No consonant-final stems in Japanese verb morphology

Kuniya Nasukawa

Tohoku Gakuin University

Abstract

Japanese has been conventionally considered to show two distinct distributional regularities holding in morpheme-final position: at the lexical level, verb stems can end with either a vowel or a consonant, whereas other morpheme types and derived/inflected forms must end with either a vowel or the placeless nasal *N*. Mono-stratal models of phonology call into question the validity of the traditional underlying-surface distinction, and respond by allowing only a single type of static pattern. Unlike the standard derivational approach, however, the mono-stratal approach cannot express the disparity between lexical and non-lexical patterns (Scobbie, Coleman and Bird, 1996; Harris, 1997, 2004). This paper adopts a mono-stratal approach to investigate regularities in the pattern of morpheme/word-final segments in Japanese, and challenges the widespread view that Japanese employs consonant-final verb stems (e.g. *jom* 'read'). Instead, it claims that the actual shape of apparently consonant-final stems is the same as that of non-past forms which end with *ɯ* (e.g. *jomɯ*), the neutral vowel of Japanese. This has the result of excluding consonant-final stems from the Japanese system — a system which prefers vowel-final forms. In terms of phonology-phonetics mapping, the stem-final *ɯ* is the phonetic manifestation of a melodically empty nucleus. This analysis has the benefit of allowing us to eliminate operations such as *ɯ*-deletion and vowel alternation. By positing a morphologically driven process which avoids empty positions at the morphological boundary, the empty position is simply filled by the initial vowel of the suffix (e.g. jo.m∅ *jomɯ* + -a.na.i *anai* NEGATIVE → jo.ma.na.i *jomanai*).

Keywords: static regularities; verb stems; mono-stratal approach, empty positions; overlapping

1. Introduction

It is widely accepted that words/morphemes contain at least two kinds of information: phonological and grammatico-semantic (Halle, 1997:92). The former specifies the phonological shape of morphemes/words in the mental lexicon, while the latter concerns information on, for example, lexical categories, grammatical peculiarities, and meanings. This paper focuses on phonological information, and in particular, on the regular patterns observed in morpheme-final position in Japanese. In addition, it considers what kind of representation is appropriate for describing these morpheme-final regularities.

The discussion adopts a mono-stratal approach to phonological derivation (Harris and Lindsey, 1995; Kaye, 1995; Takahashi, 2004; Nasukawa, 2005, 2007); this approach claims that syllable/prosodic structure cannot be excluded from the lexicon; it must be lexically specified. Furthermore, it assumes that lexical representations are sufficiently complete to be read as phonological representations that can be accessed by sensori-motor systems. Employing this approach, I challenge the traditional view that Japanese verb stems divide into two types: consonant-final (e.g. *jom* 'read') and vowel-final (e.g. ***mi*** 'watch, see').[1] Recognising both patterns requires us to assume that Japanese imposes two different regularities on the shape of verb stems during derivation: lexical and non-lexical. At the lexical level, verb stems can end with either a vowel or a consonant (such as *-r*, *-s*, *-k*, *-g*, *-m*, *-n*, *-b*, *-t*, *-w*), whereas in derived/inflected verb forms (following suffixation) the range of domain-final segments is more limited: only vowels and the placeless nasal *N* are permitted. This disparity is typically observed in multi-stratal derivational models such as those based on SPE, as well as in mono-stratal models such as Optimality

[1] Bloch (1946) employs the term 'root' (rather than 'stem') for describing morphemes such as *jom* which end with a consonant. The term 'stem' is reserved for V-ending forms such as *mi* ('see') and *jomi* (< *jom* 'read' + *i* the epenthetic vowel).

Theory.[2] These models have not postulated the reason why only verb stems lexically permit C-ending in the system which at the surface level allows only vowel and the placeless nasal ending.

By contrast, the mono-stratal approach to be employed here cannot accommodate these two distinct static regularities, since it allows only a single type of static pattern in phonology. In order to facilitate this mono-stratal approach, I adopt a licensing-constrained model of syllable structure (Harris, 1997; Takahashi, 2004; Nasukawa, 2007) and a set of monovalent phonological primes (elements: Harris, 2005; Nasukawa and Backley, 2008; Backley and Nasukawa, 2009) for representing segmental structures. I go on to claim that there are no consonant-final verb stems in Japanese, arguing that what appears to be a final consonant is always followed by a melodically empty vowel which phonetically manifests itself as the neutral vowel *ɯ*; the whole expression then corresponds to its present tense form (e.g. *jom∅* interpreted as *jomɯ*). This leads to the claim that, like other free morphemes in Japanese, verb stems can stand independently — that is, they do not require the presence of other morphemes such as suffixes in their non-past forms. This view departs from traditional analyses in which Japanese verb forms are thought to involve no free morphemes because they are seen as unable to stand alone without suffixation.[3]

The structure of this paper is as follows. Section 2 describes the static distribution of word-final/morpheme-final segments in relation to the CV dichotomy in Japanese and the notion of moraicity. In addition, it discusses how we should treat the disparity between lexical and non-lexical static regularities within a mono-stratal model of phonological derivation. In section 3, I examine the question of which prosodic/syllabic structures are appropriate for representing word-final/morpheme-final static distribution. Section 4 develops an account of Japanese verb inflection by adopting a government-based licensing theory of prosodic representation and a monovalent approach to melodic representation. It proposes that morphemes always end in a nucleus, in both their lexical and interpreted forms; this nucleus may be either empty or melodically filled. Section 5 discusses the necessity of full specification of prosodic structure in lexical representations. Section 6 concludes the discussion.

2. Static regularities

2.1. Static CV distribution in Japanese morphemes/words

Putting aside theory-internal variations, in generative grammar it is the mental lexicon which contains details of the static phonological patterns that characterise the shape of morphemes/words. This type of patterning controls the ability of sounds to occur in particular contexts within the phonological string. To describe such regularities, sounds are first of all divided into two major groups: C[onsonant]s and V[owel]s. The traditional CV dichotomy is a familiar starting point in the description of phonological patterns within morphemes/words. By employing C and V descriptions, languages naturally divide into two types according to their morpheme-final/word-final segment: (1a) languages of the CV type permit only a final V (i.e. they permit only ‘open’ syllables), whereas (1b) languages of the CVC type allow either a final V or a final C (i.e. they permit both ‘open’ and ‘closed’ syllables) (Harris and Gussmann, 1998:142). Examples of the (1a) type are Zulu and Italian, while (1b) is exemplified by languages such as Luo and English.[4]

---

[2] In both cases syllable structure is lexically unspecified, and constructed during derivation (for SPE) or via constraint interaction (for Optimality Theory).

[3] Nishiyama (1996, p.c.) claims that like other lexical categories V-final stems (e.g. *mi* ‘see’, *ne* ‘sleep’, *tabe* ‘eat’, *tazɯne* ‘ask’) can stand independently in coordinate contexts, as in Taroo-wa tookɯ-o *mi*, Hanako-wa tokei-o mita ‘Taroo looked over and Hanako looked at her watch’.

[4] Kaye (1990:323–325) claims that there are two further language types, these being identified on the basis of whether they allow *internal* closed syllables or not. Incorporating this additional parametric choice, (1) is re-tabulated as follows (a full stop indicates a syllable boundary).

VC typology

final …V<u>C</u>]?	internal …V<u>C</u>.C…?			
	No		yes	
no	I	…V.CV]	III	…V(C).CV]
yes	II	…V.CV(C)]	IV	…V(C).CV(C)

Thus, the interaction of two separate parametric choices defines a four-way (rather than a two-way) typology. Type I languages such as Zulu, Swahili and Yoruba disallow both word-internal codas (i.e. VC. sequences) and word-final Cs. In contrast, Type IV languages such as English, French and Polish allow word-internal codas and word-final Cs. Type II languages such as Luo, Malay

(1)

final …V<u>C</u>]?			
no	a	…CV]	e.g. Zulu, Italian
yes	b	…CV(C)]	e.g. Luo, English

Japanese is generally considered to be a language which favours vowel-final morphemes. Indeed, most Japanese morphemes/words end in a vowel, whether they are nouns, adjectives, case markers or postpositions:

(2) a. Nouns
sakana ‘fish’
tamago ‘egg’
tegami ‘letter’
kaze ‘wind’

b. Adjectives (*i* = adjectival suffix)
ooki-i ‘big’
taka-i ‘high’
too-i ‘far’
samɯ-i ‘chilly’

c. Case markers
-ga nominative
-o accusative
-ni dative
-no genitive

d. Postpositions
de ‘at’ (e.g. *ɯtɕi-de* ‘at home’)
e ‘to’ (e.g. *gakkoo-e* ‘to school’)
to ‘with’ (e.g. *tomodatɕi-to* ‘with a friend’)
kaɾa ‘from’ (e.g. *yama-kaɾa* ‘from mountain’)

Also, in borrowing words from other languages, Japanese add a vowel after a word-final consonant of original words: typically the high back unrounded vowel *ɯ* appears post-consonantally (e.g. *slip* ‘slip’ > *sɯɾipp**ɯ***: where a consonant sequence is broken up by inserting *ɯ*). The fact that most Japanese morphemes are vowel-final suggests that Japanese may be a CV language according to the above typology. However, there exists one ‘non-vowel’ which can appear word-finally: this is the placeless nasal.

(3) Nouns with a final *ɴ*
hoɴ ‘book’
teɴ ‘dot, spot’

As discussed in S. Yoshida (2003), *ɴ* is rare in native morphologically simplex words, but commonly found in Sino-Japanese words, such as those in (3), which are the Japanese readings of Chinese characters borrowed from Ancient Chinese. Other types of loanwords also allow word-final *ɴ* (e.g. *wɒʃɪŋtən* ‘Washington’ → *waɕintoɴ*), in which case the placeless nasal is typically restricted to nouns that are open class uninflected words in Japanese.

If we regard the placeless nasal as a consonant, then Japanese may be classified as a CVC type language which permits consonant-final endings.

