A preliminary study of postlexical syllable structure in Japanese: a view from vowel devoicing*

Manami Hirayama
University of Toronto

Prosodic structure has been the subject of a great deal of discussion in phonology, but prosodic structure at the postlexical level has perhaps received less attention than at the lexical level. This paper sheds light on postlexical aspects of prosodic structure and investigates the question of what is permissible structure there; specifically, whether the prosodic structure in the postlexical domain differs from that found in the lexical domain. The process of interest is High Vowel Devoicing/Deletion in Tokyo Japanese. Kondo (e.g., 1997) analyzes this process as a change in syllable structure involving desyllabification. I test the hypothesis that desyllabification is present in this process. Preliminary results based on production and perception data suggest that, contrary to Kondo’s proposal, desyllabification is not present. Based on this, I conclude that the syllable structure is the same in the postlexical domain as in the lexical domain. The data presented in this paper is from impressionistic observations. At this point, experimental evidence is needed for both production and perception, which is in progress.

1. Introduction

Prosodic structure has been the subject of a great deal of discussion in phonology, but the prosodic structure at the postlexical level has perhaps received relatively less attention than at the lexical level. This paper sheds light on postlexical aspects of prosodic structure and investigates the question of what is permissible structure there, specifically the question of whether the prosodic structure in the postlexical domain differs from that found in the lexical domain. The language of interest is Japanese (Tokyo

* I would like to thank Keren Rice for her comments. Earlier versions of this paper were discussed in the Workshop on Japanese Phonology and Morphology held at the University of Toronto on August 16, 2005 and Montreal-Ottawa-Toronto Phonology Workshop held on February 10–12, 2006 at the University of Toronto and York University. Many thanks go to the audience of these workshops for suggestions and comments.
The syllable is well-defended in the literature on Japanese. For example, Kubozono (1999) argues for the importance of syllable as well as mora for Japanese phonology as both are needed in describing the accent system of this language. In the lexicon and throughout the lexical phonology, the well-formed syllable shape is \((C)(j)V(N)(Q)\), where \(N\) denotes a moraic nasal and \(Q\) the first half of geminate. Onsets and codas are optional. Onsets maximally consist of two consonants with the second segment being the glide /\(j\)/ (e.g., /\(kj\)oo/ [\(k\'\(j\)\(o\)] ‘today’). Codas can optionally have a moraic nasal (N) and/or the first half of geminate (e.g., /\(ga\)'N/1 ‘wild goose’, /kappa/ ‘water imp’, /roN\(do\)N/ ‘London’ + /-ko/ ‘local people, lit. child’ \(\rightarrow\) roN\(do\)Nkk\(o\) ‘Londoner’2). I summarize the syllable shape in (1).

(1) Permissible syllable in the lexical domain of Japanese

\[(C)(j)V(N)(Q)\]

\(N\): moraic nasal \(Q\): first half of geminate

Note: vowel length contrast is omitted.

However, when we look at surface forms, we sometimes find somewhat different forms from what we expect from the syllable shape in (1). One such instance is apparent (voiceless) obstruent clusters created by a postlexical process called High Vowel Devoicing/Deletion (hereafter HVD), whereby the high vowels \(i, u\) become voiceless between voiceless obstruents or between a preceding voiceless obstruent and a following pause. Phonetic realizations are as shown in (2). The transcription in (2) is given in two ways—one treating the vowel as devoiced (to the left) and the other as deleted (to the right). Some analyze HVD as vowel devoicing, which also is the traditional analysis for HVD in Japanese (e.g., Beckman 1996:101), and some as vowel deletion (e.g., Kondo 1997, 2000, 2005). I owe the acoustic description of HVD to Kondo (2005:238). If the onset consonant is a fricative, the sequence of it followed by a devoiced vowel is realized as continuation of a preceding fricative (2a). If the onset consonant is a stop, the preceding stop is released into a fricative (2b, c, d).

