Japanese has syllables: A reply to Labrune (2012)

Abstract

Labrune (2012b) proposes a syllable-less theory of Japanese, suggesting that Japanese has no syllables and instead has only moras below feet. Labrune (2012b) argues that there is no phonetic or psycholinguistic evidence for the existence of syllables in Japanese. This reply summarizes and reexamines previous experimental findings that demonstrate that Japanese does show evidence for syllables both phonetically and psycholinguistically. After the extensive review of these previous experimental studies, this reply also takes up a few phonological and theoretical issues that merit explicit response from the syllable proponent perspective. Based on these considerations, this paper concludes that Japanese does have syllables.

1 Introduction

In a provocative article, Labrune (2012b) argues that there is little phonetic or psycholinguistic evidence for syllables in Tokyo Japanese (henceforth Japanese), and that phonological phenomena which have been hitherto analyzed in terms of syllables can be reanalyzed by deploying a distinction between a "regular/full mora" and a "deficient/special mora". She concludes that Tokyo Japanese does not have syllables, and as a further theoretical consequence of this view, she argues that not all prosodic levels are universal, extending on the suggestions by Hyman (1985, 2008). Although this proposal is very thought-provoking and its potential theoretical consequence is an important one, it does miss a substantial body of previous experimental findings about the existence of syllables in the prosodic organization of Japanese. Therefore, this reply article reexamines the evidence demonstrating that Japanese does show evidence for syllables both phonetically and psycholinguistically (sections 3 and 4). After the summary, this reply also addresses some of the phonological and theoretical issues which merit further discussion from the perspective of the syllable proponents (section 5).[1]

[1] Vance (2013), in his book review of Labrune (2012a), which advances the same syllable-less view of Japanese as Labrune (2012b), says "I look forward to seeing how syllable proponents will respond (p.171)." Here is a concrete attempt by one of the syllable proponents. See also Ito & Mester (2015) and Tanaka (2013) for other critical responses to the syllable-less theory of Japanese.

2 Background: Heavy syllables in Japanese

To offer background for the discussion that follows, this section provides an overview of different types of sequences which have been considered as "heavy syllables" in the standard phonological analyses of Japanese. It is important to focus on heavy syllables, because light syllables and moras coincide in Japanese. It is therefore the existence of heavy syllables that is crucial to posit syllables, in addition to moras, in the prosodic organization of Japanese phonology.

Since Japanese syllable structures are simple, most of the time syllables and moras coincide (e.g. [ta] is both monomoraic and monosyllabic). Moras and syllables diverge in the case of heavy syllables, which contain two moras (e.g. [taa] with a long [aa] is monosyllabic, but bimoraic). Japanese has limited types of heavy syllables, as summarized in Table 1.[2]

Table 1: The four types of heavy syllables in Japanese.

Second part of the syllable	phonemic rep.	phonetic rep.	gloss
First half of geminate : /Q/	/ha**Q+p**uN/	[ha**pp**uɴ]	'excitement'
	/ha**Q+t**atu/	[ha**tt**atsu]	'development'
	/ha**Q+k**eN/	[ha**kk**eɴ]	'discovery'
	/ha**Q+s**aN/	[ha**ss**aɴ]	'diffusion'
Moraic nasal: /N/	/he**N**/	[he**ɴ**]	'strange'
	/he**N+p**aJ/	[he**mp**ai]	'to pour back'
	/he**N+d**a/	[he**nd**a]	'is strange'
	/he**N+k**a/	[he**ŋk**a]	'change'
Second part of long vowel: /R/	/ok**aR**saN/	[ok**aa**saɴ]	'mother'
	/ot**oR**saN/	[ot**oo**saɴ]	'father'
	/on**eR**saN/	[on**ee**saɴ]	'sister'
	/on**iR**saN/	[on**ii**saɴ]	'brother'
Second part of diphthong: /J/	**/aJ/**	**[ai]**	'love'
	/oJ/	**[oi]**	'nephew'

The first type of heavy syllable contains the first half of a geminate, or long consonant (Kawagoe, 2015; Kawahara, 2015a). The traditional label for the first half of geminate is /Q/ (called *sokuon* in the traditional Japanese grammar), a convention that Labrune (2012b) also deploys. This /Q/ archiphoneme assimilates to the following consonant in manner and place, surfacing as a geminate consonant. For example, /ha**Q+t**atu/ is realized as [ha**tt**atsu],[3] and /ha**Q**+keN/ is realized as [ha**kk**eɴ]. In this example, /haQ/—or [hat] and [hak]—is monosyllabic but bimoraic.

[2]This reply sets aside syllables preceding a devoiced vowel, as in [**kas**ɯ̥] 'scam', because whether or not the syllable with a devoiced vowel loses its syllabicity resulting in resyllabification is a controversial matter (see e.g. Kondo 1997; Matsui 2014; Starr & Shih 2014; Tsuchida 1997; Vance 1987, 2008).

[3][t] is affricated in front of [u] in Japanese.

The second type of heavy syllable contains a so-called moraic nasal, which is traditionally represented with /N/ (called *hatsuon* in the traditional grammar). Word-finally before a pause, this consonant is arguably realized as a uvular nasal, [ɴ], without much oral constriction, although its exact phonetic realization is extensively debated (Okada 1999, Vance 1987, pp.37-39, Vance 2008, pp. 95-105; see also section 5.2). When a stop follows this nasal consonant, it assimilates to the following stop in terms of place of articulation; e.g. /heN+da/ is realized as [he**nd**a], as exemplified more fully in Table 1.[4] Again, a sequence like /heN/—or [hen] and [hem]—is monosyllabic but bimoraic.

The third type of heavy syllable contains a long vowel, where the second portion is phonemically labeled as /R/ (or /H/); for example, the /kaR/ portion in /okaRsan/ is realized as [kaa]. The final type of heavy syllable contains a diphthong, a tautosyllabic sequence of two vowels (e.g. [ai] and [oi]). It is debatable which vowel sequences constitute a diphthong and which sequences constitute a hiatus in Japanese, but there is a general consensus that [ai] sequences are parsed into one syllable, thus forming a diphthong; see Kubozono (2015a) for recent discussion on this topic.

As Labrune (2012b) notes, most generative work assumes the characterizations of heavy syllables reviewed in this section; i.e. those syllables in Table 1 that contain /Q/, /N/, /R/, and /J/ constitute one prosodic unit—one heavy syllable. In fact, not only have they assumed the existence of heavy syllables, they have provided a body of explicit phonological evidence for the existence of syllable structures in the phonological organization of Japanese (e.g. Haraguchi 1996, 1999; Ito 1986, 1989, 1990; Ito & Mester 1992/2003, 1993, 2015; Kubozono 1989, 1996, 1999a,b, 2003, 2011, 2015a; Kubozono et al. 2008; McCawley 1968 among many others).[5]

Arguing against this commonly-held view in the generative tradition, Labrune (2012b) instead proposed to treat these heavy syllables as a sequence of a vowel ("a regular/full mora") and the following element (/Q/, /N/, /R/, /J/), the latter being collectively referred to as "a deficient/special mora". This proposal entirely does away with the notion of syllables, an idea which has also appeared in her recent book, Labrune (2012a). The current paper refers to this proposal by Labrune (2012a,b) as "the syllable-less theory" of Japanese.[6] In essence, the syllable-less theory of Japanese

[4]The behavior of moraic nasals before fricatives is debated. Vance (2008) says that it is realized as a nasalized dorso-velar approximant, [ɰ̃], while acknowledging that it is hard to transcribe (p.97). A recent electropalatography (EPG) experiment by Kochetov (2014), on the other hand, shows that moraic nasals assimilate in place and stricture to the following fricatives, just like pre-stop moraic nasals. The exact assimilative nature of moraic nasals before fricatives in Japanese is irrelevant to the main point of this paper, however.

[5]As can be seen in this list of the references, Haruo Kubozono should perhaps be given the most credit for establishing the existence of syllables in the prosodic organization of Japanese phonology. Contributions by Junko Ito and Armin Mester have also been significant and substantial. The arguments raised for the existence of syllables in Japanese by these authors are mainly phonological, and this reply instead focuses on experimental evidence. However, some of the phonological arguments discussed by these authors are briefly taken up in section 5 of this paper; nevertheless, readers are referred to these original works for details, should they be interested.

[6]Labrune (2012b) explicitly declares that her proposal is a revival of the view in the traditional study of Japanese, also known as *kokugogaku*. She states "[i]nterestingly, the Japanese linguistic tradition, which has a long history

decomposes heavy syllables into a sequence of two moras. One main basis of this proposal is the alleged lack of the phonetic interaction between the nucleus and the following element (/Q/, /N/, /R/, /J/). This paper thus starts by reviewing and reexamining the evidence that the vowel and the following tautosyllabic element do interact phonetically in Japanese.

3 Phonetic evidence for heavy syllables in Japanese

One argument that is raised for the syllable-less theory of Japanese is the "[a]bsence of phonetic clues for the existence of a rhyme-like constituent" (Labrune, 2012b, p.120), where "a rhyme-like constituent" refers to the combination of a vowel and any of the following /Q/, /N/, /R/, and /J/. In other words, the claim put forth by Labrune (2012b) is that there is no phonetic interaction between the nucleus and the following /Q/, /N/, /R/ and /J/. However, existing evidence suggests that there is phonetic interaction between the the nucleus and the following tautosyllabic element, contrary to the claim made by the syllable-less theory of Japanese.

3.1 The interaction between /N/ and the preceding vowel

Perhaps the most convincing case for the phonetic interaction between the coda consonant and the preceding vowel comes from the behavior of coda moraic nasals, /N/. The first type of evidence comes from the durational compensation effect between this nasal consonant and the preceding vowel within the rhyme (Campbell, 1999). In order to study how prosodic structures—including syllables—may affect durational patterns of various segments in Japanese, Campbell (1999) analyzed a corpus consisting of speech by four male and four female Japanese speakers, who read 503 phrases and sentences found in Japanese newspapers and magazines.