2.2. The placeless nasal as a vowel in terms of moraicity

However, the word-final placeless nasal differs from other consonants in terms of its ‘mora’ value. It is widely acknowledged in the literature (Kubozono, 1999; et passim) that native speakers of Japanese (in most dialects, at least) detect a ‘beat’ called moraicity in the placeless nasal. Moraicity is primarily a property of vowels, where it is associated with a similar ‘beat’ or ‘timing’ value and can bear a tone. For example, *i* ‘stomach’ and *me* ‘eye’ both have one beat, whereas *ɯde* ‘arm’ has two beats and *atama* ‘head’ three beats. Evidently, moraicity corresponds to the number of vowels in a word. As ‘timing-bearing’ units, vowels also support pitch contours in Japanese. For example, *ɯde* and *atama* are segmented into two and three parts, respectively: *ɯ-de* (with the tone LH) and *a-ta-ma* (with the tone LHH).

---

and Yucatec Maya allow word-final Cs but do not tolerate word-internal codas, whereas Type III languages such as Italian, Telugu and Tamil exclude word-final Cs but do allow word-internal codas.

Unlike other consonants, the placeless nasal *N* is treated as an independent unit for moraic purposes, just as vowels are, and is considered to bear one 'beat': for example, *kabaN* 'bag' has three beats rather than two, and is segmented into three parts *ka-ba-N* rather than two chunks such as **ka-baN*. As for its pitch contour, *ka-ba-N* is realised as LHH. Also, a pitch contrast is found in the placeless nasal: an example is *o-N* (LH) 'sound' versus *o-N* (HL) 'gratitude'.

Another well-known piece of evidence which highlights the vowel-like behaviour of this special consonant is found in the pitch accentuation of foreign loanwords: pitch accent is assigned on the antepenultimate vowel in (non-nativised) words of (typically) more than four vowels, as shown below. Note that the following vowel sequences are not regarded as diphthongs or long vowels. The places in (4) where pitch accent is assigned are underlined, and the symbol ' indicates the point where a falling pitch begins.

(4)	a.	kɯ-ɾi-s<u>ɯ</u>'-ma-sɯ	LH<u>H</u>LL	'Christmas'
	b.	ko-sɯ-t<u>a</u>'-ɾi-ka	LH<u>H</u>LL	'Costa Rica'
	c.	sɯ-to-ɾ<u>a</u>'-i-ki	LH<u>H</u>LL	'strike'
	d.	pɯ-ro-gɯ-ɾ<u>a</u>'-ma-a	LH<u>H</u>LL	'programmer'

If a word contains the placeless nasal in word-final position (i.e. ...*N*]), then the accent goes on the penultimate vowel. This is exemplified below.

(5)	a.	bi-t<u>a</u>'-mi-N	L<u>H</u>LL	'vitamin'
	b.	sɯ-p<u>e</u>'-i-N	L<u>H</u>LL	'Spain'

Now if we assume that the placeless nasal is a tone-bearing unit like vowels, then we can make the generalisation that the accent falls on the antepenultimate tone-bearing unit, whether this is a vowel or the placeless nasal consonant.

Furthermore, the placeless nasal also behaves like a vowel in other phonological processing tasks such as transposition in speech errors (Kubozono, 1985), secret language games (e.g. Babibu language: Haraguchi, 1991) and particle vowel reduction in casual speech (Hasegawa, 1979). Some other examples are found in Kubozono (1999). These phenomena suggest that the placeless nasal might possibly be viewed as a vowel (Nasukawa, 1998, 2004).

In the literature, at least two proposals have been put forward in which the placeless nasal is analysed as a vowel: in S. Yoshida (1990), *N* occupies the vocalic (nuclear) position of a CV syllable, while in Nasukawa (1998, 2004) it is argued that *N* consists of a two-position sequence (an onset and a following nucleus) rather than a single syllabic position. In this two-position approach the melodic properties of nasality and stopness occupy the onset, and these are together realised as a nasal stop without any place of articulation. Its vowel-like characteristics come from the realization of a following nucleus, which contains no melodic material. Both S. Yoshida and Nasukawa are able to capture the vocalic characteristics of the placeless nasal by referring crucially to the presence of a vocalic/nuclear position. For an evaluation of these two approaches in terms of phonological restrictiveness, see Nasukawa (2004:50–52).

2.3. Verb morphology in Japanese

Verb stems do not necessarily conform to the patterns described above. It has become an established view in the literature (cf. Bloch, 1946:7–10; McCawley, 1968:93–110; Ashworth, 1976–77; Maeda, 1979; Tabata, 1983; Ito, 1986; Poser, 1986; Vance, 1987:184–193; Davis and Tsujimura, 1991; Tsujimura, 1996:40-54) that Japanese verb stems fall into two types: V[owel]-ending and C[onsonant]-ending. Unlike other types of morphemes, the lexicon contains more C-ending stems than V-ending stems. Examples of the two types are given in (6).

(6) a. Vowel-ending

stem		non-past (-*ɾɯ*)	negative (-*nai*)	past (-*ta*)
mi	'watch, see'	*mi-ɾɯ*	*mi-nai*	*mi-ta*
ki	'wear'	*ki-ɾɯ*	*ki-nai*	*ki-ta*
ne	'sleep'	*ne-ɾɯ*	*ne-nai*	*ne-ta*
tabe	'eat'	*tabe-ɾɯ*	*tabe-nai*	*tabe-ta*
tazɯne	'ask'	*tazɯne-ɾɯ*	*tazɯne-nai*	*tazɯne-ta*

b. Consonant-ending

stem		non-past (-*ɯ*)	negative (-*anai*)	past (-(*i*)*ta*/-(*i*)*da*)
tas	'add'	*tas-ɯ*	*tas-anai*	*tas-ita*
kiɾ	'cut'	*kiɾ-ɯ*	*kiɾ-anai*	*kit-ta*
keɾ	'kick'	*keɾ-ɯ*	*keɾ-anai*	*ket-ta*
kak	'write'	*kak-ɯ*	*kak-anai*	*ka-ita*
tog	'sharpen'	*tog-ɯ*	*tog-anai*	*to-ida*
jom	'read'	*jom-ɯ*	*jom-anai*	*jon-da*
ɕin	'die'	*ɕin-ɯ*	*ɕin-anai*	*ɕin-da*
tob	'fly'	*tob-ɯ*	*tob-anai*	*ton-da*
tat	'stand'	*tat-ɯ* (> *tatsɯ*)	*tat-anai*	*tat-ta*
kaw	'buy'	*kaw-ɯ* (> *kaɯ*)	*kaw-anai*	*kat-ta*

According to this established view, verb morphemes are all bound in Japanese and must have a suffix attached to them: no free morphemes including verb stems are involved in verb morphology. In non-past tense forms, either -*ɾɯ* or -*ɯ* appears at the end of stems: -*ɾɯ* is added when the verb stem ends with a vowel such as *i*, *e* or *ɯ*, while -*ɯ* appears when the stem ends with a consonant such as -*ɾ*, -*s*, -*k*, -*g*, -*m*, -*n*, -*b*, -*t* or -*w*. To form the negative, -*nai* or -*anai* attaches to the stem: -*nai* follows vowel-ending stems while -*ɯ* follows consonant-ending stems. In past forms, -*ta* is suffixed to vowel-ending stems while consonant-ending stems are typically followed either by -(*i*)*ta* or -(*i*)*da*. Furthermore, in the case of past tense forms of consonant-ending stems, processes such as voicing assimilation, stem-final consonant deletion and gemination are found. These do not operate in forms with vowel-ending stems.

In addition to these three suffixes, we also find other suffixes displaying different allomorphs that are sensitive to the presence of a V-ending or a C-ending. Examples are given below, together with further examples of the suffixes already shown above.

(7)		V-final verb stems	C-final verb stems
a.	Non-past	-*ɾɯ*	-*ɯ*
b.	Negative	-*nai*	-*anai*
	Polite	-*masɯ*	-*imasɯ*
c.	Volitional	-*joo*	-*oo*
	Causative	-*saseɾɯ*	-*aseɾɯ*
d.	Past (perfective)	-*ta*	-*ita*, -*tta*, -*da*
	Gerundive	-*te*	-*ite*, -*tte*, -*de*
	Representative	-*taɾi*	-*itaɾi*, -*ttaɾi*, -*daɾi*

An obvious regularity emerges in which a consonant (C) appears after V-final stems, whereas C-final stems are followed by a vowel (V). Only the suffixes preceded by the C-final verb stems in (7d) show a more complex pattern, in which not only V-initial but also C-initial suffix forms appear, such as -*tta*, -*tte* and -*da*, -*de*.

A further point to be noted is that in (7a) and (7c) the length of suffixes tends to be shorter after C-final stems than after the V-final stems. On the other hand, in (7b) and (7d) suffix length tends to be longer after C-final stems than after V-final stems.

With the exception of some dialect forms (Shibatani, 1990), this kind of word formation pattern, in which the type of stem-final segment is crucial, is peculiar to the verb conjugation and is never found in other morpheme types in Japanese.

2.4. A disparity between the two static patterns

We have now seen the two distinct static patterns which operate morpheme-finally in Japanese: at the lexical level in (8a), verb stems can end with either a vowel or a consonant (such as *-r*, *-s*, *-k*, *-g*, *-m*, *-n*, *-b*, *-t*, *-w*), while other morpheme types end with either a vowel or the placeless nasal *N*; by contrast, at the interpreted level in (8b), only vowels and the placeless nasal *N* are allowed to occupy the domain-final position (of derived/inflected verb forms following suffixation).

(8)

		… CV]	… V*N*]	VC]
a.	Lexical (underived)	✓ (most morphemes)	✓ (only nouns)	✓ (only verb stems)
b.	Interpreted (derived)	✓ (most morphemes)	✓ (only nouns)	

In order to explain this disparity between the lexical and interpreted levels, a standard derivational approach typically appeals to a general model of the phonetics-phonology interface such as the one shown below (Harris, 2004:103–104):

(9) a. Underlying phonology ← (8a)
↓
b. Surface phonology
↓
c. Categorial phonetics

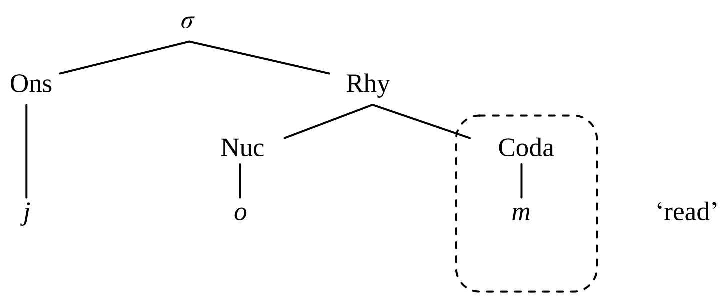
} ← (8b)

Articulation ← → Auditory perception

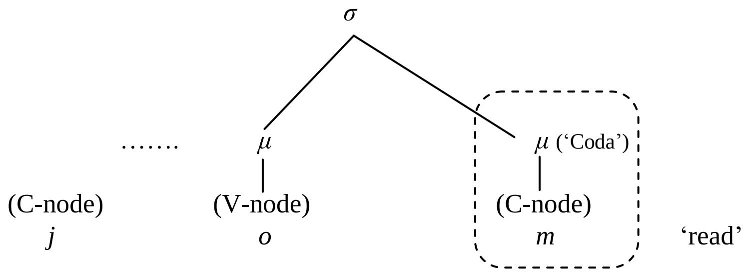
In this multi-stratal derivation, the kind associated with SPE, the regularity (8a) is assumed to operate at the level marked (9a), and (8b) is assumed to work somewhere between (9b) and (9c).