(2) Phonetic realization of HVD (devoiceable vowels are italicized and underlined)

a. /\(s\)u\(p\)o'\(o\)u\(t\)/ ‘sports’ \(\rightarrow\) [\(s\)\(t\)]oo\(ts\)u \(\sim\) [\(s\)\(p\)\(o\)\(o\)\(ts\)u]

b. /k\(t\)a/ ‘north’ \(\rightarrow\) [k\(t\)]a \(\sim\) [k\(\#\)t]a

c. /s\(j\)o\(k\)'\(k\)a\(N\)/ ‘cabinet secretary’ \(\rightarrow\) sjo[k\(k\)\(k\)]aN \(\sim\) sjo[k\(k\)k]\(a\)N

d. /s\(j\)o\(k\)ut\(k\)u\(t\)/ ‘dining table’ \(\rightarrow\) sjo[k\(\#\)t]\(a\)ku \(\sim\) sjo[k\(\#\)t]\(a\)ku

Among others, Kondo (1997, 2000, 2005) analyzes HVD as altering syllable structure. Specifically, she analyzes HVD as involving deletion of the vowel, with desyllabification followed by resyllabification of the onset consonant to an adjacent syllable.3 I turn to the details of her analysis in next section. This proposal calls for a

---

1 Accent is marked with an apostrophe (‘’) after the vowel. More about accent bearer follows.
2 The sequence of moraic nasal followed by a geminate in the coda (i.e., NQ in the coda) occurs only in derived contexts—within a morpheme, the coda is occupied (if it is) by either a moraic nasal or the first half of geminate.
3 Kondo (2005) proposes a structural change at the mora level by demoraification as well as a change at the
radical structural change in syllable structure in the postlexical domain, since the proposed onsets and codas of the type shown in the right-hand-side transcription in (2) are not permitted in the lexicon or in the lexical phonology of the language—they yield new contrasts for onset and coda inventories for the postlexical domain of Japanese. A question thus arises: Is this structural change necessary? How should the apparent consonant clusters be analyzed? I investigate this question from the perspectives of production and perception of speakers and from the perspective of learnability.

The paper is organized as follows. Section 2 gives an overview of Kondo’s (1997, 2000, 2005) analysis with critical comments on it, which leads us to a hypothesis to test her claim. Section 3 describes an experimental design to test the hypothesis and reports some results from my own speech and several sound files from a Japanese accent dictionary (NHK Hoosoo Bunka Kenkyujo 2002). In section 4, I propose a postlexical structure as the results suggest. In section 5, I investigate the question from a learnability perspective. Section 6 concludes the paper with a brief summary.


In this section, an overview of Kondo (2005) is given with critical comments. As introduced in the previous section, Kondo analyzes HVD as altering syllable structure. She proposes that HVD involves vowel deletion and desyllabification, followed by demoraification and resyllabification to a neighbouring syllable. This is illustrated in (3) with an example word .akikan ‘empty can’, in which the middle vowel i is the target of HVD. (3a) is the structure before HVD applies. (3b) shows that the vowel, being the target of HVD, is deleted, the preceding consonant becomes moraic, and the syllable is lost (i.e., desyllabification). This is followed by demoraification of the consonant and resyllabification of the consonant to the preceding syllable (3c).

(3) Kondo’s analysis—resyllabification to the preceding syllable as default case
/akikan/ ‘empty can’ (Kondo 2005, (2))

Kondo also identifies special cases where resyllabification occurs in the other direction, i.e., to the following syllable. Although the directionality issue does not concern us (partly because it is not entirely clear in Kondo’s analysis why resyllabifying to the preceding syllable should be the default case), it is worth looking at an example, since one of these special cases offers a testable case of her proposal, that HVD leads to a structural change at the syllable level. Kondo says that resyllabification is to the following syllable in two cases. Though only one is the crucial for my argument, let us syllable level. I leave the topic of demoraification for future research and focus on the syllable level here.

