

Phonology

<http://journals.cambridge.org/PHO>

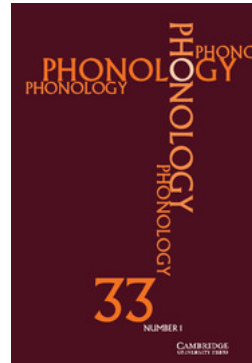
Additional services for *Phonology*:

Email alerts: [Click here](#)

Subscriptions: [Click here](#)

Commercial reprints: [Click here](#)

Terms of use : [Click here](#)



Japanese has syllables: a reply to Labrune

Shigeto Kawahara

Phonology / Volume 33 / Issue 01 / May 2016, pp 169 - 194

DOI: 10.1017/S0952675716000063, Published online: 15 June 2016

Link to this article: http://journals.cambridge.org/abstract_S0952675716000063

How to cite this article:

Shigeto Kawahara (2016). Japanese has syllables: a reply to Labrune. *Phonology*, 33, pp 169-194 doi:10.1017/S0952675716000063

Request Permissions : [Click here](#)

Squibs and replies

*Japanese has syllables: a reply to Labrune**

Shigeto Kawahara

Keio University

Labrune (2012b) proposes a syllable-less theory of Japanese, suggesting that Japanese has no syllables, with only moras below the foot. She argues that there is no phonetic or psycholinguistic evidence for the existence of syllables in Japanese. This reply summarises and re-examines previous experimental findings that demonstrate that Japanese does show evidence for syllables both phonetically and psycholinguistically. After an extensive review of previous studies, this reply also takes up a number of phonological and theoretical issues that require an explicit response from the perspective of a syllable proponent. On the basis of 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 analysed in terms of syllables can be reanalysed 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 theoretical consequence of this view, argues that not all prosodic levels are universal, extending the suggestions of Hyman (1985, 2008). Although this proposal is very thought-provoking, and its potential theoretical

* E-mail: KAHARA@ICL.KEIO.AC.JP.

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 presented in a 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 associate editor and the reviewers: Robert Daland, Donna Erickson, Osamu Fujimura, Haruka Fukazawa, Dylan Herrick, Junko Ito, Mayuki Matsui, Michinao Matsui, Armin Mester, Jeff Moore, Takashi Otake, Miho Sasaki, Helen Stickney, Yoko Sugioka and Yukiko Sugiyama. Preparation of this paper was partly supported by two JSPS Kakenhi grants, #26770147 and #26284059. Remaining errors are mine.

consequences are important, it fails to take into account a substantial body of previous experimental findings about the existence of syllables in the prosodic organisation of Japanese. This reply article re-examines the evidence demonstrating that there is both phonetic and psycholinguistic evidence for syllables in Japanese (§3 and §4), and also addresses some of the phonological and theoretical issues from the perspective of proponents of the syllable (§5).¹

2 Background: heavy syllables in Japanese

As background for the later discussion, this section provides an overview of different types of sequences which have been considered to be ‘heavy syllables’ in standard phonological analyses of Japanese. It is important to focus on heavy syllables, because light syllables and moras coincide in Japanese. It is therefore crucial to demonstrate the existence of heavy syllables in order to show that syllables are required in the prosodic organisation of Japanese phonology in addition to moras.

Since Japanese syllable structures are simple, syllables and moras generally coincide (e.g. [ta] is both monomoraic and monosyllabic). However, moras and syllables diverge in the case of heavy syllables, which contain two moras (e.g. [taa] is monosyllabic, but bimoraic). Just four types of heavy syllables are found in Japanese, as summarised in (1).²

(1) *Types of heavy syllables in Japanese*

a. <i>First half of geminate</i>	/haQ+puN/	[happun]	‘excitement’
Q/	/haQ+tatu/	[hattatsu]	‘development’
	/haQ+keN/	[hakken]	‘discovery’
	/haQ+saN/	[hassan]	‘diffusion’
b. <i>Moraic nasal</i>	/heN/	[hen]	‘strange’
Q/	/heN+paJ/	[hempai]	‘to pour back’
	/heN+da/	[henda]	‘is strange’
	/heN+ka/	[henka]	‘change’
c. <i>Second part of long vowel</i>	/okaRsaN/	[okaasan]	‘mother’
R/	/otoRsaN/	[otoosan]	‘father’
	/oneRsaN/	[oneesan]	‘sister’
	/oniRsaN/	[oniisan]	‘brother’
d. <i>Second part of diphthong</i>	/aJ/	[ai]	‘love’
J/	/oJ/	[oi]	‘nephew’

¹ Vance (2013: 171), in his 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’. This paper is a response from one such proponent. See also Ito & Mester (2015) and Tanaka (2013) for other critical responses to the syllable-less theory of Japanese.

² This paper does not consider syllables preceding a devoiced vowel, e.g. [kasu] ‘scam’, because whether or not the syllable with a devoiced vowel loses its syllabicity resulting in resyllabification is a controversial matter (see e.g. Vance 1987, 2008).

The heavy syllables in (1a) contain the first half of a geminate consonant (Kawagoe 2015, Kawahara 2015a). The traditional label for the first half of a geminate is /Q/ ('*sokuon*' in traditional Japanese grammar), a convention also employed by Labrune (2012b). The /Q/ archiphoneme assimilates to the following consonant in manner and place. For example, /haQ+tatu/ is realised as [hattatsu], and /haQ+keN/ as [hakkeN]. /haQ/ (or [hat, hak]) is thus monosyllabic, but bimoraic.

(1b) contains a 'morai nasal', traditionally represented as /N/ ('*hatsuon*' in traditional grammar). Word-finally before a pause, this consonant is arguably realised as a uvular nasal, [N], without much oral constriction – its exact phonetic realisation has been extensively debated (Vance 1987: 37–39, 2008: 95–105, Okada 1999; see also §5.2). When the nasal consonant is followed by a stop, it assimilates to it in place of articulation; e.g. /heN+da/ is realised as [henda], as (1b).³ Again, a sequence like /heN/ is monosyllabic, but bimoraic.

