllable structure corresponds to characteristic patterns of gestural timing in speech. Our case study also highlights the importance of integrating theoretical insights with phonetic experimentation.
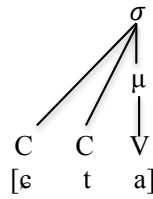
2   The two hypotheses examined in the current study

We are not the first to consider the question of how consonant clusters resulting from high vowel deletion are syllabified in Japanese, and there are two competing hypotheses in the literature. These two hypotheses are illustrated in Figure 1. The first hypothesis (H1), shown on the left side of Figure 1, is that the consonant that preceded the deleted high vowel is resyllabified into the following syllable, resulting in a complex syllable onset. Kondo (1997) argues for this view based on the observation that devoicing of two consecutive vowels is often prohibited (for which see a recent study by Nielsen 2015 and references cited therein). On Kondo's account, consecutive vowel devoicing is blocked by a constraint against tri-consonantal onsets (*CCC). This constraint can only function to block consecutive devoicing if the devoiced vowels are also deleted.

Matsui (2017) on the other hand argues that it is possible for Japanese to have consonantal syllables, as in the right side of Figure 1. His argument is primarily based on linguo-palatal patterns obtained using ElectroPalatoGraphy (EPG). He found that the pattern of lingual contact typically observed for Japanese /u/ is absent in devoiced contexts, as shown in Figure 2, implying that devoiced [u̥] is actually deleted. Moreover, when devoiced /u/ is preceded by /s/, the linguo-palatal contact pattern characteristic of /s/, which is likely to be produced with tongue groove, extends temporally throughout the syllable (Figure 3). Thus, in terms of linguo-palatal contact, it appears that /u/ is replaced by a consonant. Matsui (2017) discusses this result in the context of the C/D model of articulation (Fujimura, 2000; Fujimura & Williams, 2015), which crucially assumes that a syllable can remain even after high vowel deletion.

---

[5]   Part of the complexity of assessing deletion based on measurement is in choosing which signal to measure, as there are many relevant options, including the neural motor control signal, the activation of muscles, individually or in ensemble, the movement of the articulators, the resulting acoustic signal, or the auditory response within the cochlea or along the auditory nerve.

Hypothesis 1 (H1): Resyllabification   Hypothesis 2 (H2): Consonantal syllable
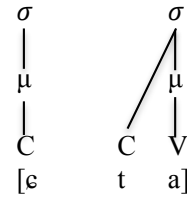
Figure 1: Two hypotheses regarding the syllabification of consonant clusters created via high vowel deletion, as in /ɕuta/ → [ɕta] .
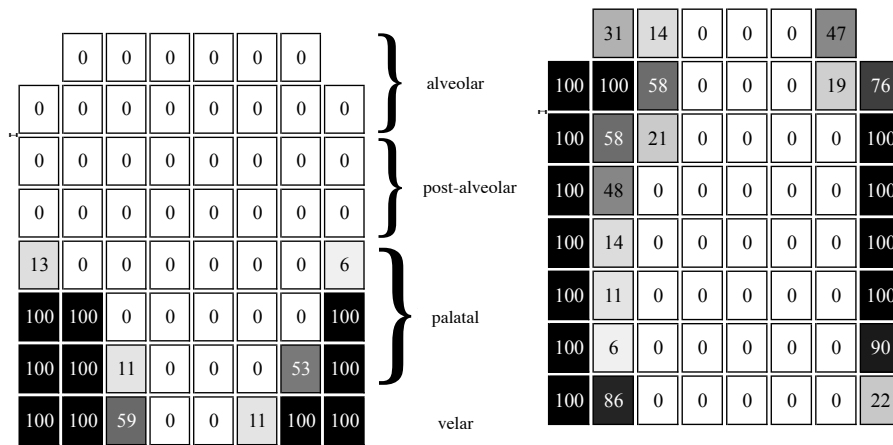
Figure 2: The linguo-palatal patterns of voiced [u] (left) and devoiced [u̥] in [su̥] (right). Adopted from Matsui (2017).
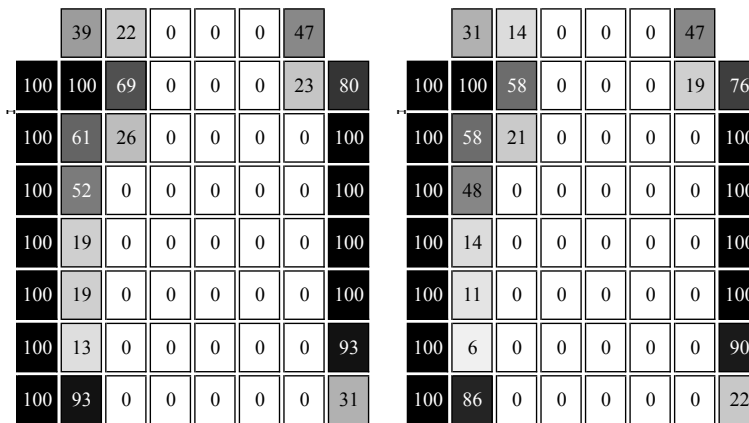
Figure 3: The linguo-palatal patterns of the devoiced syllable, [su̥]. The left panel is the first half of the syllable; the right panel is the second half of the syllable. We observe that lateral constriction for [s], likely produced with tongue groove, persists throughout the syllable. Adopted from Matsui (2017).

This paper provides further evidence for H2, drawing on a confluence of phonological and phonetic evidence.

3   Phonological considerations

We begin with phonological considerations that favor the consonantal syllable hypothesis (H2). As observed by Tsuchida (1997) and Kawahara (2015a), devoiced vowels count toward the bi-moraic requirement of some morphophonological processes. Japanese has many word formation processes that are based on a bimoraic foot (Ito, 1990; Ito & Mester, 1992/2003, 2015; Mester, 1990; Poser, 1990) and devoiced vowels count toward satisfying this requirement. The general patterns are described in (10)-(12) (the data are based on the previous works cited above, with some examples added by the first author).

(10)   loanword truncation

     a.   [demonsu̥toreeɕoɴ] → [demo] 'demonstration'
     b.   [rokeeɕoɴ] → [roke] 'location'
     c.   [rihaasaru] → [riha] 'rehearsal'
     d.   [sureddo] → [sure] 'thread'
     e.   [raboratorii] → [rabo] 'lab'
     f.   [opereeɕoɴ] → [ope] 'operation'
     g.   [ookesu̥tora] → [oke] 'orchestra'

(11)   hypocoristic formation ([-tɕaɴ] is an optional hypocoristic suffix)

     a.   [tomoko] → [tomo(-tɕaɴ)] (personal name)
     b.   [sumiko] → [sumi(-tɕaɴ)] (personal name)
     c.   [mariko] → [mari(-tɕaɴ)] (personal name)
     d.   [wasaburoo] → [wasa(-tɕaɴ)] (personal name)
     e.   [aamiɴ] → [aa(-tɕaɴ)] (personal name)
     f.   [dʑuŋko] → [dʑuɴ(-tɕaɴ)] (personal name)

(12)   Mimetics

     a.   [buru-buru] 'shivering'
     b.   [doɴ-doɴ] 'stomping'
     c.   [pasa-pasa] 'dry'
     d.   [kira-kira] 'twinkle'
     e.   [pojo-pojo] 'bouncy'
     f.   [riɴ-riɴ] 'ringing'

The following data in (13)-(15) show that devoiced vowels count toward bimoraic template patterns:

(13)   loanword truncation

     a.   [su̥toraiki̥] → [su̥to] 'strike'
     b.   [ripurai] → [ripu̥] 'reply'
     c.   [hi̥su̥terii] → [hi̥su̥] 'Hysterie (German)'
     d.   [moruhine] → [mohi̥] 'morphine'
     e.   [su̥kuriiɴ-ɕotto]   [su̥ku̥-ɕo] 'screenshot'

(14)   hypocoristic formation

     a.   [kumiko] → [ku̥ko(-tɕaɴ)] (personal name)
     b.   [tɕi̥kako] → [tɕi̥ka(-tɕaɴ)] (personal name)
     c.   [satɕi̥ko] → [satɕi̥(-tɕaɴ)] (personal name)
     d.   [satsu̥ki̥] → [satsu̥(-tɕaɴ)] (personal name)
     e.   [akira] → [aki̥(-tɕaɴ)] (personal name)

(15)   mimetics

     a.   [ɸu̥ka-ɸu̥ka] 'fluffy'
     b.   [su̥ka-su̥ka] 'empty'
     c.   [ɕi̥to-ɕi̥to] 'rainy'

      d.   [suku̥-suku̥] 'growing steadily'
      e.   [çiku̥-çiku̥] 'wining'
      f.   [pitçi̥ -pitçi̥] 'stretched'

The patterns in (13)-(15) show that the moras of the devoiced (and possibly deleted) high vowels remain. If they did not, then the bimoraic loanword truncation for, e.g., [sutoraiki̥] would be *[sutora] instead of [suto̥]; the hypocoristic for, e.g, [tçikako] would be *[tçika:] or *[tçikka] instead of [tçi̥ka]; and, similarly, the mimetic for 'rainy' would be *[çitto-çitto] or *[çito:-çito:] instead of [çito-çito].