The result of this speech corpus analysis shows that the duration of a vowel and the duration of the following moraic nasal correlate negatively with each other—the longer the vowel, the shorter the consonant (pp. 34-35). More specifically, among back vowels, in particular, the higher the vowel, the shorter it is ([u]=100ms, [o]=109ms, [a]=115ms), and the moraic nasal /N/ undergoes lengthening next to higher vowels (post-[u]=90ms, post-[o]=87ms, post-[a]=83ms). These lengthening patterns of /N/ after higher vowels are significant, according to the t-tests reported in Campbell (1999).

This phonetic interaction is a typical durational compensation effect, which is observed commonly across many languages (e.g. Broselow et al. 1997; Campbell & Isard 1991; Lehiste 1970;

of remarkable achievements in the fields of philological description and analysis, has never felt the need to refer to a unit such as the syllable in opposition to the mora in accounts of Tokyo Japanese" (p. 114). This is at best an oversimplified statement. Joo (2008) (p. 217-218) lists several traditional Japanese linguists who have proposed the notion of syllables in addition to moras in Japanese, including Hattori (1951) and Shibata (1958). Joo (2008) himself endorses a position in which Japanese both has syllables and moras.

Port et al. 1980). Indeed, Campbell (1999) suggests that "[the negative correlation between the vowel's duration and the following consonant's duration] is consistent with the view that they both occupy a space within the same higher-level framework, accommodating to each other to optimally fill this frame" (p.35). The "higher-level frame" referenced here is the syllable or the rhyme (note that the title of Campbell's paper is, "A study of Japanese speech timing from the syllable perspective"). According to Campbell's data reported above, the durational target for the Japanese rhyme falls within the range of 190 ms to 200 ms.

A vowel and the following /N/ interact not only in terms of duration, but also in terms of quality. Vowels are nasalized before a nasal consonant within the same syllable (=/N/) in Japanese. The patterns of vowel nasalization in Japanese are illustrated in (1)-(2) (Campbell, 1999; Labrune, 2012b; Starr & Shih, 2014; Vance, 1987, 2008, 2013). Vowels are nasalized before a tautosyllabic nasal consonant, but vowels are not nasalized before an onset nasal consonant. Vance (2008), citing some classic descriptions of Japanese phonetics (Bloch, 1950; Jones, 1967; Nakano, 1969), states that "the vowel before [ɴ] is clearly nasalized [in Japanese]" (quoted from p. 96 with slight modification on transcription), whereas in a few pages later, he shows oral vowels following onset nasal consonants (e.g. the second diphthong in [sãmm**ai**] 'three sheets') (p.99).

(1) Vowels are nasalized before a tautosyllabic nasal

  a. [hõɴ] 'book'
  b. [hõn.da] 'Honda'
  c. [hõm.ma] 'Homma (personal name)'

(2) Vowels are not nasalized before an onset nasal

  a. [ho.ne] 'bone'
  b. [ko.me] 'rice'

The domain of vowel nasalization in Japanese can only be defined syllabically: vowels are nasalized before a tautosyllabic nasal. The domain of nasalization cannot be explained in terms of moras, because both in (1) and (2), the first vowels—the [o]s in (1) and (2)—consist of a mora preceding a nasal consonant, but only the former undergoes nasalization. We cannot explain away the nasalization pattern in terms of feet either; in both (1) and (2), the word-initial vowels are all parsed in the same foot as the nasal consonants; e.g. [(hõn)da] and [(hone)], assuming the uncontroversial bimoraic foot parsing pattern in Japanese (Poser 1990 *et seq.*, also embraced by Labrune 2012b). Finally, vowel nasalization cannot be explained in terms of a precedence relationship either, because not all vowels preceding nasal consonants are nasalized, as in (2): a syllable boundary blocks the regressive nasalization.

An anonymous reviewer went one step further and pointed out that vowels are not nasalized

after /ɴ/, either (e.g. the second vowel in [ãɴ.i] 'easy' remains oral). When a morpheme boundary follows the moraic nasal, it is realized as a coda consonant /ɴ/, without resyllabification (see section 5.2). Even if the next morpheme begins with a vowel, that vowel is not nasalized. This example precludes the explanation of Japanese vowel nasalization in which coda nasals are phonologically /N/ throughout the phonological derivation (e.g. [honda] is actually represented as [hoNda]), and only /N/, but not /n/ or /m/, causes nasalization. If /N/ were the trigger of nasalization without domain restrictions, both the preceding and following vowel should be nasalized; however, nasalization by /N/ is blocked in the following vowel, which is separated by a syllable boundary. This patterning of nasalization—[ṽɴ.v]—is therefore consistent with the view that the domain of vowel nasalization in Japanese is a syllable.

As Vance (2013) points out, Labrune (2012a,b) acknowledges this coda-induced nasalization process explicitly: the phonetic representations used in Labrune (2012b) show that nasalization applies to the vowel if and only if the vowel and the nasal are in the same syllable. For instance, Labrune's (2012b) examples (4) are transcribed as: [ãɴ.i] 'ease', [a.ni] 'old brother', and [ãn.ni] 'implicitly' (p. 122), in which all and only the vowels that are followed by a tautosyllabic nasal are nasalized. As Vance (2013: 172) notes, it is somewhat puzzling why this coda-induced nasalization was not considered to constitute evidence for the phonetic interaction between the nucleus and the following tautosyllabic element.

3.2 The interaction between /Q/ and the preceding vowel

Further suggestive evidence for the interaction between a vowel and a coda consonant exists, and we now look at how a vowel and other types of coda consonants interact, starting with /Q/—the first part of geminates. The first half of a geminate affects the quality and duration of the preceding vowel. For example, previous experimental studies have found that singleton consonants and geminate consonants show different f0 behavior in the preceding vowels in several different ways.

First, in order to investigate the phonetic properties of voicing and geminacy in Japanese, Kawahara (2006) recorded speech of three female native speakers of Japanese, pronouncing singleton stops and geminate stops in the /kVC(C)V/ word frame. The acoustic analysis shows that f0 is higher before geminates than before singletons. This result shows that /Q/ raises the f0 of the preceding vowel within the same syllable.

Second, Idemaru & Guion (2008) show that the fall in f0 due to pitch accent is greater across a geminate consonant than across a singleton consonant; i.e. the realization of pitch accent on the preceding vowel is affected by /Q/. Idemaru & Guion's (2008) study is based on the speech of six native speakers, three females and three males. Their acoustic study examined how a singleton-geminate contrast manifests itself in various acoustic dimensions. Among other contrasts, they

found that the f0 fall due to lexical accent in Japanese is much greater across geminates than across singletons: the mean differences between geminates and singletons were 32 Hz for voiced consonants and 29 Hz for voiceless consonants (p. 180) (see also Ofuka 2003 for a similar observation).

Third, for unaccented words, Fukui (1978) reports that f0 falls toward geminates in the preceding vowels. All of these examples show that the f0 patterns of a vowel are substantially affected by the presence of a tautosyllabic /Q/.

Next, as acknowledged by Labrune (2012b), vowels are longer when followed by a geminate consonant (i.e. /Q/) than when followed by a singleton consonant (pp.120-121), and this pattern has been documented by many acoustic studies on Japanese geminates (Campbell, 1999; Fukui, 1978; Han, 1994; Hirata, 2007; Idemaru & Guion, 2008; Kawahara, 2006; Port et al., 1987). According to the acoustic study of Idemaru & Guion (2008), for example, preceding vowels were on average 75 ms before geminates, whereas they were 59 ms before singletons (i.e. 16 ms difference) (p. 176). The median values reported by Campbell (1999) are 85 ms for the vowels in CV open syllables, and 105 ms for the vowels before /Q/ (i.e. 20 ms difference) (p. 35). Therefore /Q/ interacts with the preceding vowel in the same syllable by lengthening it.

Labrune (2012b) claims that this interaction does not instantiate a true phonetic interaction between the nucleus and the following tautosyllabic element, because we should expect shortening rather than lengthening before geminates (pp.120-121). However, pre-geminate lengthening is a phonetic interaction nevertheless, and not even a cross-linguistic anomaly, as other languages show similar pre-geminate lengthening: Finnish (Lehtonen 1970: 110-111, Yoshida et al. 2015), Persian (Hansen, 2004), Shinhala (Letterman, 1994), and Turkish (Jannedy, 1995; Lahiri & Hankamer, 1988). Just because vowels lengthen, instead of shorten, in front of a geminate, it does not mean that a vowel and /Q/ do not interact. They do interact, and the way they interact is not even cross-linguistically strange (see Kawahara 2015a for a recent review on this point).

The way /Q/ interacts with the preceding vowel may offer less strong evidence for a rhyme constituent than the way /N/ interacts with a vowel, because the patterns exhibited by /Q/ can potentially be modeled, as pointed out by an anonymous reviewer, in terms of a precedence relationship. However, the fact that /Q/ and the preceding vowel interact in so many different ways is suggestive of a constituent structure that /Q/ and the vowel form together.

Moreover, there is an observation that ties /Q/ and /N/ together: vowels are longer before a moraic nasal /N/ than they are in open syllables (Campbell, 1999, pp. 34-35): the median vowel durations reported in Campbell (1999) are 90 ms for open syllables and 110 ms when closed by a coda nasal (i.e. 20 ms difference) (p. 35). Hence /N/ lengthens a preceding tautosyllabic vowel, more or less to the same degree that /Q/ lengthens the preceding vowel (16 ms or 20 ms, as cited above). To the extent that /N/ provides strong evidence for the constituent structure of

the vowel and the coda, as discussed in section 3.1 above, and to the extent that /N/ and /Q/ show the same durational patterns with respect to the preceding vowel, it makes sense to conjecture that /Q/ also forms a constituent with the preceding vowel. All in all, the fact that both /Q/ and /N/ induce tautosyllabic vowel lengthening indicates that this lengthening is a general, syllable-based phenomenon (Campbell, 1999). In other words, Japanese vowels are phonetically longer (by about 20 ms) before a consonant within the same syllable.