The heavy syllables in (1c) contains a long vowel, whose second element is phonemically labelled as /R/ (or /H/); for example, /kaR/ in /okaRsan/ is realised as [kaa]. (1d) contains a diphthong, i.e. a tautosyllabic sequence of two vowels (e.g. [ai oi]). It is debatable which vowel sequences constitute a diphthong and which a hiatus in Japanese, but there is a general consensus that an [ai] sequence is 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 characterisation of heavy syllables reviewed in this section; i.e. syllables in (1) containing /Q N R J/ constitute one prosodic unit – one heavy syllable. In fact, not only have they assumed the existence of heavy syllables, they have also provided a body of explicit phonological evidence for the existence of syllable structures in the phonological organisation of Japanese (McCawley 1968, Itô 1989, 1990, Kubozono 1989, 1999a, b, 2003, 2015a, Haraguchi 1996, Itô & Mester 2003, 2015, Kubozono *et al.* 2008, among many others).⁴

Arguing against this commonly held view, Labrune (2012b) instead proposes to treat these heavy syllables as sequences of a vowel (a 'regular' or 'full' mora) and a following /Q N R J/ element, the latter being collectively

³ The behaviour of moraic nasals before fricatives is a matter of debate. Vance (2008: 97) claims that it is realised as a nasalised dorsovelar approximant, [ũ̃], while acknowledging that it is hard to transcribe. 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 nature of the assimilation process here is irrelevant to the main point of this paper, however.

⁴ As can be seen in this list of the references, Kubozono should perhaps be given the most credit for establishing the existence of syllables in the prosodic organisation of Japanese phonology. Contributions by Itô & 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 §5 of this paper; interested readers are referred to these original works for details.

referred to as a ‘deficient’ or ‘special’ mora. This proposal does away entirely with the notion of syllable, an idea which also appears in Labrune (2012a). The current paper refers to this proposal as ‘the syllable-less theory’ of Japanese.⁵ In essence, the syllable-less theory of Japanese decomposes heavy syllables into sequences of two moras. One important basis of this proposal is the alleged lack of phonetic interaction between the nucleus and the following /Q N R J/ element. This paper thus starts by reviewing and re-examining the evidence that the vowel and the following tautosyllabic element do in fact 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 ‘absence of phonetic clues for the existence of a rhyme-like constituent’ (Labrune 2012b: 120), where ‘rhyme-like constituent’ refers to the combination of a vowel and any of /Q N R J/. In other words, Labrune’s claim is that there is no phonetic interaction between the nucleus and a following /Q N R J/. However, there is evidence which suggests that this is not the case.

3.1 The interaction between /N/ and a preceding vowel

Perhaps the most convincing case for the phonetic interaction between the coda consonant and the preceding vowel comes from the behaviour of the coda moraic nasal, /N/. The first type of evidence comes from durational compensation effects 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 analyses a corpus consisting of speech of 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 durations of the vowel and the following moraic nasal correlate negatively with each other – the longer the vowel, the shorter the consonant (Campbell 1999: 34–35). More specifically, among back vowels in particular, the higher the vowel, the shorter it is ([u] = 100 ms, [o] = 109 ms, [a] = 115 ms), while the moraic nasal /N/ undergoes lengthening following higher vowels (post-[u] = 90 ms, post-[o] = 87 ms, post-[a] = 83 ms). These

⁵ Labrune explicitly declares that her proposal is a revival of the view in the traditional study of Japanese, also known as ‘*kokugogaku*’. She states (2012b: 114): ‘interestingly, the Japanese linguistic tradition, which has a long history 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’. This is at best an oversimplified statement. Joo (2008: 217–218) lists several traditional Japanese linguists who have proposed the notion of the syllable in addition to the mora in Japanese, including Hattori (1951) and Shibata (1958). Joo (2008) himself adopts an analysis in which Japanese both has syllables and moras.

lengthening patterns of /N/ after higher vowels are statistically significant, as shown by 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. Lehiste 1970, Port *et al.* 1980, Campbell & Isard 1991). Indeed, Campbell (1999: 35) 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'. The 'higher-level framework' referred to here is the syllable or the rhyme (the title of Campbell's paper is 'A study of Japanese speech timing from the syllable perspective'). Campbell's data show that the durational target for the Japanese rhyme falls within the range of 190 ms to 200 ms.

A vowel and a following /N/ interact not only in duration, but also in nasalisation. Vowels are nasalised before a nasal consonant within the same syllable (i.e. /N/) in Japanese. The patterns of vowel nasalisation in Japanese are illustrated in (2) (Vance 1987, 2008, 2013, Campbell 1999, Labrune 2012b). Vowels are nasalised before a tautosyllabic nasal consonant, but not before an onset nasal consonant. Vance (2008: 96), citing some classic descriptions of Japanese phonetics (Bloch 1950, Jones 1967, Nakano 1969), states that 'the vowel before [N] is clearly nasalised [in Japanese]', and a few pages later shows oral vowels following onset nasal consonants (e.g. the diphthong in [sãmmai] 'three sheets') (2008: 99).

- | | | | | | |
|--------|----------|----------------|----|---------|--------|
| (2) a. | [hõN] | 'book' | b. | [ho.ne] | 'bone' |
| | [hõn.da] | 'Honda' | | [ko.me] | 'rice' |
| | [hõm.ma] | 'Homma (name)' | | | |

The domain of vowel nasalisation in Japanese can only be defined in terms of the syllable: vowels are nasalised before a tautosyllabic nasal. The domain of nasalisation cannot be explained in terms of moras, because in both (2a) and (b), the first vowel consists of a mora preceding a nasal consonant, but only the former undergoes nasalisation. We cannot explain the nasalisation pattern in terms of feet either; 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), which is also adopted by Labrune. Finally, vowel nasalisation cannot be explained in terms of a precedence relationship, because not all vowels preceding nasal consonants are nasalised, as in (1b): regressive nasalisation is blocked by a syllable boundary.

There is also anecdotal yet telling evidence for the perceptual relevance of a syllable unit containing a coda nasal. Musashi Homma at the Tokyo Metropolitan Neurological hospital has been working on a project to record ALS patients' voices before they undergo a tracheotomy. He and his colleagues developed software in which this recording allows the patients to communicate with their caretakers using their own voice,

played through a personal computer. Several phoneticians, including myself, have now been participating in this project from a linguistic perspective (Kawahara *et al.* 2016). In this project, the recording was done at the level of the mora, meaning that /ho/ and /N/, for example, were recorded separately; however, Homma noticed that the patients were much happier with the quality of the recording if the coda nasal was recorded with the preceding mora (e.g. /hoN/). This observation shows that the coda nasal constitutes a unit with a preceding vowel (i.e. a syllable), with vowel nasalisation likely to be one of the perceptual cues for this unit.