    To further corroborate this observation, Hirayama (2009) showed that moras of devoiced vowels count in *haiku*, whose rhythm is based on mora counts, in the same way as voiced vowels. To the extent that onset consonants do not project a mora (e.g. Hayes 1989, cf. Topintzi 2008, 2010), then, this observation supports H2 in Figure 1. At the very least, the patterns in (13)-(15) show that the moras of devoiced vowels remain. If these cases of devoiced vowels also contain variable deletion, as in Shaw & Kawahara (2018c), then the mora must be docked to the remaining consonant, as in H2 in Figure 1.

    Phonologically, some evidence suggests that syllables of devoiced vowels remain as well. Ito (1990) observes that the morphophological truncation pattern in (16) cannot result in monosyllabic outputs, and that a light syllable is appended in such cases, as in (17).[6] Ito & Mester (1992/2003) formalize this pattern as a result of a binarity branching condition at the prosodic word level; a PrWd must branch at the level of the syllable. As shown in (18), a syllable hosted by a devoiced vowel satisfies this prosodic branching requirement. If devoiced vowels in this context are also deleted, then the syllabic requirement is being satisfied by the final consonant in the word. This supports the syllabic consonant analysis, as in H2.

(16)   Bimoraic truncation

      a.   [ookesu̥tora] → [oke] 'orchestra'
      b.   [rihaasaru] → [riha] 'rehearsal
      c.   [rokeeçoɴ] → [roke] 'location'

(17)   No monosyllabic outputs

      a.   [daijamondo] → [dai.ja] 'diamond'
      b.   [paamanento] → [paa.ma] 'permanent'
      c.   [kombineeçoɴ] → [kom.bi] 'combination'
      d.   [çimpoʑium] → [çim.po] 'symposium'
      e.   [impotentsu̥] → [im.po] 'impotent'
      f.   [kompoonento] → [kom.po] '(stereo) compoent'

(18)   Devoiced vowels count

      a.   [maikuroɸooɴ] → [mai.ku̥] 'microphone'
      b.   [ampuriɸaiaa] → [am.pu̥] 'amplifier'
      c.   [paŋkutçaa] → [paŋ.ku̥] 'puncture'
      d.   [wam.pii.su̥] → [wam.pi̥] 'one piece'

Another piece of phonological evidence comes from patterns of pitch accent placement. Kubozono (2011) argues that the Japanese default accent pattern, which is observed in loanwords and nonce word pronunciation, generally follows the Latin Stress Rule: (i) place the accent on the penultimate syllable if it is heavy (19), (ii) otherwise place the accent on the antepenultimate syllable (20) (see also Kawahara 2015b).[7] The presence of devoiced vowels does not disrupt this pattern (21). In cases of vowel deletion, the final consonant must still count as a syllable.

(19)   Accent on penultimate syllable if heavy

      a.   [ɸu.re'n.do] 'friend'
      b.   [pu.ra'a.to] 'Praat'

---

[6] Labrune (2012) attempted to reanalyze this pattern without recourse to syllables; Kawahara (2016) argues that this reanalysis misses an important generalization, and reference to syllables is crucial.

[7] Not all loanwords follow Latin Stress Rule; for example, some four-mora words can be unaccented (Ito & Mester, 2016b; Kubozono, 1996, 2006). What is important in the current discussion is that these forms which follow Latin Stress Rule do not show accent shift one syllable to the left.

      c.   [pu.ra'i.zu] 'prize'
      d.   [pu.ro'o.zu] 'prose'
      e.   [ɸu.ro'o.zuɴ] 'frozen'
      f.   [maa.ma.re'e.do] 'marmalade'

(20)   Otherwise accent on antepenultimate syllable

      a.   [re'.ba.noɴ] 'Lebanon'
      b.   [se'.ku.taa] 'sector'
      c.   [do'.ku.taa] 'doctor'
      d.   [pa'.ku.tɕii] 'coriander'
      e.   [ɕi'.na.moɴ] 'cinnamon'
      f.   [ga'.va.doɴ] (proper name)

(21)   Devoicing does not affect LSR

      a.   [pu.ro'.se.su̥] 'process"
      b.   [bu.ra'n.ku̥] 'blank'
      c.   [su̥.pa'i.su̥] 'spice'
      d.   [su̥.pa'i.ku̥] 'spike'
      e.   [ku.re'e.pu̥] 'crepe'
      f.   [pu.ra'i.su̥] 'price'

Moreover, there is evidence from compound accentuation patterns and statistical distributions in native words that Japanese strongly disfavors accent on final syllables (Kubozono, 1995, 2011). Given this dispreference, take words like [pu.ra'i.su̥] and [ku.re'e.pu̥]. If the final syllables are lost due to high vowel deletion, it would be natural to expect that accent shifts away to the word-initial syllables, which does not occur. This lack of accentual shift also supports the view that the syllables of deleted high vowels remain phonologically.

In addition, devoiced syllables can bear pitch accents in modern Japanese (Kawahara, 2015b; Vance, 1987). For example, Japanese accented verbs predictably bear accent on the penultimate syllable; when the penultimate syllables in verbs are devoiced, accent remains on that syllable (e.g. [kaku̥'su̥] 'hide'; [tsu̥'ku̥] 'to arrive'), especially in the speech of contemporary young speakers. Since the accent bearing unit in Japanese is the syllable (Kawahara, 2016; Kubozono, 2003; McCawley, 1968), this observation too shows that syllables are maintained even in the presence of devoiced vowels. If, besides being devoiced, the vowel is also deleted in some of these cases, it must be that the remaining consonant supports the presence of the syllable.

All of these observations converge on one conclusion: morphophonological processes that make reference to prosodic structure in Japanese do not treat devoiced vowels and voiced vowels differently. To the extent that devoiced vowels are deleted (Beckman, 1982, 1996; Beckman & Shoji, 1984; Matsui, 2017; Shaw & Kawahara, 2018c), then the general conclusion should be that moras and syllables remain after the deletion of these vowels, which is consistent with H2 in Figure 1.

In the next section, we further corroborate this conclusion from the perspective of articulatory coordination. In particular, we build on previous research findings that different syllable structures show different articulatory stability patterns (Browman & Goldstein, 1988; Byrd, 1995; Goldstein et al., 2007; Hermes et al., 2013, 2017; Marin, 2013; Marin & Pouplier, 2010; Shaw & Gafos, 2015; Shaw et al., 2009).

4   Temporal stability analysis

4.1 *Approach* The following analysis is based on ElectroMagnetic Articulograph (EMA) data obtained for the study reported in Shaw & Kawahara (2018c). The general idea of the analysis is, as illustrated in Figure 4, to evaluate patterns of temporal stability in syllable-referential intervals across CV and CCV sequences. Previous studies, beginning with pioneering work by Browman & Goldstein (1988), have shown that languages that parse word-initial consonants tautosyllabically, i.e., as complex syllable onsets, tend to exhibit a specific pattern of temporal stability across CV and (C)CCV sequences. This general observation includes results for English (Browman & Goldstein, 1988; Honorof & Browman, 1995; Marin & Pouplier, 2010), Romanian (Marin, 2013), and rising sonority clusters in Italian (Hermes et al., 2013). Specifically, as illustrated schematically in the right side of Figure 4, in these languages the center-to-anchor interval is more stable across CV and CCV sequences than the left edge-to-anchor interval or the right edge-to-anchor