3.3 Long vowels

There is also evidence that long vowels cannot be treated as a non-interacting sequence of two elements, a vowel and the following /R/, as in the syllable-less theory of Japanese. Hirata & Tsukada (2009) investigated the acoustic differences between short and long vowels in Japanese. They recorded four male native speakers of Japanese, pronouncing short and long vowels in /mV(V)mV(V)/ contexts (e.g. /mimi/, /miimi/, /mimii/).

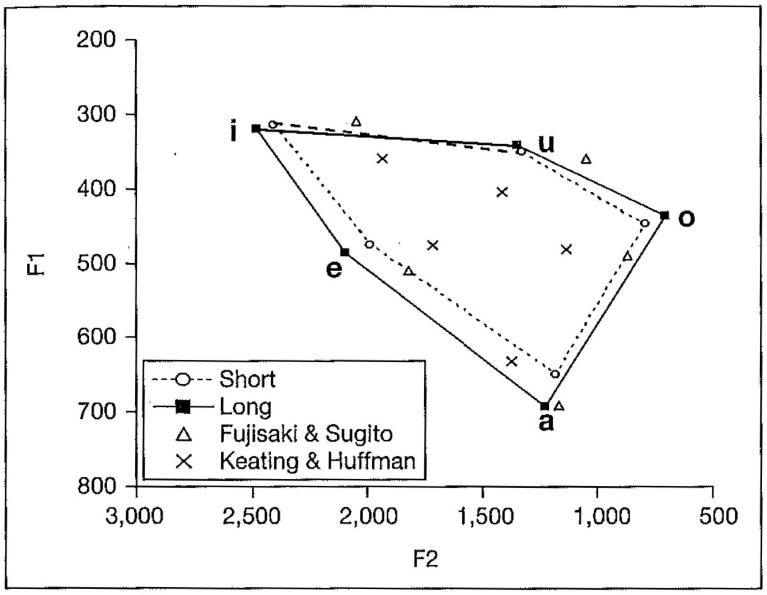


Figure 1: The results of Hirata & Tsukada (2009), based on their Figure 1 (p.137), showing the F1 and F2 values of the five Japanese vowels, short and long. The dashed lines represent the vowel space of short vowels. The solid lines represent the vowel space of long vowels. Fujisaki & Sugito (1977) and Keating & Huffman (1984) are two previous studies on vowel formant frequencies of Japanese short vowels. These data are plotted in the figure to guarantee that Hirata & Tsukada's F1 measurements are reliable, despite the presence of surrounding nasal consonants in their stimuli. This figure is reprinted with permission from S. Karger AG, Basel.

The study found that not only do short and long vowels differ in duration, but they also differ in formant frequencies. The result of their acoustic study is reproduced in Figure 1, where the dashed

lines and solid lines represent the vowel space of short and long vowels, respectively.

We observe that long vowels are generally more dispersed from each other in terms of F1 and F2 than short vowels, although not all the vowels are dispersed in the same way. [ii] and [ee] have higher F2 values than their short counterparts; [oo] has lower F2 than its short counterpart. [aa] has lower F1 than short [a]. Overall, [uu] is the only vowel whose formant frequency characteristics are almost identical to its short counterpart.

If the long vowels were represented simply as a sequence of a vowel and /R/, this dispersion effect would remain unaccounted for; i.e. pronouncing long vowels in Japanese is not simply a matter of "pronouncing a short vowel and lengthening it". Hirata & Tsukada (2009) furthermore argue that this dispersion effect cannot be relegated as a matter of a mechanical issue. To account for the dispersion effect in long vowels, one may raise an alternative theory, in which speakers cannot reach the formant targets in short vowels, because short vowels are too short for speakers to achieve their acoustic targets (i.e. short vowels involve articulatory undershoot: Lindblom 1963). This alternative is not viable, because Hirata & Tsukada (2009) did not find a comparable dispersion effect when the speakers produced short vowels at a slower speaking rate—if the speakers were showing undershoot for short vowels, they should show a similar dispersion effect when they speak slowly, because they would have enough time to reach their articulatory targets. However, this prediction was not borne out in the experiment.

In summary, then, the formant dispersion patterns of long vowels should be considered as a result of intentional articulatory control. In order to account for this observation, long vowels need to be treated as having different formant characteristics—or distinct articulatory targets—from short vowels. For this reason, long vowels cannot be treated as a mere sequence of a short vowel and /R/.

3.4 Summary

In summary, there is a non-negligible set of phonetic evidence for the interaction between the nucleus and the following tautosyllabic element, suggesting that Japanese has syllables.[7] At the

[7]To complete the discussion on the phonetics of heavy syllables in Japanese, I know of no phonetic evidence for syllables containing diphthongs; e.g. whether [ai] is phonetically different from [a.i]. Such evidence is hard to come by, partly because /VJ/ sequence is syllabified monosyllabically as [.Vi.] within a morpheme, and heterosyllabic parsing occurs across a morpheme boundary; i.e. [.V.#.i] (see Kubozono 2015a). Therefore, the effect of syllable boundary is confounded by the effect of a morpheme boundary; whenever there is a syllable boundary in this context, there is a morpheme boundary. In addition, the phonetic boundaries between any two vowels are very difficult to locate with precision because of their spectral continuity, and therefore, it is next to impossible to reliably measure, for example, the duration of [V] next to [i] (Turk et al., 2006).

Although not discussed in Labrune (2012b), in her recent book, Labrune (2012a) states that in a sequence of two vowels, "there is no significant gradual change of the quality of the first vowel towards the second one...contrary to what generally occurs with diphthongs in other languages" (p. 54), suggesting that two vowel sequences in Japanese are phonetically "more separable" than those in other languages. Labrune (2012a) thus implies that the alleged lack of

very least, the claim that there is no phonetic evidence for the interaction between a vowel and the following /Q, N, R/ is not correct.

4 Psycholinguistic evidence for syllables in Japanese

Labrune (2012b) makes a strong statement that "none of the many psycholinguistic studies which have been conducted has been able to establish the cognitive reality of the syllable in Japanese" (p.120). While it is probably true that the dominant segmentation pattern by Japanese speakers is mora-based (e.g. Otake et al. 1993),[8] this statement by Labrune (2012b) is too strong, and this section reexamines evidence that falsifies this claim. We will review psycholinguistic evidence for syllables from (i) the behavior of adult speakers, (ii) the behavior of Japanese children, and (iii) the text-setting patterns in Japanese songs.

4.1 Evidence from adult speakers: Speech production

Let us start with evidence for syllable-based segmentation patterns in adult speakers. To address the question of whether syllables affect the speech planning by Japanese speakers, Tamaoka & Terao (2004) visually presented two types of target stimuli: (i) trisyllabic, trimoraic nonce words (CVCVCV), and (ii) disyllabic, trimoraic nonce words (CVXCV where X is /Q/, /N/, /R/, or /J/). In addition, to get a baseline measure for simple disyllabic words, the experiment included disyllabic, bimoraic CVCV nonce words (see Table 2).

The task was to pronounce the visually presented words as quickly as possible. For the visual prompts, their first experiment used the *hiragana* orthography and the second experiment used the *katakana* orthography. In both writing systems, each letter corresponds to one mora. The experiment thus controlled for the number of letters across the two target conditions: both disyllabic and trisyllabic stimuli involved three letters (e.g. けっぺ and けたぺ in *hiragana*; コタモ and コーモ in *katakana*). Their stimuli also controlled for the quality of initial CV sequences and word-likeliness measures independently judged by a different set of speakers, because these are the two factors which were known to affect naming latencies rather substantially (see Tamaoka &

gradual changes between two adjacent vowels is evidence for lack of syllables in Japanese. Labrune (2012a), however, offers no quantitative support for this claim. Kawahara (2012b) argues based on spectrogram and waveform inspection that this claim is not well supported.

[8]Even this statement should be taken with caution. The work by Takashi Otake and his colleagues was influential in establishing the role of moras in Japanese speech segmentation; in later work, however, they weakened their view that moras play a dominant role in early speech processing in Japanese. For example, Cutler & Otake (2002) state that "[o]ur results indicate no role for morae in early spoken-word processing; we propose that rhythmic categories constrain not initial lexical activation but subsequent processes of speech segmentation and selection among word candidates (p. 296)."

Terao 2004 and Tamaoka & Makioka 2009 and references cited therein for extended discussion). Twenty-four native speakers of Japanese participated in these experiments.

The results of their Experiment I are summarized here as Table 2 (based on their Table 1, p.11). As we can observe, the trisyllabic, trimoraic words (the first row) showed the longest naming latencies, compared to the disyllabic words, trimoraic words (2nd-4th rows). In fact, the CVNVC condition showed shorter naming latencies than CVCV stimuli (the 5th row)—the naming latencies between these two conditions are at least very similar. Naming latencies therefore seem to correlate more with the number of syllables than with the number of moras. The error rate was also higher in the trisyllabic condition, compared to those for the disyllabic conditions.[9]

Table 2: The results of Tamaoka and Terao (2004), Experiment I.

Stimuli	Example	Visual stimuli	Naming latencies (ms)	Error rates (%)
1. CV.CV.CV	/ketape/	けたぺ	645	15.1
2. CVJ.VC	/keope/	けおぺ	590	7.8
3. CVQ.CV	/keQpe/	けっぺ	575	9.9
4. CVN.CV	/keNpe/	けんぺ	533	3.1
5. CV.CV	/kepe/	けぺ	537	1.6

The behavior of CVR syllables was only tested in Experiment II, whose results are partially reproduced here as Table 3 (their Table 2, p.17). The CVRCV stimuli showed shorter naming latencies and lower error rates than the CVCVCV stimuli, despite the fact that in both of the conditions, the stimuli had three moras.

Table 3: The results of Tamaoka and Terao (2004), Experiment II.