The disparity may be accounted for by referring to both representational and functional aspects of phonology. A typical approach to the syllabification of domain-final Cs in Japanese morphemes is adopted in Abe (1987) and Vance (2008), where the verb-stem-final consonant occupies a coda, as shown in (10a) and (10b) respectively.

(10) a.

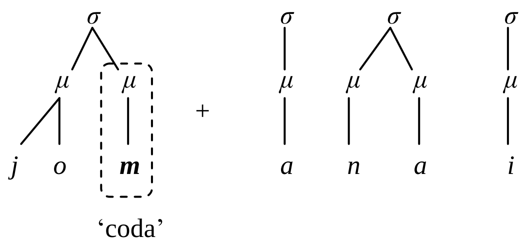


b.

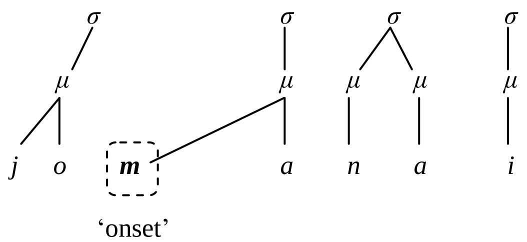


In rule-based multi-stratal model, based on lexically-given precedence relations between bundles of features, the syllabification takes place though the serial application of extrinsically ordered rules during the course of derivation (Bromberger and Halle, 1989). Assume that morphological concatenation takes place after the verb stem is syllabified. When this coda approach is employed in conjunction with the derivational scheme shown in (9), we find morphological concatenation triggering resyllabification: e.g. *jo**m***. ‘to read’ + *-a.na.i* NEGATIVE in (11a) > *jo.**m**a.na.i* ‘(somebody) does not read’ in (11b).

(11) a. *jo**m***. ‘to read’ + *-a･na･i* NEGATIVE



b. → *jo.**m**a.na.i* ‘(somebody) does not die’



This kind of syllabic reorganization suggests that there is a need to override particular structural conditions holding at the earlier stage of derivation. However, structure-changing operations of this sort are undesirable since there seems no clear motivation for losing the mora associated with *m*. This provides a negative impact on the degree of restrictiveness that the model achieves, which in turn weakens the analysis. In most versions of the licensing/government-based framework, this is interpreted as a failure to conform to the Structure Preservation principle.

(12) STRUCTURE PRESERVATION PRINCIPLE (cf. Harris, 1994; Backley and Takahashi, 1998; Takahashi, 2004; Nasukawa, to appear)
Dependency (licensing) relations holding of lexical representations also hold at the interface with performance systems.

Because this framework does not allow any reorganization of lexically-determined structural relations, a lexically given structure such as (11a) must be the same when it is interpreted by the performance facilities.

In recent years there has been a general shift in approach towards output-oriented models of phonology, as discussed in Harris (2004:103–104) and elsewhere; and this has raised serious doubts about the validity of the underlying-surface distinction. In response to this, theories which place an emphasis on restrictiveness and which abandon the distinction altogether have developed; these include Government Phonology (Kaye, Lowenstamm and Vergnaud, 1990; Harris, 1997), Declarative Phonology (Scobbie, Coleman and Bird, 1996) and classic Optimality Theory (Prince and Smolensky, 1993, 2004).

(13) Phonology

Articulation ← → Auditory perception

In the mono-stratal approach described in (13) it is not possible to employ the two distinct static regularities, since the model allows only a single type of static pattern in phonology. In the following section, therefore, I examine the kind of regularity which is appropriate for phonologically representing this pattern of word-final/morpheme-final static distribution. I employ the mono-stratal approach described above, and adopt a licensing-constrained model of syllable structure (Harris, 1997; Takahashi, 2004; Nasukawa, to appear) which, unlike classic OT, assumes that not only melodic material but also syllable structure is fully specified in lexical representations. After that, I go on to identify a suitable phonological shape for verb stems in Japanese.

3. Identifying verb stems in mono-stratal phonology: no C-final verb stems

The mono-stratal model must choose between the two ‘informal’ static regularities in (8) for capturing the shape of the domain-final pattern in Japanese morpho-phonology. On the basis of the general preference for vowel-final morphemes/words in the language, I favour the pattern in (8b) over that in (8a) because the latter is never phonetically observed. This choice leads us to reconsider the phonological shape of C-final verb stems. I assume that a possible candidate for the apparently C-final stems is one which is identical to the shape of non-past verb forms; these end in a single vowel (*ɯ*), unlike other forms which end in a longer string containing both consonants and vowels (as seen in the negative *-anai* and the past -(*i*)*ta*/-(*i*)*da*). Given that these non-past forms are themselves verb stems (alternatively considered to be infinitive forms), we can immediately rule out the need for any C-ending verb stems. Consequently, all stem types end in a vowel,[5] so that the typological split in (6) is reanalysed as in (14), where the two types of verb stems in (6) can be relabeled as Types A and B.

(14) a. Type A: ... *i*/*e* *ɾ* *ɯ*]

stem		= non-past
miɾɯ	‘watch, see’	
kiɾɯ	‘wear’	
neɾɯ	‘sleep’	
tabeɾɯ	‘eat’	
tazɯneɾɯ	‘ask’	

[5] Following Nasukawa (1998, 2004), this paper regards the moraic nasal, which typically occupies a noun-final position, not as a consonant but as the phonetic manifestation of an onset-nucleus pair; within this pair, the onset part possesses nasality and stopness while the nuclear part is melodically empty and contributes a moraic beat of the segment in question.

b. Type B: ... *a*/*i*/*ɯ*/*e*/*o* C *ɯ*]

stem		= non-past
tasɯ	‘add’	
kiɾɯ	‘cut’	
keɾɯ	‘kick’	
kakɯ	‘write’	
togɯ	‘sharpen’	
jomɯ	‘read’	
sɯmɯ	‘live’	
ɕinɯ	‘die’	
tobɯ	‘fly’	
tatɯ (> *tatsɯ*)	‘stand’	
kawɯ (> *kaɯ*)	‘buy’	

In comparison with the forms in (6a), which are considered to end either with *i* or *e*,[6] the corresponding forms in (14a) all end with *ɾɯ*. On the other hand, the forms in (14b) all end uniformly with *ɯ* but the preceding consonant varies: *s*, *ɾ*, *k*, *g*, *m*, *n*, *b*, *t* or *w*. Furthermore, in Type B forms, any vowel can appear in the preceding syllable. When it comes to distributional restrictions, there seems to be a strong correlation between the presence of *i*/*e* and the presence of *ɾ*, which will be discussed in section 4.

Now we turn to the construction of negative forms in the two types of verb stem in (14). As depicted in (6), it has been widely supposed that there are two negative suffixes, *-nai* and *-anai*. But in a derivational approach, one of these must be taken to be the lexical form, while the other is derived via morphophonemic operations. According to this view, it is often considered that *-anai* is the lexical form of the negative suffix.

(15)

-anai → -[anai] after Type B stems, as in [jomanai] and [kiɾanai]

-anai → -[nai] after Type A ending stems, as in [minai] and [kinai]

The only alternative is to posit an arbitrary rule inserting *a* between the stem and the suffix *-nai* if this is assumed to be the lexical form: however, there is no obvious reason for choosing *a* over any other segment. In fact *a* makes an unlikely choice compared with *ɯ*, which is nowadays considered to be the default vowel of Japanese (although *i* played the role more than one hundred years ago).

Given that *-anai* is the lexical form of the negative suffix, as illustrated in (16b), *-anai* remains intact after Type B stems while the Type-B-stem-final *ɯ* is not pronounced. In the case of Type A in (16a), on the other hand, not only the stem-final vowel *ɯ* but also its preceding consonant *ɾ* and the initial *a* of *-anai* are suppressed.

(16) a. Type A

stem		negative formation	negative
miɾɯ	‘watch, see’	*mi**ɾɯ** + -**a**nai*	[minai]
kiɾɯ	‘wear’	*ki**ɾɯ** + -**a**nai*	[kinai]
neɾɯ	‘sleep’	*ne**ɾɯ** + -**a**nai*	[nenai]
tabeɾɯ	‘eat’	*tabe**ɾɯ** + -**a**nai*	[tabenai]
tazɯneɾɯ	‘ask’	*tazɯne**ɾɯ** + -**a**nai*	[tazɯnenai]

---

[6] The verb stems ***sɯ**ɾɯ* ‘do’ and ***kɯ**ɾɯ* ‘come’ contain *ɯ* (rather than *i* or *e*) before the word-final *ɾɯ*. When they appear in their past tense forms, ***ɕ**ita* ‘did’ and ***k**ita* ‘came’, *i* occurs before the past-tense ending *ta*. These are usually considered to be exceptional in terms of verb inflection. It is beyond the scope of this paper to examine the historical processes responsible for these verb stems.

b. Type B

stem		negative formation	negative
tasɯ	'add'	*tas**ɯ*** + *-anai*	[tasanai]
kiɾɯ	'cut'	*kiɾ**ɯ*** + *-anai*	[kiɾanai]
keɾɯ	'kick'	*keɾ**ɯ*** + *-anai*	[keɾanai]
kakɯ	'write'	*kak**ɯ*** + *-anai*	[kakanai]
togɯ	'sharpen'	*tog**ɯ*** + *-anai*	[toganai]
jomɯ	'read'	*jom**ɯ*** + *-anai*	[jomanai]
sɯmɯ	'live'	*sɯm**ɯ*** + *-anai*	[sɯmanai]
ɕinɯ	'die'	*ɕin**ɯ*** + *-anai*	[ɕinanai]
tobɯ	'fly'	*tob**ɯ*** + *-anai*	[tobanai]
tatɯ (> *tatsɯ*)	'stand'	*tat**ɯ*** + *-anai*	[tatanai]
kawɯ (> *kaɯ*)	'buy'	*kaw**ɯ*** + *-anai*	[kawanai]

At this point, two questions arise. First, we must explain why stem-final *ɯ*-suppression is commonly observed in both Type A and Type B formations. The second question seeks to explain why the other types of suppression (*ɾ*-suppression in stems and initial *a*-suppression in suffixes) take place in Type A but not in Type B. With respect to the latter question, as we will see in section 4, it can be accounted for by claiming that, in creating an inflected form, a verb stem and the negative suffix avoid having empty prosodic categories (such as empty onset and empty nucleus) around the morphological boundary.