interval (a.k.a. "c-center effect") . In contrast, languages that enforce a heterosyllabic parse of initial CCV sequences, e.g., Moroccan Arabic and Tashlhiyt Berber (Dell & Elmedlaoui, 2002), tend to exhibit a different stability pattern. As illustrated schematically in the left side of Figure 4, these languages tend to show right edge-to-anchor stability (for Berber, see Hermes et al. 2017; for Arabic, see Shaw et al. 2009).
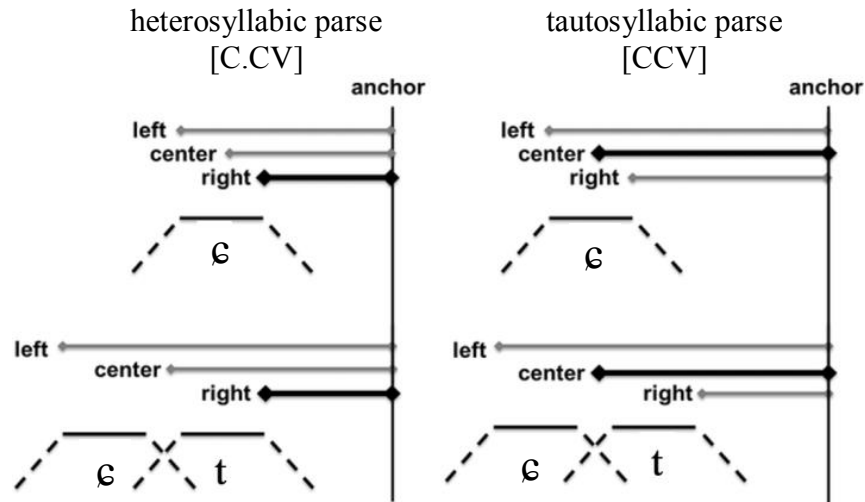

Figure 4: Illustration of stability indices—heterosyllabic parse vs. tautosyllabic parse.

The different patterns of temporal stability illustrated in Figure 4 can be derived from distinct coordination topologies organizing the relative timing of consonant and vowel gestures (Gafos et al., 2014; Shaw & Gafos, 2015). The key assumption linking syllable structure to patterns of temporal stability is an isomorphism between the arrangements of segments into syllables and the network of coordination relations that makes up the coordination topology. Specifically, onset consonants are assumed to enter into a relation of temporal coordination with the syllable nucleus, an assumption adopted from Browman & Goldstein (2000). Relevant coordination topologies are illustrated in Figure 5. Gestures are represented as vertices and coordination relations between them are represented as edges, a schema which follows the representational formalism developed in Gafos (2002). Different types of coordination relations are color-coded. The relation between adjacent consonants, i.e., C-C coordination, is shown in blue; the relation between an onset consonant and a vowel, i.e., C-V coordination, is shown in red. For completeness, a yellow edge is also included, which indicates a relation between a vowel and possible post-vocalic segment, i.e., V-C coordination, although it does not play a role in the current analysis.
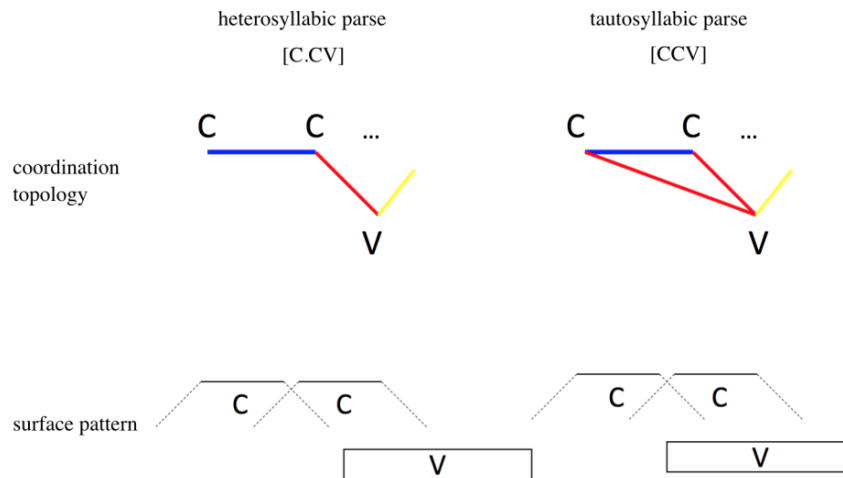
Figure 5: Relation between syllable parse, coordination topology and surface timing pattern for heterosyllabic (left) and tautosyllabic (right) parses of consonant clusters.

Under a heterosyllabic parse of initial consonants (Figure 5, left), the initial consonant is not contained in the same syllable as the following vowel—it is not a syllable onset—and, therefore, it is timed only to the following consonant (and not to the following vowel). In contrast, under a tautosyllabic parse (Figure 5, right) both pre-vocalic consonants are syllable onsets and, therefore, both enter into a coordination relation with the following vowel. Under the assumptions adopted here, complex onsets result in a coordination topology that, unlike the heterosyllable parse, places competing constraints on the temporal organization of gestures. That is, to satisfy the pattern of relative timing imposed by C-V coordination, the onset consonants would have to be temporally overlapped in time, a violation of C-C coordination. Satisfying C-C coordination, on the other hand, would entail a violation of C-V coordination. Although proposals differ in the technical details of how such competition is resolved (Browman & Goldstein, 2000; Gafos, 2002; Goldstein et al., 2009), the surface timing patterns shown at the bottom of Figure 5 derive from the coordination topologies shown in the center of the Figure. It is therefore possible to recover a syllabic parse from the pattern of relative timing in articulatory movements. We make use of this mapping to bring in phonetic data bearing on the syllabification of consonant clusters in Japanese. Specifically, we pursue a stability analysis, evaluating the stability of intervals across CV and CCV sequences (Figure 4) to assess whether consonant clusters in Japanese resulting from targetless vowels syllabify like sequences in Arabic (i.e., $C_1.C_2V$, according to H2) or sequences in English (i.e., C_1C_2V, according to H1).

4.2 *Method*   The stimuli are listed in Table 1. They contained five dyads, the members of which differ in whether they contain a devoicable high vowel (first column) or not (second column); in addition, the stimuli included singleton controls (third column).

Table 1: The list of the stimuli.

Voiced vowel	Deleted (devoiced) vowel	Singleton control
[masuda] (personal name)	[mastaa] 'master'	[bataa] 'butter'
[yakuzai] 'medication'	[haksai] 'white cabbage'	[dasai] 'uncool'
[ɕudaika] 'theme song'	[ɕtaisee] 'subjectivity'	[taisee] 'system'
[ɸuzoku] 'attachment'	[ɸsoku] 'shortage'	[kasoku] 'acceleration'
[katsudoo] 'activity'	[katstoki] 'when winning'	[mirutoki] 'when looking'

Six native speakers of Tokyo Japanese (3 male) read items in the carrier phrase [okee ⌣ to itte] 'ok, say ⌣', where the underlined blank indicated the position of the target word. Items were randomized within a block, and 10-15 blocks were recorded per participant. For additional methodological details, such as EMA sensor attachments and post-processing routines, see Shaw & Kawahara (2018c). The second author and one research assistant inspected the acoustics of the produced tokens and found that all devoicable vowels were actually devoiced.

In order to assess whether the devoiced vowels were deleted or not, Shaw & Kawahara (2018c) analyzed tongue dorsum trajectories from the vowel preceding [u], e.g., [a] in [katsudoo] or [e] from the carrier phrase in [e#ɕudaika], to the following vowel, e.g., [o] in in [katsudoo] or [a] in [ɕudaika]. A sample illustration is given in Figure 6, which plots tongue dorsum height trajectories from the preceding vowel [e] in the frame sentence [ookee] through [u] and onto the following vowel [a]. The blue lines represent tongue dorsum movement across the underlined portion of [e#ɕudaika], whereas the red lines represent tongue dorsum movement across the underlined portion of [e#ɕutaisee]. A rise in tongue dorsum height between [e] and [a], corresponding to the intervening [u] is expected if there is an articulatory target for [u]. We observe from Figure 6 that when the [u]s are devoiced (red lines), the tongue dorsum does not substantially rise between [e] and [a], at least not in some tokens, indicating a lack of [u] target. To assess this quantitatively, Shaw & Kawahara (2018c) apply an analytical technique involving machine classification of the trajectories based on competing phonological hypotheses: (i) a vowel present scenario, for which the voiced vowels (Table 1: column one) provided the training data and (ii) a vowel absent scenario, which was simulated. The simulations were guided by the assumption of phonetic interpolation (Cohn, 1993; Keating, 1988; Pierrehumbert & Beckman, 1988), i.e., if there is no [u] target, then the tongue dorsum will move from [e] to [a]. The technique for simulating trajectories based on phonetic interpolation of flanking targets (including the hypothesized vowel absent scenario) is described and justified in further detail in Shaw & Kawahara (2018a). The outcome of the classification yields a posterior probability that the trajectories contain a vowel target. Shaw & Kawahara (2018c) found that the posterior probability of a vowel target was very high for some tokens and very low for others, but there were few intermediate values. They conclude that the data support an optional process of deletion; some tokens are produced like full, voiced vowels, whereas some tokens entirely lack an articulatory target.
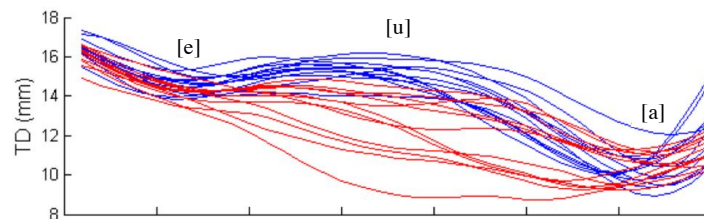

Figure 6: Sample tongue dorsum trajectories of [e#ɕudaika] (blue lines) and [e#ɕ(u)taisee] (red lines).