Stimuli	Example	Visual stimuli	Naming latencies (ms)	Error rates (%)
1. CV.CV.CV	/kotamo/	コタモ	645	2.78
2. CVR.CV	/koRmo/	コーモ	573	1.91

In summary, Tamaoka & Terao's (2004) experiments show that the disyllabic CVXCV stimuli showed shorter naming latencies and lower error rates than the trisyllabic CVCVCV stimuli, despite the fact that both conditions involve three moras and three letters. Based on these results, Tamaoka & Terao (2004) conclude that "/N/, /R/, /Q/, and /J/ are combined with a preceding

[9]The Associate Editor asked if the difference between the trisyllabic stimuli and disyllabic stimuli can be explained in terms of the number of segments. The experiment controlled for the number of letters used for the visual stimuli, but not the number of segments. This issue is hence not addressed by Tamaoka & Terao 2004, and it needs to be explored further in future experimentation. However, the result that the CVNCV stimuli and the CVCV stimuli patterned together is hard to reconcile with the segment-counting view, because the CVNCV stimuli contain an extra segment compared to the CVCV stimuli.

CV mora when native Japanese speakers pronounce visually presented non words, regardless of [whether they are] presented in the hiragana or katakana script. More specifically, syllabic units are...used for naming tasks requiring phonological production" (p.20).

In a follow-up study using 28 native speakers of Japanese, Tamaoka & Makioka (2009) replicated this finding, further showing that CVXCV disyllabic nonce words induced both shorter naming latencies and lower error rates than CVCVCV trisyllabic nonce words. In Experiment II, they found that CVNCV words are pronounced with shorter latencies and lower error rates than CVCVCV nonce words (654 ms vs. 735 ms; 1.49% vs. 7.44%) (p. 92). Their Experiment III shows a similar pattern in which CVRCV nonce words are pronounced with shorter latencies and lower error rates (643 ms vs. 737 ms; 4.17% and 11.46%) (p. 94).

4.2 Evidence from adult listeners: Perception

Nakamura & Kolinksy (2014) used a dichotic listening task to show the relevance of syllables in the speech perception by Japanese listeners. In this task, two different sounds were simultaneously presented to listeners—one sound to the left ear and the other to the right ear. It has been shown for languages other than Japanese that in such a dichotic listening paradigm, a sound or a sound unit can "migrate" from one ear to another (Kolinsky & Morais, 1996; Kolinsky et al., 1995). Nakamura & Kolinksy (2014) tested whether this migration can occur at several different prosodic levels for Japanese listeners, as illustrated in Table 4. For example, when presented with the stimulus [**ge**ɴru] to one ear and [hai**do**] to the other, the migration of the initial syllable would result in the perception of the real word in Japanese, [**gendo**] 'limit', if perceptual migration can occur at the syllabic level.

In their experiment, the task was for participants to decide whether or not they heard a target word—e.g. [gendo]. In the test condition, the target words occur after migration of sounds, as illustrated in Table 4 (based on their Table 2, p. 13). As a comparison, they included stimuli which should not result in the perception of the target words, even after migration (e.g. [**ge**ɴru] and [hai**ba**]). A total of 56 native speakers of Japanese participated in this listening experiment.

Their experimental results show that Japanese speakers do show evidence for migration at different levels: segmental, moraic, and syllabic, as shown by the high d' values in Table 4. For example, in the syllable condition, the listeners correctly identified the target words as present 90.6% of the time when the targets were present in the stimuli after migration (=hit rate); at the same time, they also correctly judged that there were no target words when the stimuli did not indeed include the target words 90.7% of the time (=correct rejection rate). These high hit rates and high correct rejection rates results in a high d'-value.

Moreover, they found that migration at the level of moras and syllables occurs more frequently than at the level of consonants, which shows that the migration at the higher level cannot be ac-

Table 4: Examples of the migration patterns tested and observed in the dichotic listening test by Nakamura & Kolinsky (2014). The migration of the units shown in bold would result in the perception of [gendo] 'limit'. The experiment used a total of five sets of pairs of this kind. Detectability of the target words, which can be perceived after the migration, was measured in terms of a detectability index, d' (Macmillan & Creelman, 2005).

Migration unit	Stimulus to one ear	Stimulus to the other ear	Detectability (d')
consonant	**g**airu	h**endo**	2.04
vowel	h**e**iru	**g**a**ndo**	2.90
mora	**ge**iru	ha**ndo**	2.76
syllable	**geN**ru	hai**do**	2.99

counted for in terms of the combination of segmental migrations. Most importantly in this context, the illusionary migration can occur at the syllabic level—/geN/ and /haJ/ can switch with one another in the dichotomous listening task.

4.3 Evidence for syllables from children's behavior

The psycholinguistic evidence for syllables in Japanese comes not only from adults' behavior, but also from children's behavior. Inagaki et al. (2000) report an experiment which addresses the issue of whether Japanese speakers make use of syllables in their speech segmentation. Since the Japanese *kana*-orthography is mora-based in that one letter usually corresponds to one mora,[10] it is reasonable to suspect that speech segmentation pattern by adult speakers is substantially influenced by this mora-based writing system. Inagaki et al. (2000) thus tested the effect of literacy acquisition on speech segmentation.

Their task was a "vocal-motor" task in which Japanese-speaking children made counting gesture as they produced stimulus words. In this particular experiment, "[c]hildren were presented words...and asked to make a doll jump on a series of different colored circles on paper while they articulated the given words" (p. 75). For example, a word like [kurejoɴ] 'crayon' can be counted with three gestures if it is segmented in terms of syllables ([ku.re.joɴ]), but with four gestures if it is segmented in terms of moras ([ku-re-jo-ɴ]). The children first went through a practice with only CV syllables, which would not bias them to make counting gestures based on syllables or moras. The stimuli in the main session included words with only light syllables, words with CVN (e.g. [ku.re.joɴ]), words with CVQ (e.g. [rap.pa] 'bugle'), and words with CVR (e.g. [çi.koo.ki] 'airplane'). They ran two experiments using different sets of stimuli. The stimuli were presented in the form of drawings. The participants were kindergardeners, with different proficiencies lev-

[10]The only exception is three letters expressing palatalization of preceding consonants (ゃ、ゅ、ょ), which do not count as one mora.

els in terms of the Japanese orthographic systems. For Experiment I, the age range was 58-80 months, and the average age was 66 months; for Experiment II, the age range was 54-78 months; the average age was 66 months.

Their results show that Japanese children exhibit a mixture of syllable-based counting and mora-based counting. The results of their Experiment II are summarized in Table 5 (based on their Table 2, p.81). The youngest children (=Level 1) show both syllable-based parsing and mora-based parsing for all the types of syllables. As their level of acquisition of Japanese orthography goes up, the numbers of the syllable-based parsing goes down, except for the /CVQ/ syllables, which show persistent syllable-based counting patterns.

Table 5: The results of Experiment II of Inagaki et al. (2000), which show the number of observed counting patterns (either mora-based or syllable-based). Levels refer to the degree of the acquisition of the Japanese orthography (see the original paper for details).

Syllable types	Segmentation patterns	Level 1 (N=22)	Level 2 (N=10)	Level 3 (N=10)	Level 4 (N=10)
CVN	Mora-based (CV.N)	16	9	10	10
	Syllable-based (CVN)	6	1	0	0
CVQ	Mora-based (CV.Q)	11	7	2	2
	Syllable-based (CVQ)	11	3	8	8
CVV	Mora-based (CV.V)	13	8	9	9
	Syllable-based (CVV)	9	2	1	1

To be concrete, units such as [joN], [rap], and [koo] were sometimes associated with one counting gesture. Thus, Japanese children may deploy syllable-based counting gestures, at least those who are not very familiar with the Japanese orthography system.[11] The results furthermore showed that as they learn Japanese orthography, the mora-based parsing becomes more dominant. This experiment overall shows that the dominance of mora-based parsing is partly due to the Japanese orthography.

Proponents of the syllable-less theory may suggest that how Japanese pre-literate children show segmental parsing effects is not relevant to the organization of adult Japanese prosody; after all, adults speakers do show dominant mora-based segmental parsing patterns, after they learn the Japanese orthographic system (e.g. Otake et al. 1993, among others; though see Cutler & Otake

[11]The Associate Editor asked if this "syllable-based" counting pattern can be explained also by resorting to "vowel-based counting" or "sonority-peak-based counting". The patterns of long vowels show, however, that the children may not be counting vowels in the "syllable-based" parse, because in this condition, there are two vowels. It is hard to exclude the "sonority-peak-based counting" hypothesis, because syllable structures and sonority peaks match exactly in Japanese. Even if the children were counting sonority peaks, that comes very close to saying that they were counting syllables—at least they were not counting moras. What is crucial here is why children could ignore moraic nasal /N/ or /R/ and count syllables/vowels/sonority peaks, if mora was the only psycholinguistic unit in Japanese.

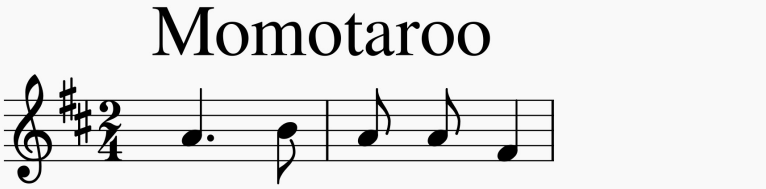
2002 who doubt this claim—see footnote 8). However, if there is no evidence for syllables in the phonetics and phonology of Japanese, how do Japanese-learning children acquire syllable-based parsing? In other words, if Japanese is entirely syllable-less, where does the syllable-based parsing pattern in Japanese children come from? One could argue that the syllable-based parsing pattern may come from the innate, universal mechanism/bias rather than from learning the phonetics and phonology of Japanese. However, this universalist view of syllables is exactly what is argued against by Labrune (2012a,b).

4.4 Evidence from text-settings in Japanese songs

Finally, Labrune (2012b) mentions that "[t]he mora is the metric unit of Japanese verse in poetry and singing" (p. 116). This is indeed a commonly-held view, and Japanese *haiku* is a famous example in which metrical units are based on moras. However, here too again, the story is not this simple, and syllables can and do play a role. There are cases in which one bimoraic heavy syllable can be associated with one, rather than two, musical notes.