Returning to the question of *ɯ*-suppression, we must consider the characteristics of the stem-final *ɯ*, which is inherently weak and susceptible to suppression. The high back unrounded vowel *ɯ* is often considered to be neutral, as it is the weakest and least prominent in the Japanese vowel system. Because of these characteristics, the vowel is often used as an unmarked vowel inserted into nativised loanwords (e.g. *slim* 'slim' → ***s**ɯ**ɾim**ɯ***, *pli:z* 'please' → ***pɯ**ɾiiz**ɯ***). Since it is considered to be the least prominent member of the Japanese vocalic inventory, it succeeds in being minimally disruptive to a loanword's original sound sequence when it is inserted between or after consonants.[7]

Also *ɯ* is regarded as weak in the sense that it is often masked by a neighbouring nasal or deleted when a nasal is adjacent (e.g. *ɯma* 'horse' → *ᴍma*: Nasukawa, 2004:52). In addition, *ɯ* (as well as the high front vowel *i*) is susceptible to devoicing, especially when flanked by two voiceless obstruents (e.g. *kɯsa* 'grass' → *k**ɯ̥**sa*) and in word-final position following *s* (e.g. *nasɯ* 'aubergine' → *nas**ɯ̥***). Furthermore, in Sino-Japanese compounds *ɯ* is deleted between identical voiceless obstruents, resulting in a geminate consonant (e.g. *kok**ɯ**-ki* → *ko**kk**i* 'national flag', *hats**ɯ**-teɴ* → *ha**tt**eɴ* 'development').[8] These phenomena are assumed to be suggestive of the weakness of *ɯ*, in that it cannot sustain its intrinsic vocalic property in certain contexts. S. Yoshida (2003) presents further arguments supporting the inherent weakness of *ɯ*, while discussions on the unmarked status of *ɯ* are to be found in Tateishi (1990) and Ito and Mester (1996). In Element Theory (Nasukawa, 1998, 2004), the vowel *ɯ* is considered to be the phonetic manifestation of a melodically unspecified (empty) nucleus. In the following section, I investigate the formal melodic structure of *ɯ* and its relation to prosodic structures.

4. Beyond alphabetic symbols

4.1. Empty positions in the CVCV model

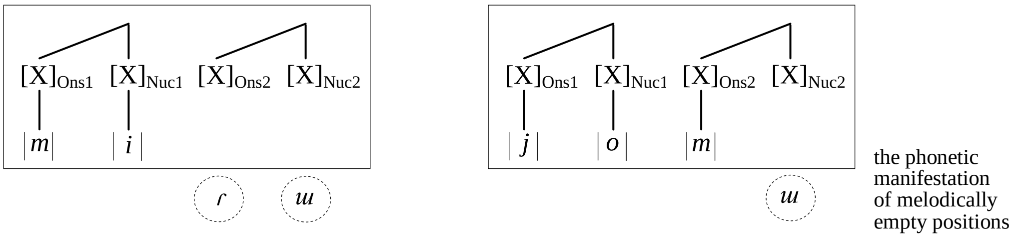
To describe the structures of the Japanese verb inflection, this paper does not utilise a template-based representation of moraic syllable structure of the kind typically found in the OT literature. Instead, I develop an account based on dependency relations of the kind associated with the theory of

[7] Besides *ɯ*, the vowel *o* is also inserted between consonants and after a word-final consonant. Unlike *ɯ*, the mid vowel has a restricted distribution: *o* is inserted into loanwords only after *t* and *d* (e.g. *peɪnt* 'paint' → *peint**o***, *tri:əu* 'trio' → *t**o**ɾio*). This is controlled by the fact that inserting *ɯ* after *t* and *d* creates the unnatural sequences *tɯ* and *dɯ*, which do not belong in the kana syllabary. To avoid these unnatural sequences, Japanese uses the vowel *o*, which is nearest to the default vowel *ɯ* in the Japanese vocalic space. This generates the sequences *to* and *do*.

[8] In Japanese, the front high vowel *i* is also deleted in the formation of a geminate (*itɕ**i**-satsɯ* → ***iss**atsɯ* 'one volume of a book') although the process is typically restricted to the root level morphology (cf. Harris, 1994:18–28).

Government Phonology, also known as Licensing/Government Phonology and Dependency/Government Phonology (henceforth GP: Harris, 1994, 1997; Takahashi, 2004; Nasukawa, 2005, to appear). This framework typically employs a strict CVCV structure for systems referred to as CV languages; these are assumed to have neither codas nor consonant clusters (e.g. Bantu languages such as Bemba: Kula, 2002; Kula and Marten, 2009; Cilungu: Nasukawa, to appear). Let us assume that this describes the sound pattern of Japanese, which gives priority to the CV-dichotomy. On this basis, *miɾɯ* 'watch, see' (a Type A verb stem) and *jomɯ* 'read' (a Type B verb stem) have the structures in (17), where Ons and Nuc stand for onset and nucleus respectively, and the arrows indicate dependency relations (heads are marked by a vertical line, dependents by an oblique line).

(17) a. Type A b. Type B



In the context of this representational model, Japanese prosodic structure is assumed to be strictly CVCV (represented by $[X]_{Ons}$ and $[X]_{Nuc}$), which rules out any branching constituents.

As for Type B stems such as *jomɯ* in (17b), which all end with a high back unrounded vowel *ɯ*, I assume that the melodically empty nucleus is present in their representation: the final position has no melodic material while the first three positions contain sets of melodic features (here abbreviated to italic phonetic symbols). As discussed in section 3, *ɯ* is a phonologically weak and quality-neutral (colourless) vowel which is best described by a structure which is maximally simple. Within GP, Japanese *ɯ* is assumed to be the phonetic manifestation of a nuclear position which contains no melodic material. So, despite the prosodic structure of a Type B verb stem *jomɯ* ending in an empty nucleus, it still conforms to the static distributional pattern for Japanese in (8b) since it is strictly CVCV (represented by $[X]_{Ons}$ and $[X]_{Nuc}$). The empty structure of *ɯ*, as we will see later, straightforwardly explains the first question raised above, which was why stem-final *ɯ*-suppression is commonly observed in both Type A and Type B formations: a melodically unspecified position can be easily suppressed or overwritten.

The notion of 'melodic emptiness' was originally introduced to account for the behaviour of *h-aspiré* in French (Clements and Keyser, 1983:107-109). Later, the application of melodic emptiness was extended to nuclear positions, and applied to the analysis of vowel-zero alternations in languages such as Moroccan Arabic (Kaye, 1990), French (Charette, 1991), Polish (Gussmann and Kaye, 1993), English (Harris, 1994) and Japanese (Nasukawa, 1998, 2005).

Empty positions play a particularly important role in the licensing/government-based model of phonological structure, which offers one of the most restrictive approaches to the representation of prosodic structure. In this framework, a phonologically legitimate position must be phonetically interpreted even if it contains no melodic material. On the other hand, the Phonological Empty Category Principle (Kaye, 1990:313, 1995:295; Harris, 1994:193) states that a melodically empty nucleus receives no phonetic interpretation if and only if it is p[rosodically]-licensed.

(18) THE PHONOLOGICAL EMPTY CATEGORY PRINCIPLE (ECP):
A p-licensed empty category (position) receives no phonetic interpretation.

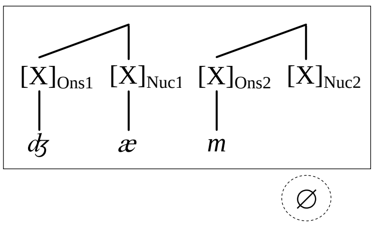
It is argued that P-licensing takes effect under various conditions, one of which involves domain-final position. A parameter can be used to describe, and at the same time, control this kind of variation. A domain-final empty nucleus is p-licensed to be phonetically silent if the setting of the Domain-final-empty-nucleus parameter in (19) is ON. Examples are found in languages such as Luo, English and French. In (19) the bracketed OFF setting indicates the default value for this parameter.

(19) Domain-final-empty-nucleus parameter（Kaye, 1990:314, Harris, 1994:162）
Domain-final nucleus licensed?

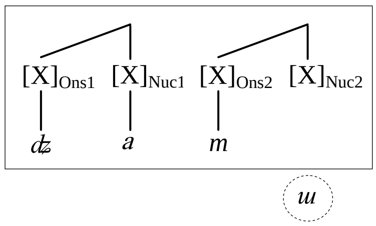
Parameter setting	Examples
ON	Luo, English, French, . . .
[OFF]	Zulu, Italian, Telugu, . . .

There is ample empirical evidence to support the existence of p-licensed empty nuclei, as discussed in Harris & Gussmann (1998, 2002). On the other hand, if the setting of the parameter is OFF, a final empty nucleus is not p-licensed; consequently, the position must receive phonetic interpretation. This is observed in languages such as Zulu[9], Italian[10] and Telugu[11], in which the neutral vowel in the relevant language usually serves as the realization of the unlicensed final empty nucleus (Archangeli, 1984; Kaye, 1990; Charette, 1990; Harris, 1994; Nasukawa, 2005). Another language with the OFF parameter setting is the Bantu language Cilungu, in which the unlicensed empty nucleus is pronounced as *i* (Nasukawa, to appear, cf. Bickmore, 2007). In the case of Japanese, the empty position is typically realised as the high back unrounded vowel *ɯ*, as illustrated in (20b) (S. Yoshida, 1990, 2003; Nasukawa, 1998, 2004). According to this argument, the vowel *ɯ* — which has elsewhere been treated as a segment inserted into nativised loanwords in order to create a CVCV string — is reanalysed as the default phonetic manifestation of an empty nucleus. This is depicted in (21b) for the Japanese loanword *ʥamɯ* ‘jam’; compare this with the original English form *ʤæm* in (21a).