The current analysis builds on the results of Shaw & Kawahara (2018c). We applied the stability analysis to the subset of tokens that had a high (> 0.5) posterior probability of linear interpolation. These tokens were taken to lack a tongue dorsum target for [u], thereby forming a consonant cluster. This resulted in different numbers of tokens from different dyads. Only [ɕtaisee], [ɸsoku] and [katstoki] exhibited sufficient numbers of such tokens. For [ɕtaisee], there were 138 tokens (from five speakers) classified as deletion (lacking an [u] target); for [ɸsoku], there were 129 tokens (from four speakers); and, for [ɸsoku], there were 88 tokens (from two speakers). The following analysis is based on tokens from these three words, classified as lacking an [u] target, and an equal number of singleton controls. Since each item in Table 1 was produced in a block, we used in the analysis the singleton control from each block that also contained a case of vowel deletion. Consequently, the stability analysis below is based on 276 tokens for the [ɕtaisee] vs. [taisee] dyad, 258 tokens for [katstoki] vs. [mirutoki], and 176 tokens for [ɸsoku] vs. [kasoku].

The three intervals schematized in Figure 4, left-edge-to-anchor (LE_A), center-to-anchor (CC_A), and right-edge-to-anchor (RE_A) were calculated for each token containing a consonant cluster as well as for the

singleton control (Table 1: third column). The stability of these intervals across CV (singleton control) and CCV provided our phonetic diagnostic of syllable affiliation. All three of the intervals were right-delimited by a common anchor, the point of maximum constriction of the post-vocalic consonant. The landmarks that left-delimit the three intervals were parsed in the following manner (Figure 7). The LE_A interval was left-delimited by the achievement of target of the first consonant in the sequence, e.g., [ɕ] in [ɕtaisee] (and [t] in the singleton control [taisee]). The RE_A interval was left-delimited by the release of the immediately pre-vocalic consonant, e.g., [t] in [ɕtaisee] (and also [t] in the singleton control [taisee]). The third interval, CC_A was left-delimited by the mean of the midpoints between the consonants in the cluster and by the midpoint of the single onset consonant in the singleton control. The midpoint was the timestamp halfway between the achievement of target and the release. The target and release landmarks were determined from the articulatory signal with reference to movement velocity, allowing us to apply a uniform criterion for all consonants, regardless of manner or place of articulation. Specifically, we used 20% of peak velocity in the movement towards/away from consonantal constrictions. Figure 7 illustrates the parse of relevant landmarks for C_1 and C_2 in a token of [ɕtaisee]. The achievement of target and release of C_1, which is [ɕ] in this case, is shown on the tongue blade (TB) trajectory (blue line). The parse of C_2, [t], is shown on the tongue tip (TT) trajectory. As an index of interval stability across CV and CCV sequences, we computed the relative standard deviation (RSD), also known as the coefficient of variance, by dividing the standard deviation of interval duration calculated across tokens of CV and CCV by the mean interval duration across these same tokens.
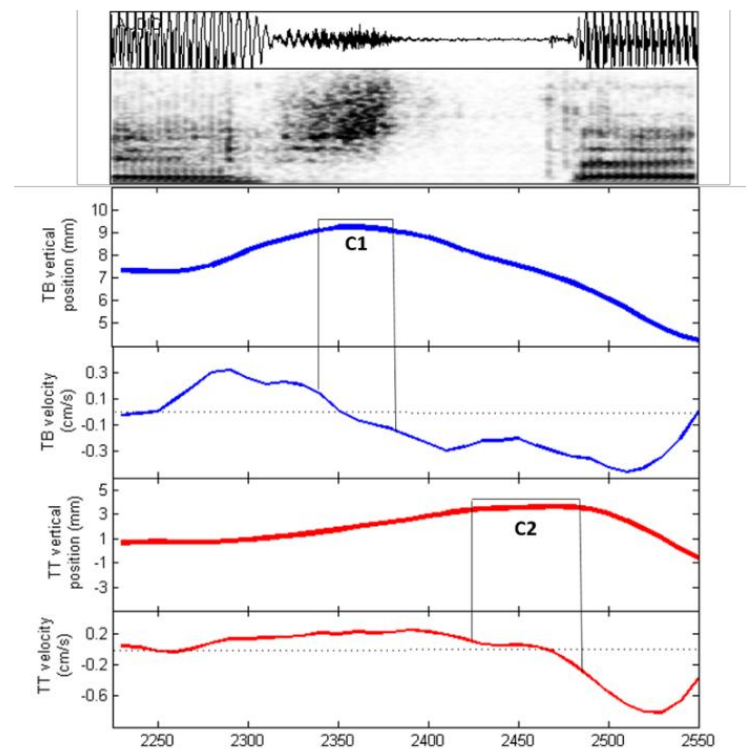

Figure 7: Illustration of how consonantal gestures were parsed based on a token of [ɕtaisee]. The portion of the signal shown begins with the [e] of the carrier phrase and ends with the [a]. The panels show, from top to bottom, the audio signal, spectrogram, tongue blade (TB) height trajectory, tongue blade (TB) velocity signal, tongue tip (TT) height trajectory, and tongue tip velocity signal. The thin black lines show the achievement of target and release of the consonants, C_1 and C_2, and the 20% threshold of the velocity peak that was used to parse them.

4.3 *Results* Figure 8 shows boxplots of interval duration for LE_A, CC_A, and RE_A intervals as calculated across CV and CCV strings in three dyads. Of course, it is always the case LE_A is the longest, followed by CC_A and then RE_A—what we are interested in is the degree of variability of these intervals, as, following the schema in Figure 4, this measure provides phonetic evidence for syllabic organization. We observe that for each of the dyads, RE_A shows the least variability (i.e. the boxplots have the smallest width). This result suggests that vowels are timed with respect to the right edge of the CC clusters, c.f., the center of CC clusters. All else equal, shorter intervals also tend to be less variable, a general property of timed events but also of other phonetic measurements (see, e.g., Nguyen & Shaw 2014 who show that variability in F1 and F2 for vowels is also correlated with the magnitude of the formant measurements). To correct for the effect that mean interval duration may have on the variability of the interval, we also computed the relative standard deviation.
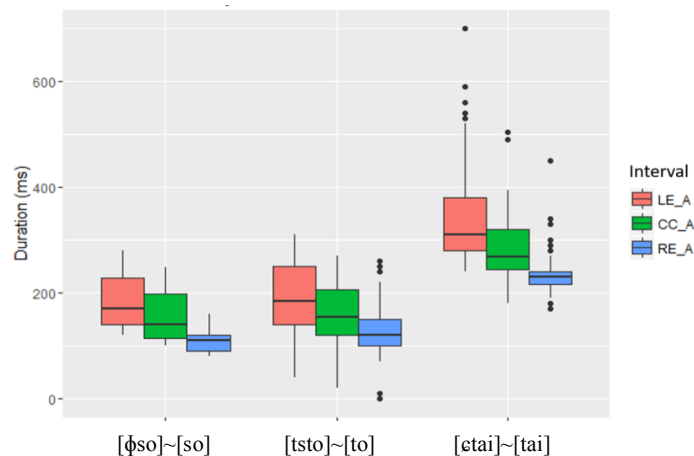

Figure 8: The durations between the three articulatory landmarks of the CC sequences and the vowel anchor.