An anonymous reviewer goes a step further, pointing out that vowels are not nasalised after /N/ (e.g. the second vowel in [ãN.i] ‘easy’ remains oral). When a morpheme boundary follows the moraic nasal, it is realised as a coda consonant /N/, without resyllabification (see §5.2). Even if the next morpheme begins with a vowel, that vowel is not nasalised. This example precludes an explanation of Japanese vowel nasalisation in which coda nasals are phonologically /N/ throughout the phonological derivation (e.g. [honda] is represented as [hoNda]), and only /N/, not /n/ or /m/, causes nasalisation. If /N/ were the trigger of nasalisation without domain restrictions, both the preceding and following vowel should be nasalised; however, nasalisation by /N/ is blocked in the following vowel, which is separated by a syllable boundary. This pattern ([ãN.V]) is therefore consistent with the view that the domain of vowel nasalisation in Japanese is the syllable.

Vance (2013) points out that Labrune (2012a, b) explicitly acknowledges this coda-induced nasalisation process: her phonetic representations show that nasalisation applies to the vowel if and only if the vowel and the nasal are in the same syllable. For instance, Labrune (2012b: 122) has the transcriptions [ãN.i] ‘ease’, [a.ni] ‘old brother’ and [ãN.ni] ‘implicitly’, in which all and only the vowels that are followed by a tautosyllabic nasal are nasalised. As Vance (2013: 172) notes, it is somewhat puzzling that this coda-induced nasalisation is not considered to constitute evidence for phonetic interaction between the nucleus and the following tautosyllabic element.

3.2 The interaction between /Q/ and a preceding vowel

We now look at how vowels interact with a following /Q/ – the first part of a geminate. The first half of a geminate affects the quality and duration of the preceding vowel. For example, experimental studies have found that singleton consonants and geminate consonants affect the F0 of preceding vowels in different ways.

In order to investigate the phonetic properties of voicing and geminates in Japanese, Kawahara (2006) recorded the speech of three female native speakers of Japanese pronouncing singleton stops and geminate stops in a /kVC(C)V/ word frame. Acoustic analysis shows that F0 is higher before geminates than before singletons, indicating that /Q/ raises the F0

of a preceding tautosyllabic vowel. Idemaru & Guion (2008) show that the fall in F₀ due to pitch accent is greater across a geminate than across a singleton consonant; i.e. the realisation of pitch accent on the preceding vowel is affected by /Q/. Their study was based on the speech of six native speakers, three females and three males, and examined how a singleton–geminate contrast is manifested 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 (2008: 180; see also Ofuka 2003 for a similar observation). Fukui (1978) reports that in unaccented words F₀ in vowels falls preceding geminates. 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: 120–121), vowels are longer when followed by a geminate consonant than when followed by a singleton consonant; this pattern has been documented in many acoustic studies on Japanese geminates (e.g. Fukui 1978, Port *et al.* 1987, Han 1994, Campbell 1999, Kawahara 2006, Hirata 2007, Idemaru & Guion 2008). According to Idemaru & Guion (2008: 176), for example, preceding vowels were on average 75 ms before geminates, whereas they were 59 ms before singletons (i.e. 16 ms difference). The median values reported by Campbell (1999: 35) are 85 ms for the vowels in CV open syllables, and 105 ms for the vowels before /Q/ (i.e. 20 ms difference). /Q/ therefore interacts with the preceding vowel in the same syllable by lengthening it.

Labrune (2012b: 120–121) claims that this interaction does not instantiate a true phonetic interaction between the nucleus and the following tautosyllabic element, because we would expect shortening rather than lengthening before geminates. However, pregeminate lengthening is nevertheless a phonetic interaction, and is not even a cross-linguistic anomaly, as other languages show similar lengthening, e.g. Finnish (Lehtonen 1970: 110–111), Persian (Hansen 2004), Shinhala (Letterman 1994) and Turkish (Jannedy 1995). The fact that vowels lengthen rather than shorten before a geminate does not mean that there is no interaction between the vowel and /Q/ (see Kawahara 2015a for a review).

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

Moreover, there is further evidence that ties /Q/ and /N/ together: vowels are longer before a moraic nasal /N/ than in open syllables (Campbell 1999: 34–35): the median vowel durations reported in Campbell (1999: 35) are 90 ms for open syllables and 110 ms when there is a coda nasal (i.e. 20 ms difference). Hence /N/ lengthens a preceding

tautosyllabic vowel, more or less to the same degree that /Q/ lengthens a preceding vowel (16 ms or 20 ms, as discussed above). The behaviour of /N/ discussed in §3.1 and the fact that /N/ and /Q/ affect the duration of the preceding vowel in similar ways support the claim 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).

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 a 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/). The study found that short and long vowels differ not only in duration, but also in formant frequency. The results of their study are reproduced in Fig. 1. It can be observed 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 a lower F2 than its short counterpart. [aa] has a lower F1 than short [a]. [uu] is the only vowel whose F1 and F2 characteristics are almost identical to its short counterpart.

If long vowels were represented simply as a sequence of a vowel + /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 seen as merely a mechanical issue. To account for the dispersion effect in long vowels, one might propose an alternative theory, whereby speakers cannot reach the formant targets in short vowels because they are too short for speakers to achieve their acoustic targets (i.e. they 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 then 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 to be the 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 mere sequences of a short vowel and /R/.

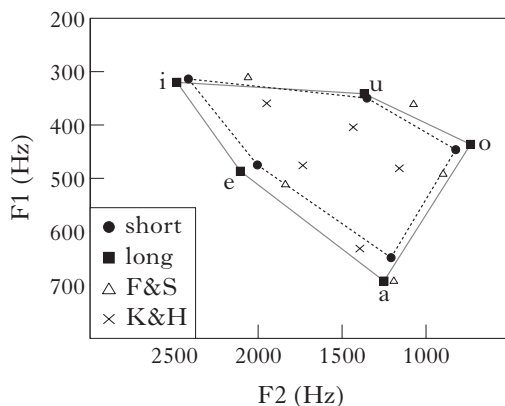


Figure 1

Results from Hirata & Tsukada (2009: 137), adapted from their Fig. 1, showing the F1 and F2 values of the five Japanese vowels, short and long. Dashed lines represent the vowel space of short vowels, and solid lines the vowel space of long vowels. Data from two previous studies on vowel-formant frequencies of Japanese short vowels (Fujisaki & Sugito 1977 (F&S) and Keating & Huffman 1984 (K&H)) are plotted in the figure, to confirm the reliability of Hirata & Tsukada's F1 measurements, despite the presence of surrounding nasal consonants in their stimuli.

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 a /VJ/ sequence is syllabified monosyllabically within a morpheme, and heterosyllabic parsing occurs across a morpheme boundary (see Kubozono 2015a). The effect of a syllable boundary is therefore 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 it is next to impossible to reliably measure the duration of a V adjacent to [i], for example (Turk *et al.* 2006).