An illustrative example is shown in (3) (adapted from Vance 1987: 68, based on Iizuka Shoten Henshuubu 1977: 85). This song *Momotaroo* is one of the most famous children's songs in Japanese. In verse 1, line 1, the last two musical notes are each associated with a heavy syllable /roo/ and /saN/, respectively. As shown in verse 2, line 1, it is not necessarily the case that one syllable needs to correspond with one musical note: /ʃoo/ in (3) is split up into two notes, /ʃo/ and /o/. It thus seems that text-setting in Japanese can be based on syllables, as well as moras.

(3) Textsetting in the "Momotaroo" song.

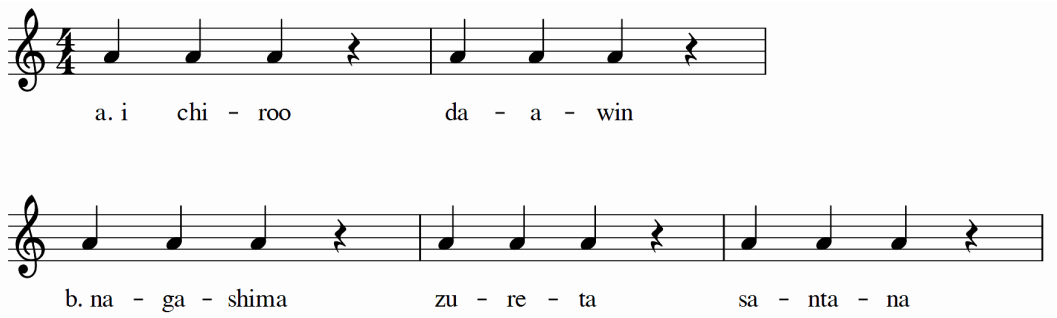


Beyond this simple example, a more systematic study has been recently conducted by Starr & Shih (2014), which shows that Japanese speakers can generally associate a heavy syllable to one musical note. Their study builds on a body of studies on how text-setting is achieved in Japanese singing traditions (Kubozono, 1999a; Manabe, 2009; Sugito, 1998; Tanaka, 2000, 2012) (all cited and discussed by Starr & Shih 2014).

Starr & Shih (2014) studied a corpus of translated Disney songs as well as native songs. They distinguished two types of text-setting patterns: (i) "syllabic" if a special mora is associated with the note together with the preceeding mora (i.e. when the entire heavy syllable is associated with one musical note), and (ii) "moraic" if a special mora is assigned to its own musical note (see (3) for these two types of text-setting). They found that syllables with /N/, /R/, and /J/ can each exhibit syllabic text-setting, both in translated songs and native songs. Syllables with /N/ and /R/ were particularly susceptible to syllabic text-setting, and in the translated Disney songs, in fact, syllabic text-setting was found about 50% of the time. Their study thus shows that text-setting at the syllabic level is not only possible, but in fact common.

Yet another interesting piece of evidence for syllables in text-setting comes from chanting patterns in baseball games (originally discovered by Tanaka 1999, cited and discussed by Kubozono 2015a). The musical template of this baseball chanting is [kattobase XXX] where [kattobase] means 'hit (a home-run)', and names of the players are mapped on to three musical notes, represented by XXX. What is interesting is when the players' names are four mora long: the last musical note is always associated with the last syllables, whether they be heavy or light. The illustrative figures are reproduced from (Kubozono 2015a: 223):

(4) Japanese baseball changing patterns. Actual notes can differ depending on the accent properties of the last names, which are abstracted away from.



The generalization about which portions of the names can be associated with the last musical note can only be stated in terms of syllables. In terms of moras, the examples in (4a) involve the last two moras, whereas the examples in (4b) involve only the last mora. In terms of feet, the examples in (4a) involve the whole final foot (e.g., (/(iti)(roo)/), whereas the examples in (4b) involve only a part of the final foot, given the standard bimoraic foot parsing pattern (e.g. /(naga)(sima)/).

4.5 Summary

All in all, it may appear at first glance that moras are very important in psychological segmentation in Japanese, so much so that the role of syllable is hard to see. Under careful experimentation and examination of the native speakers' speech behavior, however, the role of syllables becomes clear. Needless to say, I am not denying the role of moras in the speech segmentation of Japanese. Moras surely do play a role, but so do syllables.

5 Some remarks on phonological and theoretical issues

Although the focus of this reply has been the phonetic and psycholinguistic evidence for syllables in Japanese, I would like to respond to a few points regarding Japanese phonology that Labrune (2012b) makes. The specific issues addressed in this section include (i) the phonological non-isomorphism between moras and syllables in Japanese, (ii) the alleged lack of onset maximization in Japanese, (iii) evidence for syllables from the phonological minimality requirement, and (iv) the Occam's Razor argument used by Labrune (2012b).

As Ito & Mester (2015) point out, the reanalyses of syllable-based phenomena by Labrune (2012b) are simply recasting the distinction between head moras and non-head moras in syllable theory with "full moras" and "deficient moras": to borrow the words of Ito and Mester, "by eschewing the syllable, proponents of the syllable-less theory must posit different types of moras with different properties, recapitulating syllable theory in a different terminology, but unfortunately within a network of assumptions entirely specific to Japanese" (p. 371). In other words, the reanalyses by Labrune (2012b) show that the patterns previously analyzed with syllables *can* be analyzed without syllables, but it is not necessarily the case that they *must* be, and Labrune (2012b) herself carefully seems to admit this point.[12] For this reason, the experimental evidence reviewed in the previous two sections should suffice to support the evidence for syllables in Japanese.[13]

[12]For example, when Labrune (2012b) discusses tonal phrase-initial lowering patterns (p.123), she says that the syllable-based analysis "does not have more explanatory power than" the syllable-less theory, but she avoids saying that the syllable-less explanation is *better*. Moreover, at the end of section 4, she says "this [syllable-less] analysis is not uncontroversial—some readers might still feel that...[some of the phenomena discussed in the paper] make a case for the syllable, or that the cost of accepting that some languages might lack syllables would be too high for phonological theory in particular and for the theory of universals in general (p.134)".

[13]Besides the alleged lack of phonetic and psycholinguistic evidence for syllables, which has already been addressed in sections 3 and 4 of this reply, there are two more points that Labrune (2012b) raised against the existence of syllables in Japanese. First, Labrune (2012b) cites Poser (1990) and argues that it is problematic that syllable boundaries and foot boundaries may not coincide (p.122). In footnote 9 of Labrune (2012b), however, it is acknowledged that this problem disappears once a certain assumption, which is a plausible one in current phonological theories, is made. Second, Labrune (2012b) argues that speech errors are dominantly mora-based rather than syllable-based (p.120). However, it is not case that syllable-based speech error patterns are impossible. For instance, Kubozono (1989) reports an example like [se.kai.rem.**poo**.ʃim.buɴ] → [se.kai.rem.**buɴ**.ʃim.buɴ] 'World Federation Newspaper', in which a heavy syllable [poo] is replaced by another heavy syllable [buɴ] (p. 252).

Nevertheless I believe that there are a few phonological and theoretical issues which would merit explicit response from the proponents of the syllables in Japanese, which I take up on in this section. See also Ito & Mester (2015) and Tanaka (2013).

5.1 Phonological non-isomorphism between moras and syllables

The first point is concerned with the phonological non-isomorphism between moras and syllables. Although it is true that the role of moras is very important in Japanese phonology, we should not ignore the observation that sometimes, one heavy syllable patterns with one light syllable, rather than a sequence of two light syllables, contrary to what the syllable-less theory predicts. Another important observation is that a sequence of two moras can pattern differently, depending on whether they are parsed into one syllable or not.[14]

One illustrative case that instantiates the first observation is the accentual pattern of /X-taroo/ personal name compounds. As Kubozono (1999b,c, 2001) shows, the whole compound takes an unaccented form when the initial element of the compound (=E1) is monosyllabic: crucially, it does not matter if the E1 is a heavy syllable (two moras) or a light syllable (one mora), as shown in (5) and (6). In contrast, when the E1 is two light syllables, they receive final-accent on E1, as in (7) (accent is shown with an apostrophe mark after the accented vowel). If Japanese phonology solely operated on moras, then the prediction would be that (5) and (7) should pattern together, because the E1s are bimoraic in both conditions.

(5) E1=one heavy syllable (two moras): unaccented

a. [kin-taroo]
b. [kan-taroo]
c. [koo-taroo]
d. [hoo-taroo]
e. [kjuu-taroo]
f. [soo-taroo]

(6) E1=one light syllable (one mora): unaccented

a. [a-taroo]
b. [ki-taroo]
c. [ko-taroo]
d. [ja-taroo]

(7) E1=two syllables (two moras): accent on the final syllable of E1

a. [momo'-taroo]

[14] I thank an anonymous reviewer for bringing this to my attention.

b. [mitʃi'-taroo]
c. [itʃi'-taroo]
d. [asa'-taroo]

Labrune (2012b) does mention this pattern, and first argues that this pattern is "lexical rather than strictly phonological" (p. 131). Then Labrune (2012b) moves on to argue that the syllable-less theory can state the environment in which unaccented compounds appear: "when the first member is equivalent to a monomoraic foot or to a bimoraic foot ending in a special mora" (p. 131). This reanalysis, however, is merely a restatement of the descriptive facts, and the disjunctive statement (the use of "or") is hardly better than the syllable analysis presented by Kubozono (1999b,c, 2001).

Another argument against the purely moraic view can be made based on patterns of loanword adaptation. As Kubozono (1996, 1999b, 2011, 2015b) points out, given four-mora loanwords, the preferred accentual patterns differ depending on their syllabic composition. More concretely, if the final two moras constitute a heavy syllable, the words tend to be accented; on the other hand, if the final two moras are both light syllables, then the default pattern is unaccented (see also Kawahara 2015b for further discussion).