91
review both cases for the sake of completeness. The first case is when the syllable with a devoiced vowel occurs word-initially. In this case, since there is no preceding syllable available for resyllabification, the following syllable is chosen as the target for resyllabification, as in (4), with the word *kita* ‘north’. The first syllable desyllabifies (4b) (and demoraifies) and the remaining consonant associates to the second syllable, *ta*, creating an onset cluster [kçt] (4c).

(4)  If word initial syllable contains a devoiceable vowel, then resyllabification is to the following syllable.

/kita/ ‘north’ (Kondo 2005, (3), slightly modified)

Another case, and the case of interest, in which Kondo claims resyllabification occurs to the following syllable, instead of to the preceding one, is when the syllable that contains a devoiceable vowel is accented. An example derivation is given in (5) with the word *sjoki*’kaN ‘cabinet secretary’. For the reason to be studied in detail below, in this case the desyllabified segment [kç] (in (5b)) is resyllabified to the following syllable, *kaN* (5c), due to the fact that the syllable that contains the vowel that underwent HVD, i.e., middle *i*, is lexically accented.

(5)  If accented syllable contains a devoiceable vowel, then resyllabification is to the following syllable

/sjoki’kaN/ ‘cabinet secretary’ (Kondo 2005, (4), slightly modified)

The evidence that Kondo uses to justify desyllabification and resyllabification to the following syllable in (5) comes from phonetic, or acoustic, manifestation of the accented syllable when the vowel is devoiced. It has been claimed (e.g., Sugito and Hirose 1988) that when the vowel of accented syllable is devoiced, the accent is phonetically manifested in the following syllable. The accented syllable normally phonetically maps into a pitch fall, i.e., a pitch curve from high to low, and the fall starts with the accented syllable. In other words, pitch, as measured by F0, is high on the accented syllable and decreases thereafter. However, if the vowel in the accented syllable is devoiced, the pitch appears first with no F0 on the accented syllable (since there is no
vocal fold vibration due to HVD), which is followed by an “unusually high starting F0 of the following vowel that then falls very sharply (Sugito and Hirose 1988)” (Kondo 2005: 241). Sugito (1970, 1996:102) describes this pitch curve as appearing as if the F0 for the accented syllable continues on to the following syllable, resulting in the rapid pitch fall with high starting in the following syllable. Kitahara (1998) calls this pitch pattern on the following syllable Pitch Elevation Effect. Citing Sugito (1982), he shows the effect in comparison with the case without the effect (i.e., no vowel devoicing) as in (6). Focusing only on the F0 patterns, the left-hand-side figure shows that if the vowel in accented syllable does not undergo HVD, that syllable has high pitch, followed by a pitch fall. In comparison, the right-hand-side figure shows that when the vowel in the accented syllable is devoiced, F0 begins with the following syllable as high as the first syllable in the left-hand-side figure, followed by a rapid fall.4

Based on this acoustic characteristic, Kondo (1997, 2000, 2005) argues for desyllabification of voiceless vowels. Since the accent is phonetically realized on the following syllable, the original accent bearer must be deleted. Thus, her structure matches the phonetic manifestation: “[t]he acoustic cue of the lexical accent is manifested in that syllable [=the following syllable, or the CCVC syllable in (5c)]” (Kondo 2005: 241).

Is Kondo’s proposal of desyllabification (and resyllabification) well-motivated? Her proposal makes an interesting prediction that I examine in the remainder of this paper. Assuming that the syllable is the accent bearer in Japanese (e.g., McCawley 1968, Shibatani 1990), loss of a syllable followed by resyllabification with HVD as Kondo proposes predicts neutralization with a form with the accent on an adjacent syllable, thus neutralization between a minimal pair of items that have accent on the syllable that contains the devoiced vowel and items that have accent on the following syllable. For example, we should find neutralization between /hu’kiN/ ‘nearby’ (accent on the first syllable, which contains the devoiceable vowel u) (7a) and /huki’N/ ‘cleaning cloth’ (accent on the second syllable, i.e., the following syllable) (7b) when both undergo HVD.4