The relative standard deviation (RSD) of the intervals in Figure 8 is shown in Table 2. Across dyads, the right-edge to anchor (RE_A) interval is the most stable (lowest RSD). On the assumptions illustrated in Figure 5, this pattern points unequivocally to simplex onsets, i.e., a heterosyllabic parse of initial clusters. Although care must be taken when interpreting stability patterns in terms of syllable structure, a point we return to in the general discussion, the pattern of RE_A stability is one of the most straightforward to interpret. The results of the stability analysis, therefore, converge on the same hypothesis as the phonological evidence discussed in section 3. Both point to H2, a heterosyllabic parse of consonant clusters arising from high vowel deletion.

Table 2: Relative standard deviation (RSD).

	LE_A	CC_A	RE_A
[ɸso] vs. [so]	0.32	0.34	0.24
[tsto] vs. [to]	0.25	0.23	0.20
[ɕtai] vs. [tai]	0.23	0.28	0.11

5 General discussion

To summarize, the EMA study by Shaw & Kawahara (2018c) showed that Japanese [u] optionally deletes in devoicing environments, yielding consonant clusters. Both phonological and phonetic evidence reviewed here suggests that these consonant clusters are parsed heterosyllabically. The current results imply a rather surprising conclusion that Japanese allows consonantal syllables headed by a fricative or an affricate, a conclusion that is especially surprising in light of the view that considers Japanese a "strict CV-language" (cf. Ito & Mester 2016a briefly discussed in the preamble).

The current results show that Japanese consonant clusters arising from high vowel deletion behave in terms of articulatory stability like word-initial consonant clusters in Moroccan Arabic. The similarity between Japanese and Moroccan Arabic is intriguing because word-initial clusters in Moroccan Arabic arose diachronically from the loss of short vowels (Benhallam, 1980), and there have been similar debates about syllabification based on internal phonological evidence, see, e.g., Keegan (1986: 214) who argues for complex onsets (H1 in Figure 1) vs. Kiparsky (2003) who argues for moraic consonants (H2). Ultimately, the weight of the evidence, which includes now arguments from temporal stability in articulation (Shaw et al., 2009) and metrical patterns in verse (Elmedlaoui, 2014) points to H2, the same conclusion that we have drawn for Japanese. In both cases, higher level syllabic structure is preserved despite the loss of a vowel.

More generally speaking, then, our data presents a case in which prosodic and temporal stability are maintained despite loss of a segment. Previously known cases of prosodic structure preservation include those discussed under the rubric of compensatory lengthening (Hayes, 1989; Kavitskaya, 2002; Wetzels & Sezer, 1986). In this pattern, higher level structure preservation is more salient because it conditions segmental-level lengthening. In the Japanese case, loss of a vowel neither lengthens adjacent segments nor shortens the transitions between consonants (Shaw & Kawahara, 2018c). The existence of patterns that delete segments while preserving prosodic structure supports independent representations of timing (prosodic structure) and articulation (segmental content), a dissociation with a known neural basis (Long et al., 2016). Generative phonology standardly assumes that prosodic structures are built off of segments, but it may instead be that prosody provides a temporal frame into which segments are "filled in" (cf. Fujimura 2000; Roelofs 1997; Sevald et al. 1995).

We also find the convergence between the phonological evidence (section 3) and the phonetic evidence (section 4) to be generally encouraging, as it speaks to the potential to reach common conclusions from diverse data sources (see Broselow et al. 1997 for a similar argument). One can address phonological questions by examining phonetic data, and phonological questions can guide us as to where to look in phonetic research (Beckman & Kingston, 1990).

We close here by pointing out some of the key assumptions that have been adopted to supported this convergence. For starters, we assumed at times that the vowel deletion observed in Shaw & Kawahara (2018c) is present in other environments in which devoicing is observed, particularly in the word final environment. This may not be necessarily the case. Kilbourn-Ceron & Sonderegger (2018) have recently argued in fact that the devoicing processes word-finally and between voiceless consonants come from different sources/mechanisms. The EMA data supporting vowel deletion in Shaw & Kawahara (2018c) includes only vowels occurring between voiceless consonants. However, our phonological arguments assume that deletion of devoiced vowels also occurs at least some of the time in devoicing contexts word-finally. If devoiced vowels word-finally are never deleted, then the phonological arguments we presented in section 3 are less compelling. A related alternative, which we cannot rule out, is that the vowel gesture is preserved in just those cases in which it is required to fulfill a morphophonological bimoraic requirement. Testing this hypothesis would require new EMA data. As it currently stands, the full force of our argument for converging phonological and phonetic evidence rests on the assumption that the optional deletion we have observed between voiceless consonants generalizes to other devoicing environments.

A second assumption, on the side of the temporal stability analysis, is that RE_A stability reflects a heterosyllabic parse of consonants. There are by now numerous studies that have applied this phonetic heuristic, which follows from the theoretical framework summarized in Figure 3. Through computational simulation using stochastic models, Shaw & Gafos (2015) probed the range of stability patterns (expressed in terms of RSD, as we do in this paper) that can arise from different parses of initial clusters. They found that it is not always the case that simplex onsets correspond to RE_A stability while complex onsets correspond to CC_A stability. In particular, they highlight specific conditions under which simplex onsets are predicted to condition CC_A stability. This happens when there is a high level of overall variability in the data. A realistic scenario of increasing variability presents itself in language acquisition. During the acquisition of the lexicon, increasing exposure to new words and new speakers increases the overall level of temporal variability in speech experience, which can drive a shift in the aggregate statistics from RE_A stability to CC_A stability (Gafos et al., 2014). In the case at hand, that of our Japanese data, the level of variability in the data is low enough that we can be reasonably sure that simplex onset topology (Figure 5: left) maps to RE_A stability. More importantly, the conditions under which a complex onset (Figure 5: right) parse could condition RE_A stability are exceedingly rare (given our working assumption that onset consonants are timed

to the syllable nucleus). We are therefore reasonably confident of our conclusions for the Japanese data, but a more complete analysis would report as well patterns of covariation between temporal intervals predicted by the competing hypothesis (see, e.g., Shaw & Davidson 2011 and Shaw et al. 2011). To the extent that the above assumptions are valid, the results provide support for H2, the hypothesis that Japanese consonant clusters resulting from vowel deletion are parsed heterosyllabically. This conclusion follows from converging evidence from the analysis of phonological patterns sensitive to syllable structure and an analysis of temporal stability in articulation.

Finally, returning to the general issue that we reviewed at the beginning of the paper, Japanese instantiates a case of *persistence of prosody* (Garcia et al., 2016; Kager, 1997), in that the rhythmic pattern is maintained after deletion of segments. Although we do not pretend as if we were the first one to find this pattern (see section 1.1), we believe that the current case study offers stronger evidence for *persistence of prosody* than previous research did. First, our analysis is based on an EMA study Shaw & Kawahara (2018c) which directly showed that vocalic gestures are indeed deleted, eliminating the possibility that segment deletion may instead be segment reduction (see also Matsui 2017 for converging evidence from an EPG experiment). Second, we established the lack of resyllabification, again using articulatory data. The lack of resyllabification was further corroborated by examination of morphophonological patterns. With these, we conclude that Japanese does have syllables headed by a fricative or by an affricate, at least.

We have limited our claims about syllabic consonants in Japanese in this paper to fricatives/affricates as it is for these manner classes that the currently available phonetic data happens to provide the strongest support, although we cannot rule out that other consonants can also head syllables. The phonological facts reviewed here are consistent with other voiceless consonants, including /k/, the only voiceless plosive that occurs before /u/, constituting syllable heads if, in fact, the devoiced vowel is deleted following /k/. In Shaw & Kawahara (2018c), we reported some data on this environment, i.e., the devoiced /u/ following /k/ in [haku̥sai]. Our method of detecting vowel deletion found very low rates in this environment. This is possibly due to the shared articulator between /k/ and /u/, which poses methodological challenges for our approach, as discussed in the paper; however, there are plausible phonological reasons as well why we might not expect to observe deletion in [haku̥sai]; i.e. why [ha.k.sai] is disfavored. One is that [k], an oral stop, does not form a very good syllable nucleus (Dell & Elmedlaoui, 1985; Prince & Smolensky, 1993/2004), maybe because a moraic stop is marked relative to moraic versions of more sonorous consonants (Zec, 1995). It may be the case that Japanese tolerates consonantal syllables only if the consonants are either [+continuant] or [+sonorant]. So far, the only conditioning environments for vowel deletion in Japanese that we know of require that the preceding consonant be voiceless, which (accidentally) precludes the possibility of syllabic consonants that are more sonorous than voiceless fricatives. Another possible phonological explanation for the lack of deletion following /k/ is syllable contact (Gouskova, 2004; Murray & Vennemann, 1983; Vennemann, 1988): a [k.s] sequence across a syllable boundary involves a rise in sonority, which is dispreferred to a fall in sonority. A follow-up EMA experiment which was designed to test this hypothesis has been conducted, and the analysis is on its way. If either of these hypotheses are correct, then it implies that even though Japanese allows consonantal syllables, they nevertheless follow markedness restrictions—restrictions on syllable nuclei or syllable contact—that are known cross-linguistically: i.e., we may be observing the emergence of the unmarked (McCarthy & Prince, 1994) in the high vowel deletion pattern in Japanese.