Labrune (2012a: 54) 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', suggesting that two vowel sequences in Japanese are phonetically 'more separable' than those in other languages. Labrune thus implies that the alleged lack of gradual changes between two adjacent vowels is evidence for lack of syllables in Japanese. However, she offers no quantitative support for this claim. Kawahara (2012b) argues on the basis of spectrogram and waveform inspection that this claim is not well supported.

4 Psycholinguistic evidence for syllables in Japanese

Labrune (2012b: 120) states that ‘none of the many psycholinguistic studies which have been conducted has been able to establish the cognitive reality of the syllable in Japanese’. While it is probably true that the dominant segmentation pattern for Japanese speakers is mora-based (e.g. Otake *et al.* 1993),⁶ this statement is too strong, and this section re-examines evidence that falsifies this claim. We will review psycholinguistic evidence for syllables from (i) the behaviour of adult speakers, (ii) the behaviour of Japanese children and (iii) 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) presented two types of target stimuli orthographically: (i) trisyllabic trimoraic nonce words (CVCVCV), and (ii) disyllabic trimoraic nonce words (CVXCV, where X = /Q N R J/). In addition, to provide a baseline measure for simple disyllabic words, the experiment included disyllabic bimoraic CVCV nonce words, as shown in Table I.

stimulus	example	visual stimulus	naming latency (ms)	error rates (%)
CV.CV.CV	/ketape/	けたぺ	645	15.1
CVJ.VC	/keope/	けおぺ	590	7.8
CVQ.CV	/keQpe/	けっぺ	575	9.9
CVN.CV	/keNpe/	けんぺ	533	3.1
CV.CV	/kepe/	けぺ	537	1.6

Table I

Results from Tamaoka & Terao’s (2004) Experiment I.

The task was to pronounce the visually presented words as quickly as possible. For the visual prompts, their Experiment 1 used the *hiragana* orthography and the Experiment II the *katakana* orthography. In both writing systems, each letter corresponds to one mora. The experiment

⁶ Even this statement should be taken with caution. The work by 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: 296) state that ‘our 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’.

stimulus	example	visual stimulus	naming latency (ms)	error rates (%)
CV.CV.CV	/kotamo/	コタモ	645	2.78
CVR.CV	/koRmo/	コーモ	573	1.91

Table II

Results from Tamaoka & Terao's (2004) Experiment II.

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-likeness measures, independently judged by a different set of speakers, because these are the two factors which were known to have a rather substantial effect on naming latencies (see Tamaoka & Terao 2004, Tamaoka & Makioka 2009 and references cited therein for discussion). Twenty-four native speakers of Japanese participated in these experiments.

The results of Experiment I are summarised here in Table I (Tamaoka & Terao 2004: 11). The trisyllabic trimoraic words in the first row showed the longest naming latencies compared to the disyllabic trimoraic words (rows 2–4). In fact, the CVNVC condition showed shorter naming latencies than CVCV stimuli (row 5) – the naming latencies between these two conditions are 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 than in the disyllabic condition.⁷

The behaviour of CVR syllables was tested in Experiment II, whose results are partially reproduced here as Table II (Tamaoka & Terao 2004: 17). The CVRCV stimuli showed shorter naming latencies and lower error rates than the CVCVCV stimuli, despite the fact that the stimuli had three moras in both conditions.

In summary, Tamaoka & Terao's (2004) experiments show that disyllabic CVXCV stimuli showed shorter naming latencies and lower error rates than trisyllabic CVCVCV stimuli, despite the fact that both conditions involve three moras and three letters. Based on these results, Tamaoka & Terao (2004: 20) conclude that '/N/, /R/, /Q/, and /J/ are

⁷ The associate editor asks if the difference between the trisyllabic 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 needs to be explored further in future experimental work. However, the results showing that the CVNVC and CVCV stimuli pattern together are hard to reconcile with the segment-counting view, because the CVNVC stimuli contain an extra segment compared to the CVCV stimuli.

combined with a preceding CV mora when native Japanese speakers pronounce visually presented nonwords, regardless of [whether they are] presented in the hiragana or katakana script. More specifically, syllabic units are basically used for naming tasks requiring phonological production’.

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 CVNVCV words were pronounced with shorter latencies and lower error rates than CVCVCV nonce words (654 ms *vs.* 735 ms; 1.49% *vs.* 7.44%) (2009: 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% *vs.* 11.46%) (2009: 94).

4.2 Evidence from adult listeners: perception

Nakamura & Kolinsky (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 *et al.* 1995). Nakamura & Kolinsky tested whether this migration can occur at several different prosodic levels for Japanese listeners, as illustrated in Table III. For example, when presented with the stimulus [genru] to one ear and [haido] 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.

migration unit	stimulus to one ear	stimulus to the other ear	detectability (d')
consonant	g airu	h endo	2.04
vowel	he ir u	g ando	2.90
mora	ge iru	han do	2.76
syllable	geN ru	hai do	2.99

Table III

Examples of migration patterns tested in Nakamura & Kolinsky’s (2014) dichotic listening test. Migration of the units shown in bold would result in the perception of [gendo] in all four cases. The experiment used five sets of pairs of this kind.

Detectability of the target word resulting from migration was measured in terms of a detectability index, d' (Macmillan & Creelman 2005).

In their experiment, the task was for participants to decide whether or not they heard a particular target word – e.g. [gendo]. In the test condition,

the target words occurred after migration of sounds, as illustrated in Table III (derived from Nakamura & Kolinsky's Table 2). As a comparison, they included stimuli which should not result in the perception of the target words, even after migration (e.g. [genru] and [haiba]). A total of 56 native speakers of Japanese participated in this listening experiment.

The 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 III. For example, in the syllable condition, the listeners correctly identified the target words as present in 90.6% of cases when the targets were present in the stimuli after migration (hit rate); at the same time, they also correctly judged in 90.7% of cases that there were no target words when the stimuli did not include the target words (correct rejection rate). These high hit rates and high correct rejection rates result in a high d' value.

Moreover, migration at the mora and syllable levels occurred more frequently than at the consonant level, showing that migration at the higher level cannot be accounted for in terms of combinations of segmental migrations. Most importantly in this context, the illusory 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 behaviour

Psycholinguistic evidence for syllables in Japanese comes from children's behaviour as well as from adult behaviour. Inagaki *et al.* (2000) report an experiment which addressed the issue of whether Japanese speakers make use of syllables in their speech segmentation. Since the *hiragana* and *katakana* orthographies are mora-based, in that one letter usually corresponds to one mora,⁸ it is reasonable to suspect that adult speakers' speech-segmentation patterns are 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, 'children 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' (Inagaki *et al.* 2000: 75). For example, a word like [kurejoN] 'crayon' can be counted with three gestures if it is segmented in terms of syllables ([ku.re.joN]), but with four if it is segmented in terms of moras ([ku-re-jo-N]). The children first underwent a practice session 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.joN]), 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

⁸ The only exception is three letters expressing palatalisation of preceding consonants (や ゆ よ), which do not count as single moras.

syllable type	segmentation pattern	level 1 (<i>n</i> = 22)	level 2 (<i>n</i> = 10)	level 3 (<i>n</i> = 10)	level 4 (<i>n</i> = 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

Table IV

Results from Inagaki *et al.*'s (2000) Experiment II, 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 Inagaki *et al.* 2000 for details).

stimuli, which were presented in the form of drawings. The participants were kindergarten children, with different proficiency levels in terms of the Japanese orthographic systems. For Experiment I, the age range was 58–80 months; for Experiment II, the age range was 54–78 months.