---

4 Although the figure in (6) is for the Osaka dialect, which has a slightly different accent system from the Tokyo dialect, Sugito (1970/1996) observes the same pattern in Tokyo speakers.
(7) *hu’kiN* (a) and *huki’N* (b) would be neutralized if desyllabification and resyllabification are present

a. /hu’kiN/ ‘nearby’

b. /huki’N/ ‘cleaning cloth’

We can test the hypothesis that desyllabification (and resyllabification) is present by examining whether there is neutralization after HVD in a minimal pair such as the above.

(8) Hypothesis to test: Desyllabification (and resyllabification) is present

I turn to this in the next section.

3. Testing the hypothesis—from perspectives of production and perception of speakers

As I explained in the previous section, Kondo’s (1997, 2000, 2005) proposal that HVD creates a change in syllable structure at the postlexical level predicts neutralization between a pair of words minimally different from each other with respect to accentuation patterns—items with accent on a syllable that contains a devoiceable vowel and items with accent on a following syllable. Some sample words are given in (9).

(9) Some sample words

<table>
<thead>
<tr>
<th>Accent on</th>
<th>syllable with devoiceable vowel</th>
<th>following syllable</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>hu’kiN</em></td>
<td>nearby</td>
<td><em>huki’N</em></td>
</tr>
<tr>
<td><em>ki’shi-o</em></td>
<td>knight-ACC</td>
<td><em>kishi’-o</em></td>
</tr>
<tr>
<td><em>ki’shi-o</em></td>
<td>knight-ACC</td>
<td><em>kishi’-o</em></td>
</tr>
</tbody>
</table>

In examining neutralization in production, it is important to understand how the accent is phonetically realized in Japanese. An accented syllable in Tokyo Japanese maps High tone to the first moraic segment in that syllable and Low tone to the following moraic segment, whether it is in the same syllable (i.e., the accented syllable is heavy) or the following syllable (i.e., the accented syllable is short). Further, I assume that, following Pierrehumbert and Beckman (1988), the tonal realization of accent correlates with F0. Thus we should see a pitch fall, or decrease of F0 values over time, associated with the accented syllable.

In next section, I present preliminary results, based on introspection of my own speech and an examination of several sound files in an accent dictionary called the *NHK Nihongo Hatsuon Akusento Jiten* (NHK Pronunciation and Accent Dictionary of Japanese

94
(NHK Hoosoo Bunka Kenkyuujo 2002)). An experiment with more subjects for both production and perception is under way.

3.1 Results

First, in my speech, a comparison of words such as those in (9) indicates that devoicing does not carry with it neutralization of the tone—the accent is displaced from the devoiced vowel but does not ‘shift’ to the following syllable. Thus, no neutralization is found there. Assuming that accent is carried by the syllable (e.g., McCawley 1968, Shibatani 1990; Vance 1987 for review; Kubozono 2005 p.c.), this suggests that the syllable unit is not lost, i.e., there is no desyllabification.

Second, I examined sound files from the NHK accent dictionary. I hear a difference between the two items in a pair. Pitch curves of pairs such as those in (9) show a clear difference between the words of the pair. Spectrograms with transcription, wave forms and F0 track are given in (10) to (13) below. Spectrograms, wave forms and F0 track were made with software for speech synthesis called WaveSurfer (Sjölander and Beskow 2005). Transcriptions are mine. If we look at the lowest figures, which show F0 movement, we can see that the patterns in each pair are clearly distinct from each other, with parallel patterns found across the two pairs. In each pair, items that have a devoiced vowel in the accented syllable (i.e., (10) hu’kiN and (12) ki’shi-o) have an F0 curve that falls right from the beginning of the vowel of the second syllable, while items with accent on the syllable following the syllable with the devoiced vowel (i.e., (11) huuki’N and (13) ki’shi’-o) exhibit a steady high pitch for some time in the vowel of the second mora (i in ki in (11) and i in shi in (13)) that then falls. The steady part is shown by a circle in the figures.
(10) Spectrogram of ʰɻ̃’kiʼN (sound file from NHK Nihongo Hatsuon Akusento Jiten)