References

Abbott, Miriam (1991). Macushi. Derbyshire, Desmond J. & Geoffrey K. Pullum (eds.), *Handbook of Amazonian languages, Vol.3*, Mouton de Gruyter, Berlin, 23–160.

Adler, Jeff & Jesse Zymet (2017). Irreducible parallelism in phonology. Ms. University of California, Santa Cruz.

Anttila, Arto & Naomi Tachikawa Shapiro (2017). The interaction of stress and syllabification: Parallel or serial? *Proceedings of the 34th West Coast Conference on Formal Linguistics*, Cascadilla Press, 52–61.

Barnes, Jonathan & Darya Kavitskaya (2003). Phonetic analogy and schwa deletion in French. Larsen, Julie & Mary Paster (eds.), *Proceedings of Berkeley Linguistic Society 28*, BLS, Berkeley, 39–50.

Beckman, Mary (1982). Segmental duration and the 'mora' in Japanese. *Phonetica* 39, 113–135.

Beckman, Mary (1996). When is a syllable not a syllable. Otake, Takeshi & Anne Cutler (eds.), *Phonological Structure and Language Processing: Cross-Linguistic Studies*, Mouton de Gruyter, 95–124.

Beckman, Mary & John Kingston (1990). An introduction. Kingston, John & Mary Beckman (eds.), *Papers in Laboratory Phonology I: Between the grammar and physics of speech*, Cambrdige University Press, Cambridge, 1–16.

Beckman, Mary & Atsuko Shoji (1984). Spectral and perceptual evidence for CV coarticulation in devoiced /si/ and /syu/ in Japanese. *Phonetica* 41, 61–71.

Benhallam, Abderrafi (1980). *Syllable structure and rule types in Arabic*. Doctoral dissertation, University of Florida.

Blevins, Juliette (1995). The syllable in phonological theory. Goldsmith, John A. (ed.), *The Handbook of Phonological Theory*, Blackwell, Cambridge, Mass., and Oxford, UK, 206–244.

Bowers, Dustin (2015). Phonological restructuring in Odawa. Ms. Yale University.

Bowers, Dustin (2018). The Nishnaabemwin restructuring controversy: New empirical evidence. Ms. Yale University.

Broselow, Ellen, Su-I. Chen & Marie Huffman (1997). Syllable weight: Convergence of phonology and phonetics. *Phonology* 14, 47–82.

Browman, Catherine & Louis Goldstein (1988). Some notes on syllable structure in Articulatory Phonology. *Phonetica* 45, 140–155.

Browman, Catherine & Louis Goldstein (2000). Competing constraints on intergestural coordination and self-organization of phonological structures. *Les Cahiers de l'ICP, Bulletin de la Communication Parlée* 5, 25–34.

Brown, Roger & David MacNeill (1966). The 'tip of the tongue' phenomenon. *Journal of Verbal Learning and Verbal Behavior* 5:4, 325–337.

Bürki, Andrey, Cécile Fougeron & Cedric Gendrot (2007). On the categorical nature of the process involved in schwa elision in French. *Proceedings of Interspeech* .

Bürki, Andrey, Cécile Fougeron, Cedric Gendrot & Ulrich H. Frauenfelder (2011). Phonetic reduction versus phonological deletion of French schwa: Some methodological issues. *Journal of Phonetics* 39:3, 279–288.

Byrd, Dani (1995). C-centers revisited. *Phonetica* 52, 285–306.

Clements, G. N. (1986). Syllabification and epenthesis in the Barra dialect of Gaelic. Bogers, K., M. Maus & Harry van der Hulst (eds.), *The Phonological Representation of Suprasegmentals*, Foris, Dordrecht, 317–336.

Clements, Nick & Samuel Jay Keyser (1983). *CV Phonology: A Generative Theory of the Syllable*. MIT Press, Cambridge.

Cohn, Abigail (1993). Nasalisation in English: Phonology or phonetics. *Phonology* 10, 43–81.

Cutler, Anne (1980). Errors of stress and intonation. Fromkin, Victoria A. (ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen and hand*, Academic Press, New York, 67–80.

Davidson, Lisa (2005). Addressing phonological questions with ultrasound. *Clinial Linguistics & Phonetics* 19:6/7, 619–633.

Davidson, Lisa (2006). Schwa elision in fast speech: Segmental deletion or gestural overlap. *Phonetica* 63:2-3, 79–112.

Davidson, Lisa & Maureen Stone (2004). Epenthesis versus gestural mistiming in consonant cluster production. *Proceedings of the West Coast Conference on Formal Linguistics (WCCFL) 22*, Cascadilla Press.

Dehaene-Lambertz, G., Emmanuel Dupoux & A. Gout (2000). Electrophysiological correlates of phonological processing: A cross linguistic study. *Journal of Cognitive Neuroscience* 12, 635–647.

Dell, François & Mohamed Elmedlaoui (1985). Syllabic consonants and syllabification in Imdlawn Tashlhiyt Berber. *Journal of African Languages and Linguistics* 7, 105–130.

Dell, François & Mohamed Elmedlaoui (2002). *Syllables in Tashlhiyt Berber and in Moroccan Arabic*. Dordrecht, Netherlands.

DiCanio, Christian T. (2012). The phonetics of fortis and lenis consonants in Itunyoso Trique. *International Journal of American Linguistics* 78:2, 239–272.

DiCanio, Christian T. (2014). The sounds of Triqui: Quantitative approaches to language description and its ramifications for historical change. Talk give at University of Albany.

Dupoux, Emmanuel, Kazuhiko Kakehi, Yuki Hirose, Christophe Pallier & 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, Sónia Frota, Yuki Hirose & Sharon Peperkamp (2011). Where do illusory vowels come from? *Journal of Memory and Language* 64:3, 199–210.

Elfner, Emily (2006). Contrastive syllabification in Blackfoot. Baumer, Donald, David Montero & Michael Scanlon (eds.), *Proceedings of WCCFL 25*, Cascadilla Press, 141–149.

Elmedlaoui, Mohamed (2014). What does the Moroccan Malħun meter compute, and how? *The Form of Structure, the Structure of Form: Essays in honor of Jean Lowenstamm* .

Faber, Alice & Timothy Vance (2010). More acoustic traces of "deleted" vowels in Japanese. Nakayama, Mineharu & Carles Quinn (eds.), *Japanese/Korean Linguistics 9*, CSLI, 100–113.

Fougeron, Cécile & Donca Steriade (1997). Does deletion of French schwa lead to neutralization of lexical distinctions? *Euro-Speech 1997, Proceedings of the 5th European Conference on Speech Communication and Technology* 943–946.

Fujimura, Osamu (2000). The C/D model and prosodic control of articulatory behavior. *Phonetica* 57, 128–138.

Fujimura, Osamu & C.J. Williams (2015). Remarks on the C/D model. *Journal of the Phonetic Society of Japan* 19:2, 2–8.

Gafos, Adamantios (2002). A grammar of gestural coordination. *Natural Language and Linguistic Theory* 20, 269–337.

Gafos, Adamantios, Simon Charlow, Jason Shaw & Phil Hoole (2014). Stochastic time analysis of syllable-referential intervals and simplex onsets. *Journal of Phonetics* 44, 152–166.

Garcia, Guilherme D., Heather Goad & Natália Brambatti Guzzo (2016). Footing is not always about stress: Formalizing variable high vowel deletion in Québec French. *Proceedings of Annual Meeting of Phonology* .