(20) a. [ON] *ʤæm* ‘jam’ in English



b. [OFF] *ʥamɯ* ‘jam’ in Japanese



Turning to the representation of the Type A stem in (17a), the structure of *miɾɯ* is assumed to end with not only an empty nucleus but also an empty onset preceding the nucleus. As we will see below in more detail, *ɾ* of *ɾɯ* (which is phonetically described as a coronal tap) is also represented by a melodically unspecified position. As in the case of *ɯ*, we can also answer the question raised earlier concerning the reason why *ɾ*-suppression occurs in Type A negative formation: a melodically empty position can be naturally suppressed or overwritten.

Evidence for the status of *ɾ* as a weak consonant in present-day Japanese comes from the way it is often targeted by phonological processes and undergoes suppression or assimilation to a trigger (e.g. *haɕiranai* ‘someone does not run’ → *haɕi**nn**ai* in Tokyo Japanese, *haɕiɾɯbe* ‘let us run’ in general Tohoku Japanese → *haɕi**pp**e* in Sendai Japanese).[12] In addition, *ɾ* (the tap) is often identified as the lenited reflex of *t* (e.g. *mita* ‘flag’ → *miɾa*) in some dialects (as in Koshikijima Island and some Tohoku regions).

This kind of lenition is not exclusive to Japanese but is attested cross-linguistically. A typical example is *t*-tapping in English (e.g. ci*t*y → ci*ɾ*y) where the process is assumed to be represented as the suppression of non-place primes. In one version of Element Theory (Nasukawa and Backley 2008; Backley and Nasukawa, 2009, 2010; cf. Harris and Kaye, 1990; Harris, 1994; Harris and Lindsey, 1995),

---

9 Although this view is controversial, we shall assume that surface NC sequences in Zulu are single, prenasalised segments which are syllabified in an onset (cf. Herbert 1975, 1986; Maddieson, 1990; Maddieson and Ladefoged, 1993; Kula, 2002; Nasukawa, to appear).

10 Although Italian allows word-final *r*, *l* and *n* on the surface, these sonorants are here regarded as the phonetic manifestation of sequences comprising an onset (containing consonantal properties) and an empty nucleus (contributing sonorancy).

11 Here surface homorganic NC sequences in Telugu are also considered to be single, prenasalised segments which are syllabified in an onset.

12 The framework of GP rules out any structural (prosodic) changes: so in the case of *haɕi**nn**ai* < *haɕi**ra**nai*, for example, it assumes there is an empty nucleus between the two nasals (*haɕi**n**∅**n**ai*), where *a* before the nasal in *haɕi**ra**nai* is phonetically suppressed and *ɾ* assimilates to *n* (cf. Nasukawa, 2005).

a privative model of segmental structure, *t* consists of |ʔ| (edge element), |H| (noise element) and |I| (coronal/dip element). The process is represented as the suppression of the first two elements |ʔ| and |H|, which leaves the remaining element |I| to be phonetically interpreted as the tap *ɾ*.

The corresponding sound in Japanese (the coronal tap) has had little attention in the literature on Japanese phonology. This paper, however, assumes that *ɾ* in Japanese has the same structure as the English tap: *ɾ* is the phonetic manifestation of a single |I|.[13] If so, we need to answer a question why the empty onset followed by an empty nucleus in Type A forms such as in (17a) is phonetically interpreted as *ɾ* when the Japanese tap is assumed to be the phonetic realisation of a sole |I|.

To answer the question, as in the case of an empty nucleus, we need to look at the condition which allows an empty onset to be phonetically silent. It is assumed that an onset may be silent if it respects the the principle in (21), which functions at the level of the minimal prosodic domain. A minimal prosodic domain comprises an onset and a nucleus, though these do not constitute a formal unit such as 'syllable' (see Takahashi (2004) who presents arguments to challenge the formal status of the syllable constituent).

(21) CONDITION ON THE PHONETIC MANIFESTATION OF EMPTY ONSETS
An empty onset can be phonetically silent if and only if it is prosodically-licensed by a following melodically 'filled' nucleus.

This condition functions within a minimal prosodic domain comprising an onset and a nucleus, and is considered to be a type of Proper Government controlling consonant positions.[14]

When it comes to empty onsets, a parallel situation arises since these must behave like empty nuclei and conform to the ECP in (18). If an empty onset is not followed by a filled nucleus, it must be phonetically realised as a consonant sound. In the case of Type A verb stems in Japanese, the empty onset seeks out any phonological sources which contribute to its phonetic interpretation. In the context of Type A forms, the preceding nucleus of the empty onset (which is followed by an empty nucleus) is the only position which is filled by a melodic material and which can therefore act as a melodic source. As already discussed in the previous section, interestingly *i* or *e* occurs in the nucleus preceding the empty onset (see (14a)). In terms of element composition,[15] *i* consists of a single |I| (which phonetically contributes frontness in nuclei) while *e* contains the same |I| and also |A| (the mass element, which contributes openness). Crucially, both segments contain |I|, and we assume that the empty onset phonetically manifests itself as a segment with the help of this preceding |I|.

---

[13] |I| covers coronals including dentals and palatals: when |I| is non-headed it is interpreted as dental, and when headed it is phonetically associated with palatality. For a detailed discussion, see Nasukawa and Backley (2008)

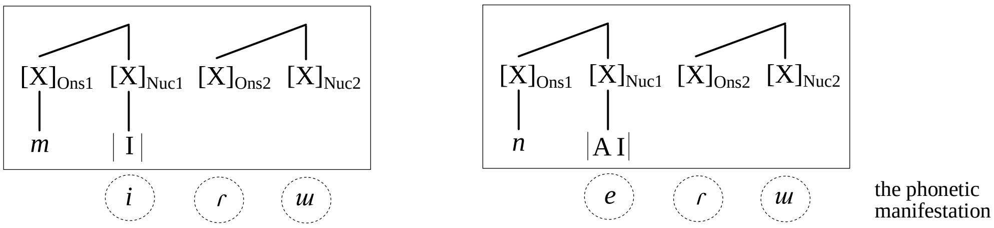
[14] In GP there are several ways to p-license (phonetically suppress) an empty position (Charette, 1998:170). One way is via Proper Government, which applies if a structure meets all the conditions below.

Proper Government (Kaye, 1990:314; Harris, 1994:191):
A nucleus α properly governs an empty nucleus β if and only if
a. α and β are adjacent on the relevant nuclear projection,
b. α is not itself p-licensed, and
c. no governing domain separates α from β.

As an example, the English word *fæməli* 'family' conforms to (a), where the word-medial empty nucleus (represented by '∅': /fæ.m∅.li/) is immediately followed by the final filled nucleus ('i') at the level of nuclear projection. In this case, p-licensing has the opportunity to apply, causing the empty nucleus to be *properly governed* by the final filled nucleus. (The structure also conforms to (b) and (c) although there is no space to discuss these.) The p-licensed empty nucleus receives no phonetic interpretation, and consequently, *fæməli* is pronounced *fæmli* (fa.m∅.li). In a variant pronunciation of the same word, p-licensing does not apply, and the empty nucleus is phonetically interpreted as a schwa *ə*, the neutral vowel in English. This type of vowel-zero alternation is reported for many other languages such as Moroccan Arabic (Kaye, 1990), French (Charette, 1991) and Japanese (Nasukawa, 1998, 2005).

[15] In principle, Element Theory allows any element to appear in any prosodic position (Nasukawa and Backley, 2008). A similar idea is also found in Clements and Hume (1995) which is based on Distinctive Feature Theory.

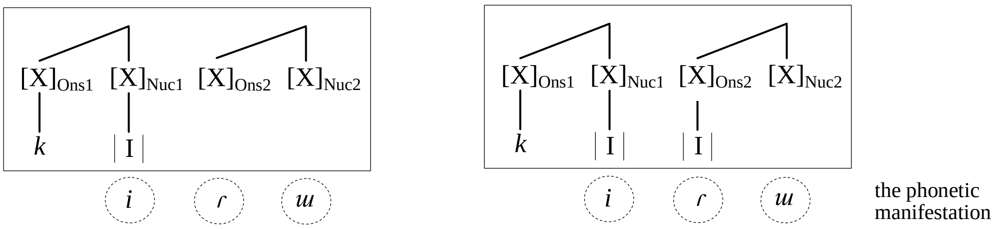
(22) a. *mi.∅∅* [miɾɯ] 'watch, see' b. *ne.∅∅* [neɾɯ] 'sleep'



Here it should be noted that no processes such as spreading or assimilation of a phonological prime are involved. As widely discussed in the literature on the analyses of harmony (Nasukawa, 2005; cf. Cohn, 1993), to conform to the empty onset condition in (21), the melodically empty nucleus does not receive |I| itself from the preceding onset, but rather, phonetically interpolates coronality (the coronal tap itself) from the phonetic realisation of |I| in the preceding nucleus. That is why the tap of Type A stems (like the second onset in (17a)) is susceptible to being targeted by phonological processes such as suppression or assimilation.

In order to justify the structures proposed for Japanese verb stems, consider the negative formation illustrated in (16). In both (16a) and (16b), we find forms which exhibit phonetically identical verb stems: *kiɾɯ* 'wear' in (16a) and *kiɾɯ* 'cut' in (16b). Although these are phonetically identical, they are phonologically different: as depicted below, *kiɾɯ* 'wear', a Type A stem ends with a sequence of an empty onset and an empty nucleus while in Type B stems only the final nucleus is melodically empty, as in *kiɾɯ* 'cut'.