(11) Spectrogram of ʰỹkiʼN (sound file from NHK Nihongo Hatsuon Akusento Jiten)
(12) Spectrogram of *ki*’o (sound file from *NHK Nihongo Hatsuon Akusento Jiten*)

(13) Spectrogram of *kishi*-o (sound file from *NHK Nihongo Hatsuon Akusento Jiten*)
Thus, in these production data from the NHK accent dictionary, there is no neutralization found. This suggests that there is no desyllabification (thus no resyllabification also). With this result, I propose a possible analysis for the apparent consonant clusters created by HVD in the next section.

4. A proposed structure for the apparent consonant clusters—no structural change

Having established that there is no neutralization in the pairs in (9), hence that the syllable unit is not lost and that the resyllabification is not present, I propose an analysis for the apparent consonant clusters created by HVD. In a nutshell, there is no structural change at the syllable level. I illustrate this in (14). In (14a) *hu’kiN* ‘nearby’, the accent, denoted by an asterisk ‘*’, is carried by the first syllable, $\sigma_1$, in the lexical domain. After HVD, this syllable $\sigma_1$ is not desyllabified but remains the same as in the lexical domain, i.e., with $\sigma_1$.

(14) An analysis of HVD—no structural change at the syllable level

a. *hu’kiN* ‘nearby’

\[
\begin{array}{c}
\sigma_1
\
\mu_1
\
C_1 V_1
\mu_2 \mu_3
C_2 V_2 N
h u k i N
\end{array}
\rightarrow
\begin{array}{c}
\sigma_1
* \sigma_2
\mu_1 H
\mu_2 \mu_3 L
\phi : C_1 V_1 C_2 V_2 N
\end{array}
\]

b. *huki’N* ‘cleaning cloth’

\[
\begin{array}{c}
\sigma_1
\
\mu_1
\
C_1 V_1
\mu_2 \mu_3
C_2 V_2 N
h u k i N
\end{array}
\rightarrow
\begin{array}{c}
\sigma_1
* \sigma_2
\mu_1 H
\mu_2 \mu_3 L
\phi : C_1 V_1 C_2 V_2 N
\end{array}
\]

I give a tentative analysis for tonal association. I assume that tonal association occurs before the application of HVD. A High tone is assigned to the first moraic segment in the accented syllable; this is $\mu_1$ in (14a). Low tones are assigned to all moras that follow $\mu_1$ (until the end of the accentuation domain). Thus $\mu_2$ and $\mu_3$ in (14a) get Low tone. When HVD applies, the High tone which is connected to the first mora, $\mu_1$, cannot bear the tone since the voicing, which is needed to realize a tone, is not available due to vowel devoicing. This is illustrated by the cross in (14a). The High tone is then associated with a nearby available mora, which is $\mu_2$ in this case. Since there is also a Low tone associated with $\mu_2$, this creates a contour tone HL on $\mu_2$. This would explain the Pitch
A PRELIMINARY STUDY ON POSTLEXICAL SYLLABLE STRUCTURE OF JAPANESE

Elevation Effect (section 2).\(^5\)

The tonal association for the other word in the pair, huki’N ‘cleaning cloth’, can be as follows (14b). I only talk about the tones related to the accent. The accent is carried by the second syllable \(\sigma_2\) in the lexical domain, and this maps High tone to the first moraic segment in the same syllable, here \(\mu_2\), and a Low tone to the following moraic segment, \(\mu_3\). HVD does not affect these tones. Thus, in terms of tonal melodies, the two words in (14) are differentiated with respect to the tones in \(\mu_2\): (14a) has HL, while (14b) has H.