Goldstein, Louis, Ioana Chitoran & Elisabeth Selkirk (2007). Syllable structure as coupled oscillator modes: Evidence from Georgian vs. Tashlhiyt Berber. Paper presented at the XVIth International Congress of Phonetic Sciences, Saabrucken, Germany.

Goldstein, Louis, Hosung Nam, Saltzman E.L. & Ioana Chitoran (2009). Coupled oscillator planning model of speech timing and syllable structure. Fant, Gunnar, Hiroya Fujisaki & J. Shen (eds.), *Frontiers in Phonetics and Speech Science: Festschrift for Wu Zongji*, Commercial Press, Beijing, 239–249.

Gouskova, Maria (2004). Relational markedness in ot: The case of syllable contact. *Phonology* 21:2, 201–250.

Hall, Nancy (2006). Cross-linguistic patterns of vowel intrusion. *Phonology* 3, 387–429.

Hayes, Bruce (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20, 253–306.

Hayes, Bruce (1995). *Metrical Stress Theory: Principles and Case Studies*. The University of Chicago Press, Chicago.

Hermes, Anne, D. Mücke & Martine Grice (2013). Gestural coordination of Italian word-initial clusters: The case of "impure s". *Phonology* 30:1, 1–25.

Hermes, Anne, D. Mücke & B. Auris (2017). The variability of syllable patterns in Tashlhiyt Berber and Polish. *Journal of Phonetics* 64, 127–144.

Hirayama, Manami (2009). *Postlexical prosodic structure and vowel devoicing in Japanese*. Doctoral dissertation, University of Toronto.

Honorof, Douglas Nathan & Catherine P. Browman (1995). The center or the edge: How are consonant clusters organised with respect to the vowel? Paper presented at the XIIIth International Congress of Phonetic Sciences, Stockholm, Sweden.

Hume, Elizabeth (1997). Vowel preservation in Leti. *Oceanic Linguistics* 36:1, 65–101.

Hyde, Brett (2012). The odd-parity input problem in metrical stress theory. *Phonology* 29:3, 383–431.

Ito, Junko (1986). *Syllable Theory in Prosodic Phonology*. Doctoral dissertation, University of Massachusetts, Amherst.

Ito, Junko (1989). A prosodic theory of epenthesis. *Natural Language and Linguistic Theory* 7, 217–259.

Ito, Junko (1990). Prosodic minimality in Japanese. Ziolkowski, Michael, Manual Noske & Karen Deaton (eds.), *Proceedings of Chicago Linguistic Society: Parasession on the Syllable in Phonetics and Phonology*, Chicago Linguistic Society, Chicago, 213–239.

Ito, Junko & Armin Mester (1992/2003). Weak layering and word binarity. Homma, Takeru, Masao Okazaki, Toshiyuki Tabata & Shin-Ichi Tanaka (eds.), *A New Century of Phonology and Phonological Theory*, Kaitakusha, Tokyo, 26–65.

Ito, Junko & Armin Mester (2015). Word formation and phonological processes. Kubozono, Haruo (ed.), *The Handbook of Japanese Language and Linguistics: Phonetics and Phonology*, Mouton, 363–395.

Ito, Junko & Armin Mester (2016a). Pitch accents as tonal complexes: Evidence from superheavies. Talk presenetd at the workshop "Syllables and prosody" at NINJAL.

Ito, Junko & Armin Mester (2016b). Unaccentedness in Japanese. *Linguistic Inquiry* 47:3, 471–526.

Jun, Sun-Ah & Mary Beckman (1993). A gestural-overlap analysis of vowel devoicing in Japanese and Korean. Paper presented at the 67th annual meeting of the Linguistic Society of America, Los Angeles.

Jun, Sun-Ah, Mary Beckman & H.-J. Lee (1998). Fiberscopic evidence for the influence on vowel devoicing of the glottal configurations for Korean obstruents. *UCLA Working Papers in Phonetics* 96, 43–68.

Kager, René (1997). Rhythmic vowel deletion in Optimality Theory. Roca, Iggy (ed.), *Derivations and Constraints in Phonology*, Oxford University Press, Oxford, 463–499.

Kahn, Daniel (1976). *Syllable-based Generalizations in English Phonology*. Doctoral dissertation, MIT.

Kaisse, Ellen M. & Patricia Shaw (1985). On the theory of lexical phonology. *Phonology* 2, 1–30.

Karins, Krisjanis (1995). Vowel deletion in Latvian. *Language Variation and Change* 7, 15–34.

Katayama, Motoko (1998). *Optimality Theory and Japanese loanword phonology*. Doctoral dissertation, University of California, Santa Cruz.

Kavitskaya, Dasha (2002). *Compensatory Lengthening: Phonetics, Phonology, Diachrony*. Routledge, New York.

Kawahara, Shigeto (2002). Similarity among variants: Output-variant correspondence. BA thesis, International Christian University, Tokyo Japan.

Kawahara, Shigeto (2015a). A catalogue of phonological opacity in Japanese. *REPORTS of the Keio Institute of Cultural and Linguistic Studies* 46, 145–174.

Kawahara, Shigeto (2015b). The phonology of Japanese accent. Kubozono, Haruo (ed.), *The Handbook of Japanese Language and Linguistics: Phonetics and Phonology*, Mouton de Gruyter, Mouton, 445–492.

Kawahara, Shigeto (2016). Japanese has syllables: A reply to Labrune (2012). *Phonology* 33, 169–194.

Kawakami, Shin (1977). *Nihongo Onsei Gaisetsu [An Overview of Japanese Phonetics]*. Ohuusha, Tokyo.

Keating, Patricia A. (1988). Underspecification in phonetics. *Phonology* 5, 275–292.

Keating, Patricia A. (2003). Phonetic encoding of prosodic structure. *Proceedings of the 6th International Seminar on Speech Production* 119–124.

Keating, Patricia A. & Stephanie Shattuck-Hufnagel (2002). A prosodic view of word form encoding for speech production. *UCLA Working Papers in Phonetics* 101, 112–156.

Keegan, John M. (1986). *The Phonology and Morphology of Moroccan Arabic*. UMI, Ann Arbor.

Kilbourn-Ceron, Oriana & Morgan Sonderegger (2018). Boundary phenomena and variability in Japanese high vowel devoicing. *Natural Language and Linguistic Theory* 36:1, 175–217.

Kiparsky, Paul (2003). Syllables and moras in Arabic. Féry, Caroline & Ruben van de Vijver (eds.), *The Optimal Syllable*, Cambridge University Press, Cambridge, 147–182.

Kirchner, Robert (1997). Contrastiveness and faithfulness. *Phonology* 14.1, 83–111.

Kondo, Mariko (1997). *Mechanisms of vowel devoicing in Japanese*. Doctoral dissertation, University of Edinburgh.

Kubozono, Haruo (1995). Constraint interaction in Japanese phonology: Evidence from compound accent. Walker, Rachel, Ove Lorentz & Haruo Kubozono (eds.), *Phonology at Santa Cruz [PASC] 4*, Linguistics Research Center, UC Santa Cruz, Santa Cruz, vol. 4, 21–38.

Kubozono, Haruo (1996). Syllable and accent in Japanese: Evidence from loanword accentuation. *The Bulletin (Phonetic Society of Japan)* 211, 71–82.

Kubozono, Haruo (1999). Mora and syllable. Tsujimura, Natsuko (ed.), *The Handbook of Japanese Linguistics*, Blackwell, Oxford, 31–61.

Kubozono, Haruo (2003). The syllable as a unit of prosodic organization in Japanese. Fery, Caroline & Ruben van der Vijver (eds.), *The Syllable in Optimality Theory*, Cambridge University Press., Cambridge, 99–122.

Kubozono, Haruo (2006). Where does loanword prosody come from?: A case study of Japanese loanword accent. *Lingua* 116, 1140–1170.

Kubozono, Haruo (2011). Japanese pitch accent. van Oostendorp, Marc, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, Blackwell-Wiley, Oxford, 2879–2907.

Labrune, Laurence (2012). Questioning the universality of the syllable: Evidence from Japanese. *Phonology* 29:1, 113–152.

Levelt, Willem, Ardi Roelofs & Antje Myer (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22, 1–75.

Levin, Juliette (1985). *A Metrical Theory of Syllabicity*. Doctoral dissertation, MIT.

Long, Michael A., Kalman A. Katlowitz, Marlo A. Svirsky, Rachel C. Clary, Tara McAllister Byun, Najib Majaj, Hiroyuki Oya, Howard Matthew A III & Jeremy D.W. Greenlee (2016). Functional segregation of cortical regions underlying speech timing and articulation. *Neuron* 89:6, 1187–1193.