The results show that Japanese children exhibit a mixture of syllable-based counting and mora-based counting. The results of their Experiment II are summarised in [Table IV](#) (Inagaki *et al.* 2000: 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 number of syllable-based parsings goes down, except for the /CVQ/ syllables, which show persistent syllable-based counting patterns.

Units such as [jɔ̃n], [rap] and [koo] were sometimes associated with one counting gesture. Thus Japanese children may employ syllable-based counting gestures, at least those who are not very familiar with the Japanese orthography system.⁹ The results also show that, as children learn Japanese orthography, the mora-based parsing becomes more dominant. This experiment shows that the dominance of mora-based parsing is partly due to the Japanese orthography.


⁹ The associate editor asks 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 have been 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 – in any case, 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 the mora was the only psycholinguistic unit in Japanese.

Proponents of the syllable-less theory might suggest that how Japanese preliterate children show segmental parsing effects is not relevant to the organisation of adult Japanese prosody; after all, adult speakers, who have learned the Japanese orthographic system, do show dominant mora-based segmental parsing patterns (see Otake *et al.* 1993, among others, although Cutler & Otake 2002 cast doubt on this claim – see note 6). 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 might argue that the syllable-based parsing pattern come from an innate universal mechanism or 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: 116) mentions that ‘the mora is the metric unit of Japanese verse in poetry and singing’. This is indeed a commonly held view, and the 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: /foo/ in (3) is split between two notes. It thus seems that text-setting in Japanese can be based on syllables, as well as moras.

(3) 

 mo - mo - ta - roo - saN (verse 1, line 1)

 ja - ri - ma - fo - o (verse 2, line 1)

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 (Sugito 1998, Kubozono 1999a, Tanaka 2000, 2012, Manabe 2009; all discussed by Starr & Shih).

Starr & Shih 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 preceding 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 J/ can exhibit syllabic text-setting, both in translated songs and native songs. Syllables with /N R/ were particularly susceptible to syllabic text-setting, and in the translated Disney songs 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 observed 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 players’ names are mapped on to three musical notes, represented by XXX. When the players’ names contain four moras, the last musical note is always associated with the last syllables, whether they are heavy or light. The examples in (4) are adapted from Kubozono (2015a: 223).

(4) a. 
i tji - roo da - a - win

b. 
na - gaji - ma zu - ree - ta sa - nta - na

The generalisation 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)(jima)/).

4.5 Summary

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 behaviour, 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 certainly 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 of

Labrune's points regarding Japanese phonology. 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 maximisation in Japanese, (iii) evidence for syllables from the phonological minimality requirement and (iv) the Occam's razor argument used by Labrune.

As Ito & Mester (2015) point out, the reanalyses of syllable-based phenomena by Labrune (2012b) simply recast the distinction between head moras and non-head moras in syllable theory as 'full' *vs.* 'deficient' moras: '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' (2015: 371). In other words, Labrune's reanalyses show that the patterns previously analysed with syllables can be analysed without syllables, but not necessarily that they must be, and she herself is careful to admit this point.¹⁰ For this reason, the experimental evidence reviewed in the previous two sections should suffice to support the evidence for syllables in Japanese.¹¹

Nevertheless, I believe that there are a number of phonological and theoretical issues which require an explicit response from the proponents of the syllables in Japanese; I take these up in this section (see also Tanaka 2013 and Ito & Mester 2015).

5.1 Phonological non-isomorphism between moras and syllables

The first issue concerns the phonological non-isomorphism between moras and syllables. Although it is true that the role of the mora is very important in Japanese phonology, we should not ignore the observation that one heavy syllable patterns sometimes with one light syllable, rather than with a sequence of two light syllables, contrary to what the syllable-less theory predicts. Another important observation is that a sequence of two

¹⁰ For example, when Labrune (2012b: 123) discusses tonal phrase-initial lowering patterns, 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, 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' (2012: 134).

¹¹ Besides the alleged lack of phonetic and psycholinguistic evidence for syllables, which has already been addressed in §3 and §4, there are two more points that Labrune (2012b) raises against the existence of syllables in Japanese. First, citing Poser (1990), she argues that it is problematic that syllable boundaries and foot boundaries may not coincide (2012b: 122). In her note 9, however, she acknowledges that this problem disappears if an output–output correspondence theory of compounding is adopted. Second, she argues that speech errors are predominantly mora-based rather than syllable-based (2012b: 120). However, it is not case that syllable-based speech error patterns are impossible. For instance, Kubozono (1989: 252) reports an example like [se.kai.rem.poo.jim.buɴ] → [se.kai.rem.buɴ.jim.buɴ] 'World Federation Newspaper', in which a heavy syllable [poo] is replaced by another heavy syllable [buɴ].

moras can pattern differently depending on whether they are parsed into one syllable or not.¹²

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

(5) a. *One heavy syllable (two moras): unaccented*

[kin-taroo]	[koo-taroo]	[kjuu-taroo]
[kan-taroo]	[hoo-taroo]	[soo-taroo]

b. *One light syllable (one mora): unaccented*

[a-taroo]	[ko-taroo]
[ki-taroo]	[ja-taroo]

c. *Two syllables (two moras):*

accent on the final syllable of the initial element

[momo'-taroo]	[itʃi'-taroo]
[mitʃi'-taroo]	[asa'-taroo]

Labrune does mention this pattern, and argues that this pattern is ‘lexical rather than strictly phonological’ (2012b: 131). She then goes 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’. 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, 2001).

Another argument against the purely moraic view involves patterns of loanword accentuation. As Kubozono (1999b, 2015b) points out, given four-mora loanwords, the preferred accentual patterns differ, depending on their syllabic composition. 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 Kawahara 2015b for further discussion).

¹² I thank an anonymous reviewer for bringing this to my attention.