(8) LLH=Accented on the initial syllable

a. [a'.ma.zoɴ] 'Amazon'
b. [do'.ra.ɡoɴ] 'dragon'
c. [re'.ba.noɴ] 'Lebanon'
d. [te'.he.raɴ] 'Teheran'
e. [he'.ru.paa] 'helper'
f. [ʃa'.ru.pii] 'Charpy (impact factor)'
g. [se'.ra.pii] 'therapy'

(9) LLLL=Unaccented

a. [a.me.ri.ka] 'America'
b. [mo.na.ri.za] 'Mona Liza'
c. [a.ri.zo.na] 'Arizona'
d. [ma.ka.ro.ni] 'macaroni'
e. [i.ta.ri.a] 'Italy'
f. [bu.ra.ʤi.ru] 'Brazil'
g. [me.ki.ʃi.ko] 'Mexico'

The syllable-less theory would not be able to distinguish the examples in (8) and (9), because both types of words contain four moras and two feet.

The bottom line message from the works by Kubozono discussed in this section boils down to

two general observations: (i) heavy syllables and light syllables can pattern together, and (ii) two mora sequences can behave differently depending on whether these moras are parsed into the same syllable or not. The notion of syllables is crucial to capture both of these observations.

5.2 Japanese maximizes onsets

The second issue is concerned with the proposed lack of onset maximization. Labrune (2012b) argues that Japanese lacks "onset optimization" (p.121-122), and this constitutes evidence for the lack of syllables in Japanese. What is meant by the "the lack of onset optimization" is a lack of resyllabification across a morpheme boundary, or what is more generally known as "onset maximization" (Clements & Keyser, 1983; Ito, 1986, 1989; Kahn, 1976; Prince & Smolensky, 2004; Steriade, 1982). Contrary to the asserted lack of onset maximization, Japanese does usually maximize onsets (Ito, 1986, 1989), so that underlying monomorphemic sequences like /ana/ 'whole' are syllabified as [a.na] rather than [aɴ.a].[15] Japanese even shows evidence for resyllabification across a morpheme boundary in the verbal patterns of stems that end with a consonant (Ito & Mester, 2015), as shown in Table 6. In this sense, Japanese does maximize onsets and avoid codas.

Table 6: Resyllabification and onset maximization in Japanese verbal inflectional paradigms.

	morpheme composition	syllabification	gloss
negative	/na**k+a**nai/	[na.**ka**.nai]	'not cry'
polite	/na**k+i**masu/	[na.**ki**.ma.su]	'cry (polite)'
non-past	/na**k+u**/	[na.**ku**]	'cry'
conditional	/na**k+e**ba/	[na.**ke**.ba]	'if cry'
volitional	/na**k+o**o/	[na.**koo**]	'let's cry'

What Labrune (2012b) discusses is the lack of syllabification across a stem boundary in Sino-Japanese compounds; e.g. /aN/ 'safe'+/i/ 'easy' is syllabified as [ãɴ.i], rather than [a.ni] (Ito & Mester, 1996, 2015; Kurisu, 2000). However, the prohibition against resyllabification across a stem boundary is hardly surprising: it is a cross-linguistically well-observed alignment effect, in which syllable boundaries and stem boundaries are required to be aligned with each other (McCarthy &

[15] One could raise a concern here that I am arguing for a syllabification pattern based on my own intuition, rather than attempting to show independent evidence. However, we know that the first syllabification—[a.na]—is the right analysis, as the second syllabification—[aɴ.a]—does occur when a morpheme boundary follows the nasal consonant. In other words, both [a.na] and [aɴ.a] are possible syllabifications in Japanese, but the latter occurs only when a morpheme boundary is present; nevertheless, the default syllabification pattern is the former pattern with onset maximization. More generally, when Japanese speakers divide words into smaller phonological chunks, they do maximize onsets: /arigatoR/ 'thank you' is separated into chunks as [a]-[ri]-[ga]-[to]-[o], rather than [ar]-[ig]-[at]-[o]-[o]. See also Ito (1989: 223) who argues that the Onset Principle, which avoids onsetless syllables, is operative in Japanese, although its requirement is not "absolute", but instead "relative".

Prince, 1993).

5.3 Phonological minimality

One argument that has been advanced for the role of syllables in Japanese comes from the prosodic minimality requirement (Ito, 1990; Ito & Mester, 1992/2003, 2015). The observation is that in loanword truncation, words can be truncated to disyllabic, bimoraic forms, as in (10). However, this truncation does not yield monosyllabic, bimoraic forms; when the initial syllables of the base forms are heavy, the truncated form takes an extra light syllable, as in (11).

(10) Bimoraic truncation patterns

a. [de.mon.su.to.ree.ʃoɴ] → [de.mo] 'demonstration'
b. [ri.haa.sa.ru] → [ri.ha] 'rehearsal'
c. [ro.kee.ʃoɴ] → [ro.ke] 'location'
d. [bi.ru.diŋ.gu] → [bi.ru] 'building'
e. [bu.ra.ʒaa] → [bu.ra] 'brassiere'
f. [pu.ro.ɸeʃ.ʃo.na.ru] → [pu.ro] 'professional'

(11) Monosyllabic outputs are not allowed: a light syllable is appended

a. [mai.ku.ro.ɸo.oɴ.] → [mai.ku], *[mai] 'microphone'
b. [dai.ja.mon.do] → [dai.ja], *[dai] 'diamond'
c. [paa.ma.nen.to] → [paa.ma], *[paa] 'permanent (hair style)'
d. [kom.bi.nee.ʃoɴ] → [kom.bi], *[koɴ] 'combination'
e. [am.pu.ri.ɸai.aa] → [am.pu], *[aɴ] 'amplifier'
f. [ʃim.po.ʤi.u.mu] → [ʃim.po], *[ʃiɴ] 'Symposium'

This truncation pattern shows that Japanese prosodic minimality is defined based on the number of syllables in addition to the well-known bimoraic requirement—one syllable, despite being bimoraic, is too small (Ito, 1990; Ito & Mester, 1992/2003, 2015).

Labrune (2012b) discusses this pattern (p.126-128), and argues that this apparent prohibition against monosyllabic forms should instead be attributed to a general ban against heavy syllables (or /N/, /R/ or /J/ in the syllable-less theory) at the end of a prosodic word. This argument, however, faces an empirical problem, because when compounds are truncated with an additional morpheme, truncated forms can end with heavy syllables, as in (12).

(12) Examples of truncated words ending with a heavy syllable

a. [ku.so] + [gee.mu] → [ku.so.**gee**] 'crappy game'
b. [mo.bai.ru] + [gee.mu] → [mo.ba.**gee**] 'mobile game (a company name)'

c. [a.ra.un.do] + [saa.tii] → [a.ra.**saa**] ‘around thirty’
d. [a.ra.un.do] + [ɸoo.tii] → [a.ra.**ɸoo**] ‘around forty’
e. [ʤi.mii] + [hen.do.rik.ku.su] → [ʤi.mi.**heɴ**] ‘Jimmy Hendrix’
f. [go.ku.doo] + [sen.see] → [go.ku.**seɴ**] ‘Yakuza teacher (TV drama)’

No examples of bi-morpheme truncation that I know of take an extra light syllable when they end with a heavy syllable, unlike the mono-morphemic truncation in (11) (i.e. *[ara**saati**]).[16] To eschew this problem, Labrune (2012b) argues that different principles operate on mono-morphemic truncation patterns and bi-morphemic truncation patterns.

However, the syllable-based minimality explanation is superior in that it automatically explains why bi-morphemic truncated compounds can end in heavy syllables, without postulating that mono-morphemic truncation and bi-morphemic truncation are governed by different principles: bi-morphemic compounds can result in heavy syllables, because one morpheme can supply enough material so that the minimality requirement at the word level is independently satisfied.

One remaining observation to be accounted for is the lack of light-light-heavy (LLH) forms in the mono-morphemic truncation pattern, which Labrune (2002, 2012b) argued to be absent. However, it is questionable as to whether there are many base forms in the first place that are long enough so that truncation results in light-light-heavy forms (i.e., mono-morphemic words of the form LLHX). Moreover, there do exist a number of examples that instantiate this pattern, as in (13):

(13) Examples of monomorphemic LLH truncation

a. [pu.re.zen.tee.ʃoɴ] → [pu.re.zeɴ] ‘presentation’
b. [hi.po.kon.de.rii] → [hi.po.koɴ] ‘hypochondriasis’
c. [sa.pi.en.ti.a] → [sa.pi.eɴ] ‘Sapientia’
d. [o.ri.en.tee.ʃoɴ] → [o.ri.eɴ] ‘orientation’
e. [re.pu.re.zen.tiŋ.gu] → [re.pe.zeɴ] ‘representing (rap jargon)’
f. [re.ko.men.dee.ʃoɴ] → [re.ko.meɴ] ‘recommendation (name of a radio show)’
g. [se.pu.tem.baa] → [se.pu.teɴ] ‘September student (ICU jargon)’

These examples show that monomorphemic LLH truncation pattern is in fact not impossible. These examples also show that the ban against monosyllabic truncation shown in (11) results from a syllable-based minimality requirement (Ito, 1990), rather than a general ban against heavy syllables

[16] An anonymous reviewer pointed out two potential counter-examples from base-ball terms: [pa-riigu] from [paʃiɸikku riigu] ‘Pacific League’ and [se-riigu] from [sentoraru riigu] ‘Central League’. However, these are not strong counterexamples, because these do not involve truncation of the second elements. These examples are also exceptional in that the first elements leave only one mora rather than two; the first elements look as if they were prefixes. These words should thus probably be viewed as prefixation.

word-finally (Labrune, 2012b).

5.4 On the Occam's razor argument

Ultimately, it is hard to prove the absence of anything in linguistics or elsewhere (though see Gallistel 2009). Labrune (2012b) fully acknowledges this difficulty: "Of course, lack of positive evidence does not automatically provide negative evidence", yet Labrune (2012b) goes on to say "but it should at least lead us to question the initial postulate that Japanese is a syllable language (p.119)." Behind this logic is Occam's razor—everything else being equal, we should not posit a theoretical device, here syllables, unless there is explicit need to do so.[17] This paper has argued throughout that there is indeed phonetic, phonological and psycholinguistic evidence to posit syllables, but even if these arguments were not to hold, it is important to bear in mind that everything else is *not* equal: it has consequences for other aspects of phonological theorization.