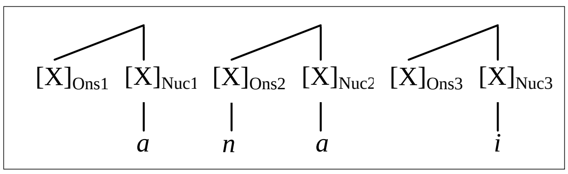
(23) a. Type A: *ki.∅∅* [kiɾɯ] 'wear' b. Type B: *ki.ɾ∅* [kiɾɯ] 'cut'



The crucial difference is obviously the second onset: the position is melodically empty in (23a) whereas the corresponding position in (23b) is filled with the coronal/dip element |I|. As discussed earlier, *ɾ* in (23b) is the phonetic realisation of the melodic prime |I| whereas *ɾ* in (23a) is the phonetic manifestation the empty onset (Ons_1) which interprets the |I| from the preceding nucleus (Nuc_1).

When it comes to negative formation, this representational difference impacts on the phonetic realisation of Ons_2 in (23a) and Ons_2 in (23b): *kiɾɯ* 'wear' plus the negative suffix *-anai* becomes *kinai* 'someone does not wear' while *kiɾɯ* 'cut' plus *-anai* appears as *kiɾanai* 'someone does not cut'. Let us now take a closer look at these differences by first considering the representation of the negative suffix *-anai*, which is illustrated in (24).

(24) The negative suffix: *-∅a.na.∅i* [anai]



This structure contains two empty onsets, both of which can be phonetically silent since they are followed by (dependent on) a filled nucleus, conforming to the condition in (21).

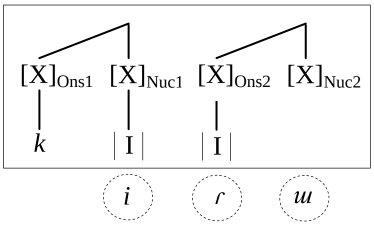
The following section employs the structures in (23a), (23b) and (24) in its demonstration of how Japanese verb inflection proceeds in the present approach.

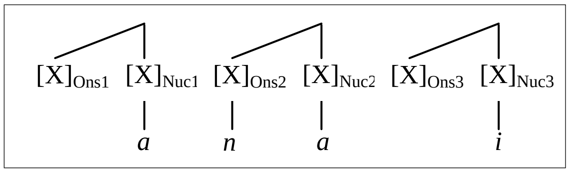
4.2. Interpret empty positions

As I mentioned in section 3, this analysis involves an operation which prevents empty categories from appearing at the boundary between the verb stem and the negative suffix. This morphological operation is proposed as an alternative to the more conventional concatenation of stem plus suffix.

First, let us consider the concatenation involving a Type B verb stem *ki.ɾɯ* ‘cut’ and the negative suffix -*a.na.i*. Both morphemes have a strict CVCV structure, as shown in (25):

(25) a. *ki.ɾ∅* [kiɾɯ] ‘cut’ b. -*∅a.na.∅i* -[anai]



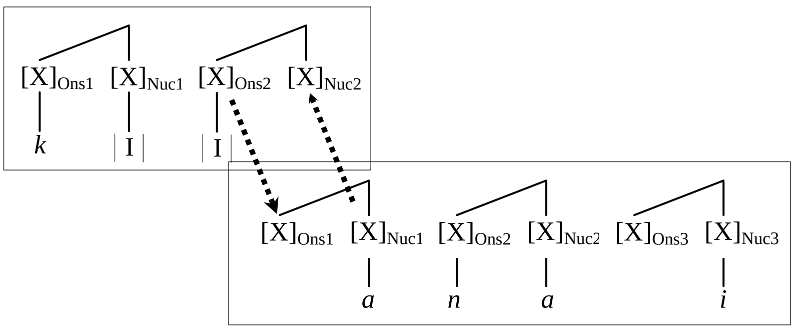


Employing the structures in (25), I claim that an inflected verb form results from a morphological concatenation process in which melodically empty positions must be interpreted with phonological prime(s) at the morphological boundary. The operation is assumed to be triggered by the following requirement on morphological concatenation.

(26) INTERPRET EMPTY POSITIONS
At a morphological boundary, interpret an empty position using melodic material from another prosodic position of the same type.

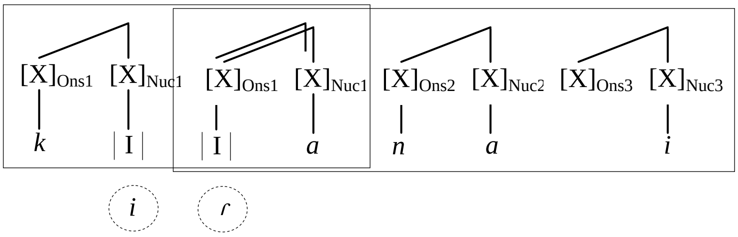
In order to conform to (26), the empty positions at the morphological boundary receive melody from another prosodic position of the same type. In the case of (25), Nuc_2 of the stem and Ons_1 of the suffix which are both melodically empty must phonetically manifest themselves with phonological material: the empty nucleus Nuc_2 of the stem is phonetically interpreted with the melodic content of the suffix Nuc_1; on the other hand, the empty onset Ons_2 of the stem manifests itself with the melodic content of the suffix Ons_1. The resulting shape of the negative form seems to involve an overlapping of stem and suffix at the boundary, as shown below.

(27) *ki.ɾ∅* + -*∅a.na.∅i*



Here, no position is left empty at the boundary and lexically-given prosodic dependency relations in both the suffix and the stem remain unchanged.

(28) *ki.ɾa.na.∅i* [kiɾanai]

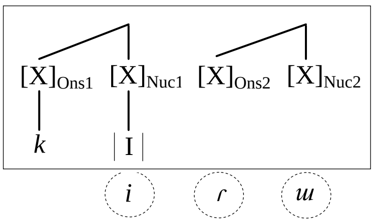


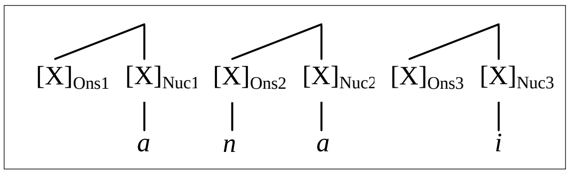
Unlike the conventional concatenation process in (11a), the overlapping process leaves all lexically-given prosodic structure intact: no deletion or addition of phonological objects is involved.[16]

This analysis also offers an explanation for why the stem-final weak vowel *ɯ* is replaced by the suffix-initial vowel *a*: this supposedly marked process is accounted for straightforwardly by identifying the melodic structure of *ɯ* as empty; then the position is easily filled by the following suffix-initial vowel *a*.[17]

We now turn to the concatenation involving a Type A verb stem and the negative suffix. In principle, this can be analysed in a parallel fashion, but with one difference concerning the melodic representation of the stem-final nucleus: Type B verb stems all end with an empty nucleus which is phonetically interpreted as *ɯ* while Type A verb stems end with a sequence of an empty onset and an empty nucleus. In Type A negative formation three empty positions are present at the morphological boundary: a sequence of an empty onset and its following empty nucleus (referred to as a minimal prosodic domain) occupy the final positions of the stem; and the suffix-initial onset is still empty, as illustrated below.

(29) a. *ki.∅∅* [kiɾɯ] 'wear' b. -*∅a.na.∅i* -[anai]





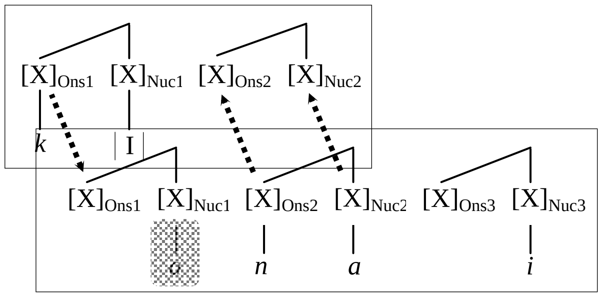
In accordance with the requirement in (26), the three empty positions at the boundary must be phonetically filled with melodic content. A possible concatenation is (30) where Ons_2 and Nuc_2 the empty positions of the stem are phonetically interpreted with the melodic content of Ons_2 and Nuc_2 of the suffix respectively, and the suffix-initial empty onset Ons_1 is superimposed by the filled first onset of the stem.

---

[16] An anonymous reviewer suggests that the claim of this paper (inalterability of lexically-established prosodic structure) can also be upheld without non-canonical morphological operations by using a standard OT approach to deletion and epenthesis; this would involve subjecting the phenomena in question to the ranking ONSET, REAL(IZE) MORPH(EME), IDENT, MAX $\text{V}_{\text{R(OO)T}}$ >> DEP >> MAX $\text{V}_{\text{AF(FIX)}}$ >> MAX $\text{ɯ}_{\text{RT/AF}}$, where REAL MORPH requires that every morpheme has at least one phonological exponent (cf. Oostendorp, 2005 and references therein). The validity of this type of analysis depends not only on what kind of lexical representations are assumed in a given system, but also on what kind of model is assumed to describe the phonological module.

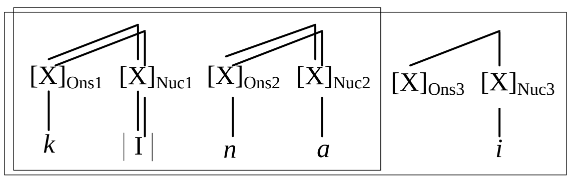
[17] A similar overlapping operation takes place in one type of affixation in Lungu (autonym *Cilungu*, spoken in parts of Zambia and Tanzania). The process involving the Class 9/10 preprefix *i*- is captured by assuming that the preprefix is affixed to the Class 9/10 prefix *n*- via a superimposing process rather than by linear concatenation (for a detailed discussion, refer to Nasukawa (to appear)).

(30) *ki*.∅∅ + -∅*a*.*na*.∅*i* → *ki*.*na*.∅*i* [kinai]



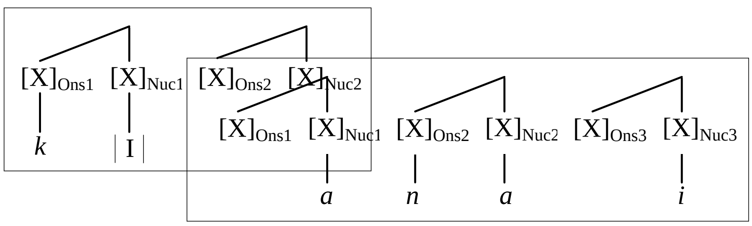
Unlike in (27), in this configuration the overlapping nuclei (Nuc_1 of the stem and Nuc_1 of the suffix) are both melodically filled with melodic material. In this case, as discussed in section 3, the *i* of the stem and the *a* of the suffix do not fuse. Rather, dependency relations holding between the two morphemes determine the output form. Following Nishiyama (this volume), there is a dependency relation between the stem and the suffix in morphological concatenation: it is the stem that is the head of composites of a concatenated form while the suffix is the dependent of the stem. This relation reflects on the manifestation of melodic material. In the case of (30), Nuc_1 in the stem (i.e. the morphological head) retains its lexical melody *i* whereas Nuc_1 in the suffix (i.e. the dependent) is not phonetically interpreted. Thus the resulting string [kinai] is generated.