(6) a. LLH: *accented on the initial syllable*

[a ¹ .ma.zoN]	‘Amazon’	[he ¹ .ru.paa]	‘helper’
[do ¹ .ra.goN]	‘dragon’	[ja ¹ .ru.pii]	‘Charpy (name)’
[re ¹ .ba.noN]	‘Lebanon’	[se ¹ .ra.pii]	‘therapy’
[te ¹ .he.raN]	‘Teheran’		

b. LLLL: *unaccented*

[a.me.ri.ka]	‘America’	[i.ta.ri.a]	‘Italy’
[mo.na.ri.za]	‘Mona Lisa’	[bu.ra.ɕi.ru]	‘Brazil’
[a.ri.zo.na]	‘Arizona’	[me.ki.ʃi.ko]	‘Mexico’
[ma.ka.ro.ni]	‘macaroni’		

The syllable-less theory would not be able to distinguish the examples in (6a) and (b), because both types of words contain four moras and two feet.

The bottom-line message from Kubozono’s work 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 syllable is crucial for the characterisation of both of these observations.

5.2 Onset maximisation in Japanese

The second issue is the proposed lack of onset maximisation. Labrune (2012b: 121–122) argues that Japanese lacks ‘onset optimisation’, and that this constitutes evidence for the lack of syllables in Japanese. What is meant by the ‘lack of onset optimisation’ is a lack of resyllabification across a morpheme boundary, or what is more generally known as ‘onset maximisation’ (Prince & Smolensky 1993 and references cited therein). Contrary to the asserted lack of onset maximisation, Japanese does usually maximise onsets (Itô 1989), so that underlying monomorphemic sequences like /ana/ ‘whole’ are syllabified as [a.na] rather than [an.a].¹³ 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 (7). In this sense, Japanese does maximise onsets and avoid codas.

¹³ It might be objected 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 correct, as [an.a] occurs when a morpheme boundary follows the nasal consonant. In other words, both [a.na] and [an.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, with onset maximisation. More generally, when Japanese speakers divide words into smaller phonological chunks, they do maximise onsets: /arigatoR/ ‘thank you’ is separated into chunks as [a-ri-ga-to-o], rather than [ar-ig-at-o-o]. See also Itô (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’.

(7) *Resyllabification and onset maximisation in verbal inflectional paradigms*

<i>negative</i>	/nak+anai/	[na.ka.nai]	‘not cry’
<i>polite</i>	/nak+imasu/	[na.ki.ma.su]	‘cry (polite)’
<i>non-past</i>	/nak+u/	[na.ku]	‘cry’
<i>conditional</i>	/nak+eba/	[na.ke.ba]	‘if cry’
<i>volitional</i>	/nak+oo/	[na.koo]	‘let’s cry’

Labrune discusses the lack of syllabification across a stem boundary in Sino-Japanese compounds; e.g. /aN/ ‘safe’ + /i/ ‘easy’ is syllabified as [ã̃n.i], rather than [a.ni] (Itô & Mester 1996, 2015). 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 & 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 (Itô 1990, Itô & Mester 2003, 2015). In loanword truncation, words can be truncated to disyllabic bimoraic forms, as in (8a). 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 (8b).

(8) a. *Bimoraic truncation patterns*

[de.mon.su.to.ree.ʃoN]	→ [de.mo]	‘demonstration’
[ri.haa.sa.ru]	→ [ri.ha]	‘rehearsal’
[ro.kee.ʃoN]	→ [ro.ke]	‘location’
[bi.ru.diŋ.gu]	→ [bi.ru]	‘building’
[bu.ra.ʒaa]	→ [bu.ra]	‘brassiere’
[pu.ro.ʃeʃ.ʃo.na.ru]	→ [pu.ro]	‘professional’

b. *Monosyllabic outputs not allowed*

[mai.ku.ro.ʃo.oN]	→ [mai.ku]	*[mai]	‘microphone’
[dai.ja.mon.do]	→ [dai.ja]	*[dai]	‘diamond’
[paa.ma.nen.to]	→ [paa.ma]	*[paa]	‘permanent (hair-style)’
[kom.bi.nee.ʃoN]	→ [kom.bi]	*[koN]	‘combination’
[am.pu.ri.ʃai.aa]	→ [am.pu]	*[aŋ]	‘amplifier’
[ʃim.po.ʒi.u.mu]	→ [ʃim.po]	*[ʃiN]	‘symposium’

This truncation pattern shows that Japanese prosodic minimality is defined on the basis of the number of syllables, in addition to the well-known bimoraic requirement – one syllable, even if it is bimoraic, is too small (Itô 1990, Itô & Mester 2003, 2015).

Labrune (2012b: 126–128) discusses this pattern, and argues that the apparent prohibition on monosyllabic forms should instead be attributed to a general ban on heavy syllables (or /N R 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, the truncated forms can end with a heavy syllable, as in (9).

(9) *Truncated words ending with a heavy syllable*

[ku.so + gee.mu]	→ [ku.so.gee]	‘crappy game’
[mo.bai.ru + gee.mu]	→ [mo.ba.gee]	‘mobile game (name)’
[a.ra.un.do + saa.tii]	→ [a.ra.saa]	‘around thirty’
[a.ra.un.do + φoo.tii]	→ [a.ra.φoo]	‘around forty’
[ɕʒi.mii + hen.do.rik.ku.su]	→ [ɕʒi.mi.hen]	‘Jimmy Hendrix’
[go.ku.doo + sen.see]	→ [go.ku.sen]	‘Yakuza teacher (TV drama)’

No examples of bimorphemic truncation that I know of take an extra light syllable when they end with a heavy syllable, unlike the monomorphemic truncation in (8b) (e.g. forms like [arasaati] are not found).¹⁴ To eliminate this problem, Labrune (2012b) argues that different principles operate on monomorphemic and bimorphemic truncation patterns.

However, the syllable-based minimality explanation is superior, in that it automatically explains why bimorphemic truncated compounds can end in heavy syllables, without postulating that monomorphemic truncation and bimorphemic truncation are governed by different principles: bimorphemic 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 monomorphemic truncation pattern, a pattern which Labrune (2002, 2012b) argues is absent. However, it is questionable whether there are many base forms in the first place that are long enough for truncation to result in LLH forms (i.e. monomorphemic words of the form LLHX). Moreover, there do exist a number of examples that instantiate this pattern, as shown in (10).