To be concrete, from the perspective of language acquisition, it would be simpler if learners can start the language acquisition process with the assumption that every language has syllables; it eliminates their task of discerning whether the target language has syllables or not (see Ito & Mester 2007, 2012, 2013; Kawahara 2012a; Kawahara & Shinya 2008; Selkirk 2005 for related discussion; cf. Hyman 1985, 2008, which are discussed by Labrune 2012b, and Jun 2005, 2006 for views that prosodic hierarchies are language-specific). Another argument is that a linguistic theory which does not admit language-specificity in terms of prosodic levels is more restrictive. As Ito & Mester (2007) succinctly put it, "[a] universal hierarchy cannot easily admit language-specific gaps" (p.97), and there have been attempts to establish a universal theory of prosodic hierarchy in the face of apparent language-specific patterns (Ito & Mester, 2007, 2009, 2012, 2013; Kawahara & Shinya, 2008; Selkirk, 2005, 2011). I do not mean to imply that these conceptual arguments alone should suffice to assume the universality of syllables; it is intended as a warning against the use of Occam's Razor argumentation in this context.

6 General conclusion

In conclusion, Japanese has syllables. There is evidence for syllables from both phonetic and psycholinguistic points of view. Phonological consideration also suggests that Japanese has syllables. Although the syllable-less theory of Japanese phonetics and phonology is very thought-provoking, evidence remains that indicates Japanese has syllables.

[17] On p.119, where the relevant quote is discussed, Labrune (2012b) does not explicitly mention Occam's Razor. An explicit reference to Occam's Razor appears on p.139, however. As discussed below, Occam's Razor argumentation should be used with caution in linguistic theorization in general, because it is barely the case that everything else is the same—eliminating one theoretical apparatus often has non-negligible consequences elsewhere.

Acknowledgements

I am very grateful for the critical yet constructive comments that I received from three anonymous reviewers as well as the Associate Editor. They helped me flesh out the arguments that I presented in the previous version of this paper. Comments from the following people were also helpful in either developing my initial ideas and/or addressing the comments that I received from the AE and the reviewers: Robert Daland, Donna Erickson, Osamu Fujimura, Haruka Fukazawa, Dylan Herrick, Junko Ito, Takashi Otake, Mayuki Matsui, Michinao Matsui, Armin Mester, Jeff Moore, Miho Sasaki, Yoko Sugioka, Helen Stickney, and Yukiko Sugiyama. Preparation of this paper is partly supported by two JSPS Kakenhi grants #26770147 and #26284059. Remaining errors are mine.

References

Bloch, Bernard (1950) Studies in Colloquial Japanese, IV: Phonemics. *Language* **26**: 86–125.

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

Campbell, Nick (1999) A study of Japanese speech timing from the syllable perspective. *Onsei Kenkyu [Journal of the Phonetic Society of Japan]* **3**(2): 29–39.

Campbell, Nick & Stephen Isard (1991) Segment durations in a syllable frame. *Journal of Phonetics* **19**: 37–47.

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

Cutler, Anne & Takashi Otake (2002) Rhythmic categories in spoken-word recognition. *Journal of Memory and Language* **46**(2): 296–322.

Fujisaki, Hiroya & Miyoko Sugito (1977) Onsei-no butsuriteki seishitsu [Physical natures of sounds]. In *On-in [Iwanami Kouza Nihongo 5]*, Tokyo: Iwanami, 63–103.

Fukui, Seiji (1978) Nihongo heisaon-no encho/tanshuku-niyoru sokuon/hisokuon toshite-no chooshu [Perception for the Japanese stop consonants with reduced and extended durations]. *Onsei Gakkai Kaihou [Phonetic Society Reports]* **159**: 9–12.

Gallistel, Randy C. (2009) The importance of proving the null. *Psychological Review* **116**(2): 439–453.

Han, Mieko (1994) Acoustic manifestations of mora timing in Japanese. *Journal of the Acoustical Society of America* **96**: 73–82.

Hansen, Benjamin B. (2004) Production of Persian geminate stops: Effects of varying speaking rate. In *Proceedings of the 2003 Texas Linguistics Society Conference*, Augustine Agwuele, Willis Warren, & Sang-Hoon Park, eds., Somerville: Cascadilla Press, 86–95.

Haraguchi, Shosuke (1996) Syllable, mora and accent. In *Phonological Structure and Language Processing: Cross-Linguistic Studies*, Takeshi Otake & Anne Cutler, eds., Berlin: Mouton de Gruyter, 45–75.

Haraguchi, Shosuke (1999) Accent. In *The Handbook of Japanese Linguistics*, Natsuko Tsujimura, ed., Oxford: Blackwell, 1–30.

Hattori, Shiroo (1951) *Onseigaku [Phonetics]*. Tokyo: Iwanami.
Hirata, Yukari (2007) Durational variability and invariance in Japanese stop quantity distinction: Roles of adjacent vowels. *Onsei Kenkyu [Journal of the Phonetic Society of Japan]* **11**(1): 9–22.
Hirata, Yukari & Kimiko Tsukada (2009) Effects of speaking rate and vowel length on formant frequency displacement in Japanese. *Phonetica* **66**(3): 129–149.
Hyman, Larry (1985) *A Theory of Phonological Weight*. Dordrecht: Foris.
Hyman, Larry M. (2008) Universals in phonology. *The Linguistic Review* **25**: 83–137.
Idemaru, Kaori & Susan Guion (2008) Acoustic covariants of length contrast in Japanese stops. *Journal of International Phonetic Association* **38**(2): 167–186.
Iizuka Shoten Henshuubu (1977) *Nihon Shooka Dooyoo Shuu [Collection of Japanese Children's Songs]*. Tokyo: Iizuka.
Inagaki, Kayoko, Giyoo Hatano, & Takashi Otake (2000) The effect of kana literacy acquisition on the speech segmentation unit used by Japanese young children. *Journal of Experimental Child Psychology* **75**: 70–91.
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. In *Proceedings of Chicago Linguistic Society: Parasession on the Syllable in Phonetics and Phonology*, Michael Ziolkowski, Manual Noske, & Karen Deaton, eds., Chicago: Chicago Linguistic Society, 213–239.
Ito, Junko & Armin Mester (1992/2003) Weak layering and word binarity. In *A New Century of Phonology and Phonological Theory*, Takeru Homma, Masao Okazaki, Toshiyuki Tabata, & Shin-Ichi Tanaka, eds., Tokyo: Kaitakusha, 26–65.
Ito, Junko & Armin Mester (1993) Licensed segments and safe paths. *Canadian Journal of Linguistics* **38**(2): 197–213.
Ito, Junko & Armin Mester (1996) Stem and word in Sino-Japanese. In *Phonological Structure and Language Processing: Cross-Linguistic Studies*, T. Otake & A. Cutler, eds., Berlin: Mouton de Gruyter, 13–44.
Ito, Junko & Armin Mester (2007) Prosodic adjunction in Japanese compounds. *Proceedings of Formal Approaches to Japanese Linguistics* **4**: 97–112.
Ito, Junko & Armin Mester (2009) The onset of the prosodic word. In *Phonological Argumentation*, Steve Parker, ed., London: Equinox, 227–260.
Ito, Junko & Armin Mester (2012) Recursive prosodic phrasing in Japanese. In *Prosody Matters*, Toni Borowsky, Shigeto Kawahara, Takahito Shinya, & Mariko Sugahara, eds., London: Equinox Publishing, 280–303.
Ito, Junko & Armin Mester (2013) Prosodic subcategories in Japanese. *Lingua* **124**: 20–40.
Ito, Junko & Armin Mester (2015) Word formation and phonological processes. In *The Handbook of Japanese Language and Linguistics: Phonetics and Phonology*, Haruo Kubozono, ed., Mouton, 363–395.
Jannedy, Stephanie (1995) Gestural phasing as an explanation for vowel devoicing in Turkish. *Ohio State University Working Papers in Linguistics* **45**: 56–84.
Jones, Daniel (1967) *The phoneme: Its nature and use*. Cambridge: Cambridge University Press.
Joo, Hakutaro (2008) *Ippan Onseigaku Koogi [Lectures on General Phonetics]*. Tokyo: Bensei.
Jun, Sun-Ah (2005) Prosodic typology. In *Prosodic Typology: The Phonology of Intonation and*