(31) *ki*.*na*.∅*i* [kinai]



In contrast, if the final minimal domain (Ons_2 and Nuc_2) of the stem overlaps with the first prosodic domain (Ons_1 and Nuc_1) of the suffix, as depicted below, we arrive at an ill-formed concatenated structure which allows the second onset at the boundary to remain empty. This ill-formed structure would be phonetically interpreted as unattested *[kianai], which violates the requirement in (26).

(32) *ki*.∅∅ + -∅*a*.*na*.∅*i* → **ki*.∅*a*.*na*.∅*i* *[kianai]



The same type of operation under the requirement INTERPRET EMPTY POSITIONS can be used to explain other types of verb inflection such as causative and conditional formations. In the former, the lexical form of the conditional suffix is assumed to be *-sase*. When it concatenates with Type A stems such as *kiɾɯ* 'wear', which end phonologically with an empty onset followed by an empty nucleus, as shown in (33a), the empty positions overlap with *sa* in the initial filled onset-nucleus domain of the suffix. This generates the grammatical string [kisase].

(33) Causative formation: *-sase* (the causative suffix)

a.	Type A:	*kiɾɯ* 'wear' + *-sase*		[kisase]
		ki.∅∅ + *-sa.se*	→	*ki.sa.se*
b.	Type B:	*kiɾɯ* 'cut' + *-sase*		[kiɾase]
		ki.ɾ∅ + *-sa.se*	→	*ki.ɾa.se*

On the other hand, as illustrated in (25a), the Type B stem *kiɾɯ* 'cut' contains only a final empty nucleus preceded by an onset that is melodically filled with a single coronal/dip element |I| (which phonetically manifests itself as the tap *ɾ*) . In order to conform to the requirement in (26), it is assumed that *ɾ∅* of the stem *ki.ɾ∅* overlaps with *sa* of the suffix *-sa.se* and, like the case in (30), *ɾ* in the stem (the morphological head) is retained whereas *s* in the suffix (the dependent) is not phonetically interpreted. This generates the grammatical string [kiɾase].

When it comes to the conditional formation, there is no phonetic distinction in the form of the concatenated structure between *kiɾɯ* 'wear' + *-ɾeba* and *kiɾɯ* 'cut' + *-ɾeba*, both of which manifest themselves as [kiɾeba]. The formation involving Type A *kiɾɯ* 'wear' is straightforward: ∅∅ of *ki.∅∅* is reinterpreted as *ɾe* of *-ɾe.ba*; then the resulting form becomes [kiɾeba].

(34) Conditional formation: *-reba* (the conditional suffix)

a.	Type A:	*kiɾɯ* 'wear' + *-ɾeba*		[kiɾeba]
		ki.∅∅ + *-ɾe.ba*	→	*ki.ɾe.ba*
b.	Type B:	*kiɾɯ* 'cut' + *-ɾeba*		[kiɾeba]
		ki.ɾ∅ + *-ɾe.ba*	→	*ki.ɾe.ba*

In the formation involving Type B *kiɾɯ* 'cut', on the other hand, *ɾ∅* of *ki.ɾ∅* and *ɾe* of *-ɾe.ba* must overlap. In this case, *ɾ* in the morphological head (the stem) is retained whereas *ɾ* in the dependent (the suffix) is not phonetically interpreted. Although the operation is different from the Type A case, the phonetic shape of the concatenated string is [kiɾeba], identical to the string for Type A.[18]

5. Why prosodic information is necessary in lexical representations

The cases of concatenation shown above require dependency relations of prosodic (syllabic) structure to be fully specified. By contrast, most rule-based and/or multi-stratal approaches do not specify syllable structure in lexical representations: instead, syllabification takes place through the serial application of extrinsically ordered rules during derivation (Bromberger and Halle, 1989). This procedure relies on the linear order of segments being shown in lexical forms.

Although the authors do not discuss the point explicitly, McCarthy and Prince (1986) assumes that some organizational units such as morae are lexically given, while other units such as syllabic constituents are assigned during derivation. Yet in this scenario too the assignment of syllable structure relies on precedence information being stated (e.g. the linear order of feature geometry Root nodes). This type of representation is generally seen as having the advantage of excluding representational redundancy in derivations. Here, a lexical representation consisting of a string of segments and moraic structure constitutes the minimal amount of information necessary for building syllable structure. The reasoning behind this approach to lexical representations in generative phonology seems to have been the assumption that constraints on long-term memory prompt speakers to minimise storage (by keeping only idiosyncratic information in the lexicon) and maximise the computation of predictable information. Interestingly, this view has never been seriously defended in the psycholinguistics literature On the contrary, all the psycholinguistic evidence in fact points to the specification of (some) prosodic information in the lexicon.

Within mono-stratal approaches to phonology, it is impossible to dispense with information

[18] A similar operation takes place in past tense formation in Japanese where the past tense suffix is assumed to be *-ita* rather than *-ta*. Although the formation involving Type A stems is straightforward, that for Type B stems is rather complex since phenomena such as voicing and gemination are involved. For a detailed discussion, refer to Nasukawa (2005).

regarding precedence relations between positions/nodes. For example, the mono-stratal model OT (McCarthy and Prince, 1993; et passim) employs the same type of lexical representation as McCarthy and Prince (1986): without specifying syllable structure, a lexical/input representation consists of a string of segments and the partial specification of morae. Evaluation via universal constraint interaction selects an optimal syllabification from candidates based on the precedence relations of positional strings.

However, arguments supporting the minimal specification of syllable information in the lexicon are circular, because it turns out that a converse approach may also be theoretically possible: i.e. syllable structure may be assigned first, then on the basis of this syllable structure we can determine the distribution of morae and segments. In this regard, Takahashi (2004) claims that syllable structure should be fully specified in lexical representations even in the framework of OT. This is because an empirically impossible evaluation may be deemed optimal if syllable structure is not specified at the input level in OT. For example, consider a constraint ranking which evaluates the candidate strings [ai] and [ja] from the input form /ai/. In order to get [ai] from the input /ai/, we may assume that faithfulness constraints such as MAX-IO are ranked higher than markedness constraints such as ONSET and COMPLEX. On the other hand, to get [ja] from the input /ai/, the constraint LINEARITY which forbids metathesis must crucially be ranked lower than ONSET and COMPLEX. However, we have no reports of phenomena which violate LINEARITY but at the same time satisfy ONSET. That is, not all possible constraint rankings are allowed.[19] Further arguments in favour of the lexical specification of syllable structure even in OT is found in Golston (1996). This avoids the general problems which inhere in the OT framework.

In order to avoid this kind of problem, I assume that syllable structure should also be lexically specified in OT. Then faithfulness constraints such as Structure Preservation would not allow for the possibility of generating [ja] from the input form /ai/.

Following this argument, consider one version of the Government/licensing-based framework GP (Harris, 2004; Takahashi, 2004; Nasukawa, 2005) which takes a mono-stratal approach. Unlike OT and the standard multi-stratal approach, it permits the specification of syllable structure in lexical representations. In this framework, the only syllable properties deemed necessary are those required for expressing lexical contrasts.

6. Conclusion

In discussing verb inflection in Japanese, this paper has claimed that prosodic (syllabic) structures are given at the lexical level rather than during the course of derivation. Prosodic structure is assumed to be specified not only for verb stems but also for suffixes — in other words, for all morphemes relevant to the inflection process in question.

In a departure from conventional analysis, I have argued that all Japanese verb stems have representations ending in a nucleus: this nucleus is melodically empty in stems ending phonetically with a consonant. This approach succeeds in excluding stems which end in a coda consonant, the coda pattern having been marked as untypical among stems (and, more generally, among free morphemes) in the language. In addition, this paper has claimed that verb stems which have been conventionally considered to end with a vowel should actually end with a minimal prosodic domain consisting of an empty onset followed by an empty nucleus.

Using the proposed representations, I have argued that verb inflection involves an overlapping process rather than a conventional process of parallel concatenation. This operation is triggered by the requirement INTERPRET EMPTY POSITIONS. The phonetic outcome of overlapping domains is determined by dependency relations holding between morphological constituents: the melodic material in the stem (head of the domain) must be realised whereas that in the suffix (dependent status) may be phonetically suppressed. On the other hand, material in the stem can be phonetically manifested if the overlapped position in the stem is melodically empty. I believe that a similar mechanism can also apply to past-tense verb formation in Japanese (for a description of past-tense verb inflection involving nasality, see

---

[19] Another example is given in Kaye (1995:320–321) as follows:

> Consider French *watt* ‘watt’ and *oiseau* ‘bird’. Their initial portion is pronounced identically, [wa]. If claim (a) is applied to French then their initial portion ought to have the same syllable structure. It does not, cf. *le watt* vs. *l’oiseau* and *les watts* vs. *les oiseaux*. Consider also Italian pairs such as *fato* ‘fate’ vs. *fatto* ‘fact’. Both contain the sounds ‘f’, ‘a’, ‘t’, and ‘o’. The first syllable of *fato* is open, while that of *fatto* is closed. Such examples could be easily multiplied.

Nasukawa, 2005:Ch4; for more general data, see Tsujimura, 1996:40–54). In addition, the formation of Japanese adjectives may be analysed in a similar manner.

Further research will be needed to investigate whether the same mechanism is employed in other languages of the world.

Acknowledgements

The author would like to thank Jeroen van de Weijer, Phillip Backley, Kunio Nishiyama and two anonymous reviewers for their insightful and helpful input and feedback, most of which I have been able to incorporate in the present paper. This research is partially supported by the Ministry of Education, Culture, Sports, Science and Technology of the Japanese government under grant 19520429.