(10) *Monomorphemic LLH truncation*

[pu.re.zen.tee.ʃoN]	→ [pu.re.zen]	‘presentation’
[hi.po.kon.de.rii]	→ [hi.po.kon]	‘hypochondria’
[sa.pi.en.ti.a]	→ [sa.pi.en]	‘Sapientia’
[o.ri.en.tee.ʃoN]	→ [o.ri.en]	‘orientation’
[re.pu.re.zen.tiŋ.gu]	→ [re.pe.zen]	‘representing (rap jargon)’
[re.ko.men.dee.ʃoN]	→ [re.ko.men]	‘recommendation’
[se.pu.tem.baa]	→ [se.pu.ten]	‘September student (jargon)’

¹⁴ An anonymous reviewer points out two potential counterexamples from baseball terms: [pa-riigu] from [paʃiʃikku riigu] ‘Pacific League’ and [se-riigu] from [sentoraru riigu] ‘Central League’. However, these examples do not involve truncation of the second elements. They are also exceptional in that the first elements leave only one mora rather than two; the first elements look like prefixes. These words should thus probably be viewed as involving prefixation.

These examples show that monomorphemic LLH truncation pattern is not in fact impossible. These examples also show that the ban against monosyllabic truncation shown in (8b) results from a syllable-based minimality requirement (Itô 1990), rather than a general ban on heavy syllables 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 fully acknowledges this difficulty: 'of course, lack of positive evidence does not automatically provide negative evidence', yet she goes on to say 'but it should at least lead us to question the initial postulate that Japanese is a syllable language' (2012b: 119). Behind this logic is Occam's razor – all else being equal, we should not posit a theoretical device, here syllables, unless there is a need to do so.¹⁵ 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 all else is not equal: abandoning the syllable in Japanese has consequences for other aspects of phonological theory.

From the perspective of language acquisition, it is 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 Selkirk 2005, Kawahara & Shinya 2008, Ito & Mester 2012 and Kawahara 2012a for related discussion; cf. also Hyman 1985, 2008, discussed by Labrune 2012b). 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: 97) succinctly put it, 'a universal hierarchy cannot easily admit language-specific gaps', and there have been attempts to establish a universal theory of prosodic hierarchy in the face of apparent language-specific patterns (Selkirk 2005, Ito & Mester 2007, 2012, Kawahara & Shinya 2008). 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 an Occam's razor argument in this context.

6 General conclusion

Although the syllable-less theory of Japanese phonetics and phonology is very thought-provoking, there is both phonetic and psycholinguistic evidence for syllables, and phonological consideration also suggests that Japanese has syllables.

¹⁵ On page 119, where the relevant quote is discussed, Labrune (2012b) does not mention Occam's razor. An explicit reference appears on page 139, however. As discussed below, Occam's razor arguments should be used with caution in linguistic theorising in general, because it is hardly ever the case that all else is the same – eliminating one piece of theoretical apparatus often has consequences elsewhere.

REFERENCES

- Bloch, Bernard (1950). Studies in colloquial Japanese IV: phonemics. *Lg* **26**. 86–125.
- Borowsky, Toni, Shigeto Kawahara, Takahito Shinya & Mariko Sugahara (eds.) (2012). *Prosody matters: essays in honor of Elisabeth Selkirk*. London: Equinox.
- 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 & S. D. Isard (1991). Segment durations in a syllable frame. *JPh* **19**. 37–47.
- Cutler, Anne & Takashi Otake (2002). Rhythmic categories in spoken-word recognition. *Journal of Memory and Language* **46**. 296–322.
- Fukui, Seiji (1978). Nihongo heisaon-no encho/tanshuku-niyoru sokuon/hisokuon toshite-no chooshu. [Perception of 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**. 439–453.
- Han, Mieko S. (1994). Acoustic manifestations of mora timing in Japanese. *JASA* **96**. 73–82.
- Hansen, Benjamin B. (2004). Production of Persian geminate stops: effects of varying speaking rate. In Augustine Agwuele, Willis Warren & Sang-Hoon Park (eds.) *Proceedings of the 2003 Texas linguistics society conference*. Somerville, Mass.: Cascadilla. 86–95.
- Haraguchi, Shosuke (1996). Syllable, mora and accent. In Otake & Cutler (1996). 45–75.
- 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**. 129–149.
- Hyman, Larry M. (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 the International Phonetic Association* **38**. 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.
- Itô, Junko (1989). A prosodic theory of epenthesis. *NLLT* **7**. 217–259.
- Itô, Junko (1990). Prosodic minimality in Japanese. *CLS* **26:2**. 213–239.
- Itô, Junko & Armin Mester (1996). Stem and word in Sino-Japanese. In Otake & Cutler (1996). 13–44.
- Ito, Junko & Armin Mester (2003). Weak layering and word binarity. In Takeru Honma, Masao Okazaki, Toshiyuki Tabata & Shin'ichi Tanaka (eds.) *A new century of phonology and phonological theory: a Festschrift for Professor Shosuke Haraguchi on the occasion of his sixtieth birthday*. Tokyo: Kaitakusha. 26–65.
- Ito, Junko & Armin Mester (2007). Prosodic adjunction in Japanese compounds. *MIT Working Papers in Linguistics* **55**. 97–111.
- Ito, Junko & Armin Mester (2012). Recursive prosodic phrasing in Japanese. In Borowsky *et al.* (2012). 280–303.
- Ito, Junko & Armin Mester (2015). Word formation and phonological processes. In Kubozono (2015c). 363–395.