Phrasing, Sun-Ah Jun, ed., Oxford: Oxford University Press.
Jun, Sun-Ah (2006) Intonational phonology of Seoul Korean revisited. In *Japanese/Korean Linguistics 14*, Timothy Vance & K. Jones, eds., Stanford: CSLI, 15–26.
Kahn, Daniel (1976) *Syllable-based Generalizations in English Phonology*. Doctoral dissertation, MIT, published by Garland Press, New York, 1980.
Kawagoe, Itsue (2015) Phonology of sokuon, or geminate obstruents. In *The Handbook of Japanese Language and Linguistics: Phonetics and Phonology*, H. Kubozono, ed., Mouton de Gruyter, 79–119.
Kawahara, Shigeto (2006) A faithfulness ranking projected from a perceptibility scale: The case of [+voice] in Japanese. *Language* **82**(3): 536–574.
Kawahara, Shigeto (2012a) The intonation of nominal parentheticals in Japanese. In *Prosody Matters: Essays in honor of Elisabeth Selkirk*, Toni Borowsky, Shigeto Kawahara, Takahito Shinya, & Mariko Sugahara, eds., London: Equinox, 304–340.
Kawahara, Shigeto (2012b) A review of The Phonology of Japanese. *Phonology* **29**(3): 540–548.
Kawahara, Shigeto (2015a) The phonetics of *sokuon*, obstruent geminates. In *The Handbook of Japanese Language and Linguistics: Phonetics and Phonology*, Haruo Kubozono, ed., Mouton, 43–73.
Kawahara, Shigeto (2015b) The phonology of Japanese accent. In *The Handbook of Japanese Language and Linguistics: Phonetics and Phonology*, Haruo Kubozono, ed., Mouton: Mouton de Gruyter, 445–492.
Kawahara, Shigeto & Takahito Shinya (2008) The intonation of gapping and coordination in Japanese: Evidence for Intonational Phrase and Utterance. *Phonetica* **65**(1-2): 62–105.
Keating, Patricia A. & Marie Huffman (1984) Vowel variation in Japanese. *Phonetica* **41**: 191–207.
Kochetov, Alexei (2014) Japanese in the typology of nasal place assimilation: Electropalatographic evidence. Talk presented at FAJL 7.
Kolinsky, Régine & José Morais (1996) Migrations in speech recognition. *Language and Cognitive Processes* **11**: 611–619.
Kolinsky, Régine, José Morais, & M. Cluytens (1995) Intermediate representations in spoken word recognition: Evidence from word illusions. *Journal of Memory and Language* **34**: 19–40.
Kondo, Mariko (1997) *Mechanisms of vowel devoicing in Japanese*. Doctoral dissertation, University of Edinburgh.
Kubozono, Haruo (1989) The mora and syllable structure in Japanese: Evidence from speech errors. *Language and Speech* **32**: 249–278.
Kubozono, Haruo (1996) Syllable and accent in Japanese: Evidence from loanword accentuation. *The Bulletin (Phonetic Society of Japan)* **211**: 71–82.
Kubozono, Haruo (1999a) Kayoo-ni okeru moora-to onsetsu [The syllable and mora in song]. In *Bunpou-to Onsei [Speech and Grammar]*, Speech and Grammar Study Group.
Kubozono, Haruo (1999b) Mora and syllable. In *The Handbook of Japanese Linguistics*, Natsuko Tsujimura, ed., Oxford: Blackwell, 31–61.
Kubozono, Haruo (1999c) *Nihongo-no Onsei: Gendai Gengogaku Nyuumon 2 [Japanese Phonetics: An Introduction to Modern Linguisitcs 2]*. Tokyo: Iwanami.
Kubozono, Haruo (2001) Epenthetic vowels and accent in Japanese: Facts and paradoxes. In *Issues in Japanese Phonology and Morphology*, Jeroen Van de Weijer & Tetsuo Nishihara, eds., Berlin & New York: Mouton de Gruyter, 113–142.

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

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

Kubozono, Haruo (2015a) Diphthongs and vowel coalescence. In *The Handbook of Japanese Language and Linguistics: Phonetics and Phonology*, Haruo Kubozono, ed., Mouton, 215–249.

Kubozono, Haruo (2015b) Loanword phonology. In *The Handbook of Japanese Language and Linguistics: Phonetics and Phonology*, Haruo Kubozono, ed., Mouton, 313–361.

Kubozono, Haruo, Junko Ito, & Armin Mester (2008) Consonant gemination in Japanese loanword phonology. In *Current Issues in Unity and Diversity of Languages. Collection of Papers Selected from the 18th International Congress of Linguists*, The Linguistic Society of Korea, ed., Republic of Korea: Dongam Publishing Co, 953–973.

Kurisu, Kazutaka (2000) Richness of the base and root-fusion in Sino-Japanese. *Journal of East Asian Linguistics* **9**: 147–185.

Labrune, Laurence (2002) The prosodic structure of simple abbreviated loanwords in Japanese: A constraint-based account. *Onsei Kenkyu [Journal of the Phonetic Society of Japan]* **6**(1): 98–120.

Labrune, Laurence (2012a) *The phonology of Japanese*. Oxford: Oxford University Press.

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

Lahiri, Aditi & George Hankamer (1988) The timing of geminate consonants. *Journal of Phonetics* **16**: 327–338.

Lehiste, Ilse (1970) *Suprasegmentals*. Cambridge: MIT Press.

Lehtonen, Jaakko (1970) *Aspects of Quantity in Standard Finnish*. Jyvaeskylaessae: Osakeyhtioen kirjapainossa.

Letterman, Rebecca S. (1994) A phonetic study of Sinhala syllable rhymes. *Working Papers of the Cornell Phonetics Laboratory* **9**: 155–181.

Lindblom, Björn (1963) Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America* **35**: 1773–1781.

Macmillan, Neil & Douglas Creelman (2005) *Detection Theory: A User's Guide. 2nd Edition*. Mahwah: Lawrence Erlbaum Associates Publishers.

Manabe, Noriko (2009) *Western music in Japan: The evolution of styles in children's songs, hip-hop and other genres*. Doctoral dissertation, City University of New York.

Matsui, Michinao (2014) Nihongo-no boin museika-ni kansuru C/D moderu-no nyuuryoku jyoohoo-ni tsuite [On the Input Information of the C/D model for Vowel Devoicing in Japanese]. Ms. Osaka Health Science University.

McCarthy, John J. & Alan Prince (1993) Generalized alignment. In *Yearbook of Morphology*, Geert Booij & Jaap van Marle, eds., Dordrecht: Kluwer, 79–153.

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

Nakamura, Miyoko & Régine Kolinksy (2014) Multiple functional units in the preattentative segmentation of speech in Japanese: Evidence from word illusions. *Language and Speech* **57**: 513–543.

Nakano, Kazuo (1969) A phonetic basis for the syllabic nasal in Japanese. *Onsei no Kenkyuu [Studies in Phonetics]* **14**: 215–228.

Ofuka, Etsuko (2003) Sokuon /tt/-no chikaku: Akusento gata-to sokuon/hisokuongo-no onkyooteki tokuchoo niyoru chigai [Perception of a Japanese geminate stop /tt/: The effect of pitch type and acoustic characteristics of preceding/following vowels]. *Onsei Kenkyu [Journal of the Phonetic Society of Japan]* **7**(1): 70–76.

Okada, Hideo (1999) Japanese. *Tha Handbook of the Ineternational Phonetic Association* : 117–119.

Otake, Takashi, Giyoo Hatano, Anne Cutler, & Jacques Mehler (1993) Mora or syllable? Speech Segmentation in Japanese. *Journal of Memory and Language* **32**: 258–278.

Port, Robert, Salman Al-Ani, & Shosaku Maeda (1980) Temporal compensation and universal phonetics. *Phonetica* **37**: 235–252.

Port, Robert, Jonathan Dalby, & Michael O'Dell (1987) Evidence for mora timing in Japanese. *Journal of the Acoustical Society of America* **81**: 1574–1585.

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

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

Selkirk, Elisabeth (2005) Comments on Intonational Phrasing in English. In *Prosodies: With Special Reference to Iberian Languages*, Sónia Frota, Marina Vigário, & Maria João Freitas, eds., Berlin: Mouton de Gruyter, 11–58.

Selkirk, Elisabeth O. (2011) The syntax-phonology interface. In *The Handbook of Phonological Theory, 2nd Edition*, John Goldsmith, Jason Riggle, & Alan Yu, eds., Oxford: Blackwell-Wiley, 435–484.

Shibata, Takeshi (1958) Onsei–Sono Honshitsu to Kinoo [Sounds: Its Natures and Functions]. In *Onsei-no Riron-to Kyooiku [Sounds: Theories and Education]*, Asakura, 3–46.

Starr, Rebecca L & Stephanie S. Shih (2014) The syllable as a prosodic unit in Japanese lexical strata: Evidence from text-setting. Ms. National University of Singapore and University of California, Merced. (Also talk presented at LabPhon 14).

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

Sugito, Miyoko (1998) *Nihongo onsei no kenkyu 5: "hana" to "hana" [Studies on Japanese Phonetics 5]*. Osaka: Izumi Shoin.

Tamaoka, Katsuo & Shogo Makioka (2009) Japanese mental syllabary and effects of mora, syllable, bi-mora and word frequencies on Japanese speech production. *Language and Speech* **52**(1): 79–112.

Tamaoka, Katsuo & Yasushi Terao (2004) Mora or syllable? Which unit do Japanese use in naming visually presented stimuli? *Applied Psycholinguistics* **25**(1): 1–27.

Tanaka, Shin-Ichi (2013) Review of Labrune, Laurence (2012) *The Phonology of Japanese*, Oxford University Press, 320p. *Onsei Kenkyu [Journal of the Phonetic Society of Japan]* **17**(1): 70–80.

Tanaka, Shin'chi (2000) Nihonkayookyoku-ni okeru rizumu-no hensen [The changing rhythm of Japanese popular songs]. *Proceedings of the 14th meeting of the Phonetic Society of Japan* : 153–158.

Tanaka, Shin'ichi (1999) Nihongo no onsetsu to yonhaku no tenpurreeto: Senryuu to puroyakyuu seien ni okeru jiamari no bunseki [Syllables and four-mora templates in Japanese: Analysis of 'ji-amari' in senryu poems and baseball cheering]. In *Bunpou to onsei II [Speech and grammar*

2], Spoken Language Working Group, ed., Tokyo: Kuroshio, 261–290.
Tanaka, Shin'Ichi (2012) Syllable neutralization and prosodic unit in Japanese Senryu poems. Talk presented at Metrics, Music and Mind, Rome.
Tsuchida, Ayako (1997) *Phonetics and Phonology of Japanese Vowel Devoicing*. Doctoral dissertation, Cornell University.
Turk, Alice, S. Nakai, & Mariko Sugahara (2006) Acoustic segment durations in prosodic research: A practical guide. In *Methods in Empirical Prosody Research*, Stefan Sudhoff, Denisa Lenertova, Roland Meyer, Sandra Pappert, Petra Augurzky, Ina Mleinek, Nicole Richter, & Johannes Schliesser, eds., Berlin: Walter de Gruyter, 1–27.
Vance, Timothy (1987) *An Introduction to Japanese Phonology*. New York: SUNY Press.
Vance, Timothy (2008) *The Sounds of Japanese*. Cambridge: Cambridge University Press.
Vance, Timothy (2013) Book review: The Phonology of Japanese. *Lingua* **123**: 168–174.
Yoshida, Kenji, Kenneth de Jong, John Kruschke, & Pia-Maria Päiviö (2015) Cross-language similarity and difference in quantity categorization of Finnish and Japanese. *Journal of Phonetics* **50**: 81–98.