References

Abe, Y., 1987. Metrical structure and compounds in Japanese. In: Imai, T., Saito M. (Eds.), Issues in Japanese linguistics, Foris, Dordrecht, pp. 5–52.

Archangeli, D., 1984. Underspecification in Yawelmani phonology and morphology. PhD dissertation, Massachusetts Institute of Technology. Published 1988, Garland, New York.

Ashworth, D., 1976–77. Historical perspective on voicing and the gerund in Japanese. Papers in Japanese Linguistics 5, 27–39.

Backley, P., Nasukawa, K. 2009. Headship as melodic strength. In: Nasukawa, K., Backley, P, (Eds.), Strength Relations in Phonology. Mouton de Gruyter, Berlin/New York, pp. 47–77.

Backley, P., Nasukawa, K. 2010. Consonant-vowel unity in Element Theory. Phonological Studies 13.

Backley, P., Takahashi, T., 1998. Element activation. In: Cyran, E. (Ed.), Structure and Interpretation ― Studies in Phonology. Folium, Lublin, pp. 13–40.

Bickmore, L.S., 2007. Cilungu Phonology. Center for the Study of Language and Information, Stanford, CA.

Bloch, B., 1946. Studies in colloquial Japanese III: derivation of inflected words. Journal of the American Oriental Society 66, 304–315.

Bromberger, S., Halle, M., 1989. Why phonology is different. Linguistic Inquiry 20, 51–70.

Charette, M., 1991. Conditions on Phonological Government. Cambridge University Press, Cambridge.

Charette, M., 1998. Empty and pseudo-empty categories. SOAS Working Papers in Linguistics 8, 167–176.

Clements, G.N., Hume, E., 1995. The internal organization of speech sounds. In: Goldsmith, J.A. (Ed.), The Handbook of Phonology. Blackwell, Oxford, pp. 245-306.

Clements, G.N., Keyser, S.J., 1983. CV Phonology: a Generative Theory of the Syllable. MIT Press, Cambridge, MA.

Cohn, A.C., 1993. Nasalisation in English: phonology or phonetics. Phonology 10, 43–81.

Davis, S., Tsujimura, N., 1991. An autosegmental account of Japanese verbal conjugation. Journal of Japanese Linguistics 13, 117–144.

Golston, C,. 1996. Direct Optimality Theory: representation as pure markedness. Language 72, 713–748.

Gussmann, E., Kaye, J.D., 1993. Polish notes from a Dubrovnik café. SOAS Working Papers in Linguistics and Phonetics 3, 427–462.

Halle, M., 1997. Some consequences of the representation of words in memory. Lingua 100, 91–100

Haraguchi, S. 1991. A theory of Stress and Accent. Foris, Dordrecht.

Harris, J., 1994. English Sound Structure. Blackwell, Oxford.

Harris, J., 1997. Licensing Inheritance: an integrated theory of neutralisation. Phonology 14, 315–370.

Harris, J., 2004. Release the captive coda: the foot as a domain of phonetic interpretation. In: Local, J., et al. (Eds.), Phonetic Interpretation: Papers in Laboratory Phonology 6. Cambridge University Press, Cambridge, pp. 103–129.

Harris, J., 2005. Vowel reduction as information loss. In: Carr, P., Durand, J., Ewen, C.J. (Eds.), Headhood, Elements, Specification and Contrastivity. Benjamins, Amsterdam, pp. 119–132.

Harris, J., Lindsey, G., 1995. The elements of phonological representation. In: Durand, J., Katamba, F. (Eds.), Frontiers of Phonology: Atoms, Structures, Derivations. Longman, Harlow, Essex, pp.

34–79.
Harris, J., Gussmann, E., 1998. Final codas: why the west was wrong. In: Cyran, E. (Ed.), Structure and Interpretation ― Studies in Phonology. Folium, Lublin, pp. 139–162.
Harris, J., Gussmann, E., 2002. Word-final onsets. UCL Working Papers in Linguistics 14, 1–42. [ROA, #575-0203, http://ruccs.rutgers.edu/roa.html]
Hasegawa, N., 1979. Casual speech vs. fast speech. Papers from the 15th Regional Meeting of Chicago Linguistic Society. pp. 126–137.
Herbert, R.K., 1975. Reanalyzing prenasalized consonants. Studies in African Linguistics 6, 105–123.
Herbert, R.K., 1986. Language Universals, Markedness Theory, and Natural Phonetic Processes. Mouton de Gruyter, Berlin.
Ito, J., 1986. Syllable theory in Prosodic Phonology. PhD dissertation, University of Massachussetts. Published 1988, Garland, New York.
Ito, J., 1990. Prosodic minimality in Japanese. Chicago Linguistic Society 26-II. pp. 213–239.
Ito, J., Mester, R-A., 1996. Stem and word in Sino-Japanese. In: Otake, T., Cutler, A. (Eds.), Phonological Structure and Language Processing: Cross Linguistic Studies. Mouton de Gruyter, Berlin/New York, pp. 13–44.
Kaye, J.D., 1990. 'Coda' licensing. Phonology 7, 301–330.
Kaye, J.D., 1995. Derivations and interfaces. In: Durand, J., Katamba, F. (Eds.), Frontiers of Phonology: Atoms, Structures, Derivations. Longman, Harlow, Essex, pp. 289–332.
Kaye, J.D., Lowenstamm, J., Vergnaud, J-R., 1990. Constituent structure and government in phonology. Phonology 7, 193–232.
Kubozono, H., 1985. Speech errors and syllable structure. Linguistics and Philology 6, 220–243.
Kubozono, H., 1999. Mora and syllable. In: Tsujimura, N. (Ed.), The Handbook of Japanese Linguistics. pp. 31–61.
Kula, N.C., 2002. The phonology of verbal derivation in Bemba. PhD dissertation, Leiden University.
Kula, N.C., Marten, L., 2009. Defining initial strength in clusterless languages in Strict CV. In: Nasukawa, K., Backley, P, (Eds.), Strength Relations in Phonology. Mouton de Gruyter, Berlin/New York, pp. 251–284.
Maddieson, I., 1990. Prenasalized stops and speech timing. Journal of the International Phonetic Association 19.2, 57-66.
Maddieson, I., Ladefoged, P., 1993. Phonetics of partially nasal consonants. In: Huffman, M.K., Krakow, R.A. (Eds.), Nasals, Nasalization and the Velum (Phonetics and Phonology 5), Academic Press, San Diego, CA, pp. 251-301.
Maeda, S., 1979. Boundaries in Japanese Phonology, Gengo Kenkyu 75, 54–70.
McCarthy, J.J., Prince, A.S., 1986. Prosodic Morphology. Ms, University of Massachusetts, Amherst and Brandies University.
McCarthy, J.J., Prince, A,S., 1993. Prosodic Morphology I: constraint interaction and satisfaction. Report no. RuCCS-TR-3. Rutgers University Centre for Cognitive Science, New Brunswick, NJ.
McCawley, J.D., 1968. The Phonological Component of a Grammar of Japanese. Mouton, Hague.
Nasukawa, K., 1998. Prosodic integrity and melodic representation in nasals. Interdisciplinary Information Sciences 4.1, Tohoku University, 1–8.
Nasukawa, K., 2004. Word-final consonants: arguments against a coda analysis. Proceedings of the Tohoku English Literary Society. pp. 47–53.
Nasukawa, K., 2005. A Unified Approach to Nasality and Voicing. Mouton de Gruyter, Berlin/New York.
Nasukawa, K., 2007. Relational properties in phonology: precedence and dependency. Paper presented at the CUNY Conference on Precedence Relations.
Nasukawa, K., to appear. Prosodic affiliation of NC sequences in Lungu (Cilungu). In: Legère, K. (Ed.), Proceedings of the International Conference on Bantu Languages: Analysis, Description and Theory. Rüdiger Köppe Verlag, Köln.
Nasukawa, K., Backley, P., 2008. Affrication as a performance device. Phonological Studies 11, 35–46.
Nishiyama, K., 1996. Historical change of Japanese verbs and its implications for Optimality Theory. In: Koizumi, M., Oishi, M., Sauerland, U. (Eds.), Formal Approaches to Japanese Linguistics 2: MIT Working Papers in Linguistics 29, pp. 155-171.
Nishiyama, K., Penultimate Accent in Japanese Predicates and the Verb-Noun Distinction, this volume.

Oostendorp, M. v., 2005. Expressing inflection tonality. Catalan Journal of Linguistics 4, 107–126.
Poser, W.J., 1986. Japanese evidence bearing on the compensatory lengthening controversy. In: Wetzels, L., Sezer, E. (Eds.), Studies in compensatory Lengthening. Foris, Dordrecht, pp. 167–186.
Prince, A.S., Smolensky, P., 1993. Optimality Theory: constraint interaction in generative grammar. Technical report #2 of the Rutgers Center for Cognitive Science, Rutgers University.
Prince, A.S., Smolensky, P., 2004. Optimality Theory: Constraint Interaction in Generative Grammar. Blackwell, Oxford.
Scobbie, J., Coleman, J., Bird, S., 1996. Key aspects of declarative phonology. In: Durand, J., Laks, B. (Eds.), Current Trends in Phonology: Models and Methods, Volume 2. European Studies Research Institute (ESRI), University of Salford, Manchester, pp. 685–709.
Shibatani, M., 1990. The Languages of Japan. Cambridge University Press, Cambridge.
Tabata, T., 1983. Tagen koozo onin riron to nihongo (A multi-tiered theory of phonology in Japanese), Gengo Kenkyu 84, 77–104.
Takahashi, T., 2004. Syllable theory without syllables. PhD dissertation, University College London.
Tateishi, K., 1990. Phonology of Sino-Japanese morphemes. University of Massachusetts Occasional Papers in Linguistics 13, 209–235.
Tsujimura, N., 1996. An Introduction to Japanese Linguistics. Blackwell, Oxford.
Vance, T.J., 1987. An Introduction to Japanese Phonology. SUNY Press, New York.
Vance, T.J., 2008. The Sounds of Japanese. Cambridge University Press, Cambridge.
Yoshida, S., 1990. The 'mora' in Japanese. Phonology 7, 331–351.
Yoshida, S., 2003. The syllabic nasal in Japanese. In: Ploch, S. (Ed.), Living on the Edge: 28 Papers in Honour of Johnathan Kaye. Mouton de Gruyter, Berlin/New York, pp. 527–542.