- Jannedy, Stefanie (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*. 3rd edn. Cambridge: Heffer.
- Joo, Hakutaro (2008). *Ippan onseigaku koogi*. [Lectures on general phonetics.] Tokyo: Bensei.
- Kawagoe, Itsue (2015). The phonology of sokuon, or geminate obstruents. In Kubozono (2015c). 79–119.
- Kawahara, Shigeto (2006). A faithfulness ranking projected from a perceptibility scale: the case of [+voice] in Japanese. *Lg* 82. 536–574.
- Kawahara, Shigeto (2012a). The intonation of nominal parentheticals in Japanese. In Borowsky *et al.* (2012). 304–340.
- Kawahara, Shigeto (2012b). Review of Labrune (2012b). *Phonology* 29. 540–548.
- Kawahara, Shigeto (2015a). The phonetics of sokuon, or geminate obstruents. In Kubozono (2015c). 43–73.
- Kawahara, Shigeto (2015b). The phonology of Japanese accent. In Kubozono (2015c). 445–492.
- Kawahara, Shigeto, Musashi Homma, Takaki Yoshimura & Takayuki Arai (2016). MyVoice: rescuing voices of ALS patients. *Acoustical Science and Technology* 37.
- Kawahara, Shigeto & Takahito Shinya (2008). The intonation of gapping and coordination in Japanese: evidence for Intonational Phrase and Utterance. *Phonetica* 65. 62–105.
- Kochetov, Alexei (2014). Japanese in the typology of nasal place assimilation: electropalatographic evidence. Paper presented at the 7th Formal Approaches to Japanese Linguistics Meeting, Tokyo.
- Kolinsky, Régine, José Morais & Mireille Cluytens (1995). Intermediate representations in spoken word recognition: evidence from word illusions. *Journal of Memory and Language* 34. 19–40.
- Kubozono, Haruo (1989). The mora and syllable structure in Japanese: evidence from speech errors. *Language and Speech* 32. 249–278.
- Kubozono, Haruo (1999a). Kayoo-ni okeru moora-to onsetsu. [The syllable and mora in song.] In *Bunpou to onsei*. [Speech and grammar.] Vol. 2. Tokyo: Kuroshio. 241–260.
- Kubozono, Haruo (1999b). Mora and syllable. In Natsuko Tsujimura (ed.) *The handbook of Japanese linguistics*. Malden, Mass. & Oxford: Blackwell. 31–61.
- Kubozono, Haruo (2001). Epenthetic vowels and accent in Japanese: facts and paradoxes. In Jeroen van de Weijer & Tetsuo Nishihara (eds.) *Issues in Japanese phonology and morphology*. Berlin & New York: Mouton de Gruyter. 113–142.
- Kubozono, Haruo (2003). The syllable as a unit of prosodic organization in Japanese. In Caroline Féry & Ruben van de Vijver (eds.) *The syllable in Optimality Theory*. Cambridge: Cambridge University Press. 99–122.
- Kubozono, Haruo (2015a). Diphthongs and vowel coalescence. In Kubozono (2015c). 215–249.
- Kubozono, Haruo (2015b). Loanword phonology. In Kubozono (2015c). 313–361.
- Kubozono, Haruo (ed.) (2015c). *The handbook of Japanese phonetics and phonology*. Berlin: de Gruyter Mouton.
- Kubozono, Haruo, Junko Ito & Armin Mester (2008). Consonant gemination in Japanese loanword phonology. In The Linguistic Society of Korea (ed.) *Current issues in unity and diversity of languages: collection of the papers selected from the 18th International Congress of Linguists*. Republic of Korea: Dongam. 953–973.
- 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**. 113–152.
- Lehiste, Ilse (1970). *Suprasegmentals*. Cambridge, Mass.: MIT Press.
- Lehtonen, Jaakko (1970). *Aspects of quantity in Standard Finnish*. Jyväskylä: Jyväskylä University Press.
- 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. *JASA* **35**. 1773–1781.
- McCarthy, John J. & Alan Prince (1993). Generalized alignment. *Yearbook of Morphology 1993*. 79–153.
- McCawley, James D. (1968). *The phonological component of a grammar of Japanese*. The Hague & Paris: Mouton.
- Macmillan, Neil A. & C. Douglas Creelman (2005). *Detection theory: a user's guide*. 2nd edn. Mahwah: Erlbaum.
- Manabe, Noriko (2009). *Western music in Japan: the evolution of styles in children's songs, hip-hop and other genres*. PhD dissertation, City University of New York.
- Nakamura, Miyoko & Régine Kolinsky (2014). Multiple functional units in the pre-attentive 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 the Japanese geminate stop /tt/: the effect of pitch type and acoustic characteristics of preceding/following vowels.] *Onsei no Kenkyuu [Studies in Phonetics]* **7**. 70–76.
- Okada, Hideo (1999). Japanese. In *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press. 117–119.
- Otake, Takashi & Anne Cutler (eds.) (1996). *Phonological structure and language processing: cross-linguistic studies*. Berlin & New York: Mouton de Gruyter.
- 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 F., Salman Al-Ani & Shosaku Maeda (1980). Temporal compensation and universal phonetics. *Phonetica* **37**. 235–252.
- Port, Robert F., Jonathan Dalby & Michael O'Dell (1987). Evidence for mora timing in Japanese. *JASA* **81**. 1574–1585.
- Poser, William J. (1990). Evidence for foot structure in Japanese. *Lg* **66**. 78–105.
- Prince, Alan & Paul Smolensky (1993). *Optimality Theory: constraint interaction in generative grammar*. Ms, Rutgers University & University of Colorado, Boulder. Published 2004, Malden, Mass. & Oxford: Blackwell.
- Selkirk, Elisabeth (2005). Comments on intonational phrasing in English. In Sónia Frota, Marina Vigário & Maria João Freitas (eds.) *Prosodies: with special reference to Iberian languages*. Berlin & New York: Mouton de Gruyter. 11–58.
- Shibata, Takeshi (1958). Onsei-sono honshitsu to kinoo. [Sounds: their natures and functions.] In Ryu Kumazawa *et al.* (eds.) *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 & University of California, Merced.
- Sugito, Miyoko (1998). *Nihongo onsei no kenkyu*. [Studies on Japanese phonetics.] Vol. 5: 'Hana' to 'hana'. ['Flower' and 'noise']. 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**. 79–112.
- Tamaoka, Katsuo & Yasushi Terao (2004). Mora or syllable? Which unit do Japanese use in naming visually presented stimuli? *Applied Psycholinguistics* **25**. 1–27.
- Tanaka, Shin-Ichi (2013). Review of Labrune (2012a). *Onsei Kenkyu* [*Journal of the Phonetic Society of Japan*] **17:1**. 70–80.
- 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: the analysis of 'ji-amari' in senryu poems and baseball cheering.] In *Bunpou to onsei*. [*Speech and grammar*.] Vol. 2. Tokyo: Kuroshio. 261–290.
- Tanaka, Shin'ichi (2000). Nihonkayookyoku-ni okeru rizumu-no henshen. [The changing rhythm of Japanese popular songs.] *Proceedings of the 14th Meeting of the Phonetic Society of Japan*. 153–158.
- Tanaka, Shin'ichi (2012). Syllable neutralization and prosodic unit in Japanese Senryu poems. Paper presented at the Metrics, Music and Mind conference, Rome.
- Turk, Alice, Satsuko Nakai & Mariko Sugahara (2006). Acoustic segment durations in prosodic research: a practical guide. In Stefan Sudhoff, Denisa Lenertová, Roland Meyer, Sandra Pappert, Petra Augurzky, Ina Mleinek, Nicole Richter & Johannes Schließer (eds.) *Methods in empirical prosody research*. Berlin & New York: de Gruyter. 1–27.
- Vance, Timothy J. (1987). *An introduction to Japanese phonology*. Albany: State University of New York Press.
- Vance, Timothy J. (2008). *The sounds of Japanese*. Cambridge: Cambridge University Press.
- Vance, Timothy J. (2013). Review of Labrune (2012a). *Lingua* **123**. 168–174.