Maekawa, Kikuo (1989). Boin-no museika [vowel devoicing]. *Kouza: Nihongo-to Nihongo Kyooiku, vol 2.*, Meiji Shoin, Tokyo, 135–153.

Marantz, Alec (1982). Re reduplication. *Linguistic Inquiry* 13, 483–545.

Marin, Stefania (2013). The temporal organization of complex onsets and codas in Romanian: A gestural approach. *Journal of Phonetics* 41:3-4, 211–227.

Marin, Stefania & Marianne Pouplier (2010). Temporal organization of complex onsets and codas in American English: Testing the predictions of a gesture coupling model. *Motor Control* 14, 380–407.

Massaro, Dominic & Michael Cohen (1983). Phonological context in speech perception. *Perception & Psychophysics* 34, 338–348.

Matsui, Michinao (2017). On the input information of the C/D model for vowel devoicing in Japanese. *Journal of the Phonetic Society of Japan* 21:1.

McCarthy, John J. (1981). A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry* 12, 373–418.

McCarthy, John J. (2003). Sympathy, cumulativity, and the Duke-of-York gambit. Féry, Caroline & Ruben van de Vijver (eds.), *The Syllable in Optimality Theory*, Cambridge University Press, Cambridge, 23–76.

McCarthy, John J. (2007). *Hidden Generalizations: Phonological Opacity in Optimality Theory*. Equinox Publishing, London.

McCarthy, John J. (2008). The serial interaction of stress and syncope. *Natural Language and Linguistic Theory* 26:3, 499–546.

McCarthy, John J. (2010). An introduction to Harmonic Serialism. *Language and Linguistic Compass* 4:10, 1001–1018.

McCarthy, John J. & Alan Prince (1986). Prosodic morphology. Ms., University of Massachusetts and Rutgers University.

McCarthy, John J. & Alan Prince (1990). Foot and word in prosodic morphology: The Arabic broken plural. *Natural Language and Linguistic Theory* 8, 209–283.

McCarthy, John J. & Alan Prince (1994). The emergence of the unmarked: Optimality in prosodic morphology. Gonzalez, Merce (ed.), *Proceedings of the North East Linguistic Society 24*, GLSA Publications, Amherst, Mass., 333–379.

McCawley, James D. (1968). *The Phonological Component of a Grammar of Japanese*. Mouton, The Hague.

Mester, Armin (1990). Patterns of truncation. *Linguistic Inquiry* 21, 475–485.

Moreton, Elliott (2002). Structural constraints in the perception of English stop-sonorant clusters. *Cognition* 84, 55–71.

Murray, Robert & Theo Vennemann (1983). Sound change and syllable structure: Problems in Germanic phonology. *Language* 59, 514–28.

Nguyen, Nhung & Jason Shaw (2014). Why the SQUARE vowel is the most variable in Sydney. Paper presented at the 15th Australasian International Conference on Speech Science and Technology.

Nielsen, Kuniko (2015). Continuous versus categorical aspects of Japanese consecutive devoicing. *Journal of Phonetics* 52, 70–88.

Pierrehumbert, Janet B. & Mary Beckman (1988). *Japanese Tone Structure*. MIT Press, Cambridge.

Poser, William (1990). Evidence for foot structure in Japanese. *Language* 66, 78–105.

Prince, Alan & Paul Smolensky (1993/2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Blackwell, Malden and Oxford.

Pruitt, Kathryn (2010). Serialism and locality in constraint-based metrical parsing. *Phonology* 27:3, 481–526.

Rhodes, Richard (1985). *Eastern Ojibwa-Chippewa-Ottawa dictionary*. Mouton.

Rialland, A. (1986). Schwa et syllabes en Français. Wetzels, Leo & Engin Sezer (eds.), *Studies in Compensatory Lengthening*, Dordrecht, 187–226.

Roelofs, Ardi (1997). The WEAVER model of word-form encoding in speech production. *Cognition* 64:3, 249–284.

Roelofs, Ardi & Antje S. Meyers (1998). Metrical structure in planning the production of spoken words. *ournal of Experimental Psychology: Learning, Memory, and Cognition* 24:4, 922–939.

Rosenthall, Sam (1997). The distribution of prevocalic vowels. *Natural Language and Linguistic Theory* 15, 139–180.

Sawashima, Masayuki (1971). Devoicing of vowels. *Annual Bulletin of Research Institute of Logopedics and Phoniatrics* 5, 7–13.

Selkirk, Elisabeth (1982). The syllable. Hulst, Harry van der & Norval Smith (eds.), *The Structure of Phonological Representations*, Foris Publications, Dordrecht, vol. Part II, 337–385.

Selkirk, Elisabeth (1984). On the major class features and syllable theory. Aronoff, Mark & Richard T. Oehrle (eds.), *Language Sound Structures*, MIT Press, Cambridge, MA, 107–136.

Sevald, C.A., G.S. Dell & J.S. Cole (1995). Syllable structure in speech production: Are syllables chunks or schemas. *Journal of Memory and Language* 34:6, 807–820.

Shattuck-Hufnagel, Stephanie (2006). Prosody first or prosody last? evidence from the phonetics of word-final /t/ American English. Goldstein, Louis, D. H. Whalen & Catherine Best (eds.), *Laboratory Phonology 8*, Mouton, 445–472.

Shaw, Jason & Lisa Davidson (2011). Perceptual similarity in input-output mappings: A computational/experimental study of non-native speech production. *Lingua* 121:8, 1344–1358.

Shaw, Jason & Adamantios Gafos (2015). Stochastic time models of syllable structure. *PLOS ONE* 10:5.

Shaw, Jason & Shigeto Kawahara (2018a). Assessing feature specification in surface phonological representations through simulation and classification of phonetic data. *Phonology* 35.

Shaw, Jason & Shigeto Kawahara (2018b). Consequences of high vowel deletion for syllabification in Japanese. *Proceedings of Annual Meeting of Phonology 2017* .

Shaw, Jason & Shigeto Kawahara (2018c). The lingual gesture of devoiced [u] in Japanese. *Journal of Phonetics* 66, 100–119.

Shaw, Jason, Adamantios Gafos, Phil Hoole & Chakir Zeroual (2009). Syllabification in Moroccan Arabic: evidence from patterns of temporal stability in articulation. *Phonology* 26:1, 187–215.

Shaw, Jason, Adamantios Gafos, Phil Hoole & Chakir Zeroual (2011). Dynamic invariance in the phonetic expression of syllable structure: A case study of Moroccan Arabic consonant clusters. *Phonology* 28:3, 455–490.

Smolensky, Paul (1996). The initial state and 'Richness of the Base' in Optimality Theory. Ms. Johns Hopkins University.

Steriade, Donca (1982). *Greek Prosodies and the Nature of Syllabification*. Doctoral dissertation, MIT.

Steriade, Donca (2000). Paradigm uniformity and the phonetics-phonology boundary. Broe, Michael B. & Janet B. Pierrehumbert (eds.), *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, Cambridge University Press, Cambridge, 313–334.

Topintzi, Nina (2008). On the existence of moraic onset geminates. *Natural Language and Linguistic Theory* 26:1, 147–184.

Topintzi, Nina (2010). *Onsets: Suprasegmental and prosodic behaviour*. Cambrdige University Press, Cambridge.

Tsuchida, Ayako (1997). *Phonetics and Phonology of Japanese Vowel Devoicing*. Doctoral dissertation, Cornell University.

Urbanczyk, Suzanne (1996). *Patterns of Reduplication in Lushootseed*. Doctoral dissertation, University of Massachusetts.

Vance, Timothy (1987). *An Introduction to Japanese Phonology*. SUNY Press, New York.

Vennemann, Theo (1988). *Preference Laws for Syllable Structure and the Explanation of Sound Change: With Special Reference to German, Germanic, Italian, and Latin*. Mouton de Gruyter, Berlin.

Verluyten, Sylvain Paul Marcel (1982). *Recherches sur la prosodie et la métrique du français*. Doctoral dissertation, Universiteit Antwerpen.

Wetzels, Leo & Engin Sezer (eds.) (1986). *Studies in Compensatory Lengthening*. Foris, Dordrecht.

Whang, James (2017). *Shaping speech patterns via predictability and recoverability*. Doctoral dissertation, New York University.

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

Zec, Draga (1995). Sonority constraints on syllable structure. *Phonology* 12, 85–129.