Slightly tangential to the main issue, the effect of HVD on syllable structure, a word has to be said here. I assumed in proposing a contour tone HL in \(\mu_2\) in (14a) that the Pitch Elevation Effect exists (see section 2). However, a careful reading of literature about Japanese phonetics (e.g., articles in Sugito 1996) finds a very similar optional phenomenon for the realization of accent in general, and that is called ososagari, literally meaning ‘delayed fall’.\(^6\) Ososagari is a phenomenon where the pitch fall starts not with the first mora of the syllable that carries accent but with the following mora or even later. This is not a compulsory feature that is found in all accented syllables, as Sugito (1969, 1996:81–85) shows. A cursory examination of the words that are pronounced with ososagari, though not all test words are given in the article, finds that they have voiceless consonant in the onset of the syllable following the accented syllable. However, more careful work with the full inventory of consonants is needed, according to which the analysis of contour tone given above may have to be revised.\(^7\) The issue of tonal realization and its analysis is left for future research.

5. Testing the hypothesis—from the perspective of learnability

In this section, I investigate the question of whether a structural change of syllable in the postlexical domain is necessary from the learnability perspective. Given the postlexical structure as input, the learner’s task is to construct the grammar of his/her language. If the structural change occurs in the postlexical domain, new contrasts are introduced for onsets and codas in this domain, ‘new’ in that they do not exist in the lexical domain. Note that the syllables created as Kondo proposes in (3) to (5) are not permitted in the lexicon or in the lexical phonology of Japanese. For example, in (3c), the coda consonant is neither moraic nasal nor first half of geminate. In other words, a new contrast is introduced for the coda inventory—coda can be occupied by any voiceless obstruents whether or not it is first half of geminate, in addition to the moraic nasal. A new contrast is introduced for onsets also: the inventory for onset consonants in the postlexical domain includes consonant clusters with two segments whose second member is a voiceless obstruent (e.g., [k\(\overset{\mathrm{C}}{\text{t}}\)] in (4), [k\(\overset{\mathrm{C}}{\text{k}}\)] in (5)), while /j/ is the only permitted second segment in a complex onset cluster in the lexical domain. I summarize the new contrasts introduced in the postlexical domain as expected by Kondo’s analysis in (15).

\(^5\) Kitahara (1998) independently proposed a contour tone HL for the phonetic extra high pitch.
\(^6\) Thanks to Yukiko Sugiyama for drawing my attention to this phenomenon.
\(^7\) Another point around the contour tone, which is questioned by Peter Avery, is that this is a new structure only for the postlexical domain; there is no words with contour tone, not at least in the Tokyo dialect (Cf. in other dialects, they may exist). I have no explanation about this at the moment.
Permitted segments in onsets and coda inventories—New contrasts in postlexical domain (as expected by Kondo’s proposal)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Coda:</td>
<td>N, Q</td>
<td>N, Q, C[-voice, -sonorant]</td>
<td></td>
</tr>
</tbody>
</table>

If we adopt a model in which the surface forms correspond to the mental representations unless there is evidence to the contrary, a structural change in syllabification in the postlexical domain is not expected; a discrepancy between the structures in the lexical and postlexical domains such as a new contrast for onsets and codas would require a complicated learning process. Instead, the apparent cluster in a form such as [φk] in /hukiN/ serves as a cue to syllabification, signalling to the learner that a syllable with a devoiced vowel is present.

6. Conclusion

To the question that I opened this paper with, i.e., whether or not structural change is needed in the postlexical domain of Japanese, I conclude in the negative, from the perspectives of both production and perception and learnability. Thus, syllable structure is the same in this domain and in the lexical domain. The production/perception evidence presented in this paper to draw this conclusion is from impressionistic observations. At this point, experimental evidence is needed for both production and perception, and this work is in progress.

References


