Syllable structure and its acoustic effects on vowels in devoicing environments

Mariko Kondo

1. INTRODUCTION

Vowel devoicing is a common phonological process in many languages and typically involves high vowels and schwa. High vowels and schwa are inherently short (Bell, 1978; Dauer 1980) and the process usually occurs when the vowels are either adjacent to, or surrounded by, voiceless consonants, during which the glottis is fully open. It is thought that vowel devoicing is a consequence of articulatory undershoot of the glottal movements. It also suggests that vowel devoicing processes are the results of the glottal gestural overlap between voiceless consonants and short vowels. The movements of glottal muscles for the short high vowels /i/ and /u/ blend with those of the adjacent voiceless sounds or a pause (Jun, 1993; Jun and Beckman, 1994). In many languages, the process is also considered to be part of the vowel neutralization and reduction processes in which vowels are first reduced in duration and centralized in quality, typically in the unaccented position, and then eventually devoiced and/or deleted in fast or casual speech (Hyman 1975, Wheeler 1979, Dauer 1980, Kohler 1990).

The Japanese high vowels /i/ and /u/ also become voiceless when surrounded by voiceless consonants, or when preceded by a voiceless consonant and followed by a pause: i.e. /C̥VC̥/ or /C̥V#/ (where the Vs are [+high]). However, in Japanese, the vowel devoicing processes do not involve apparent centralization of vowels. There is no obvious vowel reduction in the unstressed positions of vowels in Japanese, nor does vowel quality depend on accentuation. However, the vowel devoicing process is very common in many Japanese dialects, especially in eastern dialects including voicing Standard Japanese. The process occurs even in slow or formal speech (Kondo, 1995). This suggests that Japanese high vowel devoicing is not merely an optional process in fast or casual speech, but a phonologically controlled process.

Previous phonetic and phonological studies have suggested that Japanese vowel devoicing can be affected by various phonetic and phonological factors, such as type of preceding and following consonants (Kuwabara and Takeda, 1988; Yoshida and Sagisaka, 1990), presence of accent on the vowel (Takeda and Kuwabara, 1987; Hattori 1989), position in a word or utterance (Maekawa, 1989; Takeda and Kuwabara, 1987) and following word boundary (Sakurai, 1985). However, Kondo (1997) found that Japanese vowel devoicing was almost obligatory even when the vowel was accented and followed by an internal word boundary, as long as there were no devoiceable vowels in adjacent syllables (the single devoicing environment), e.g. *ashita* /as*i̲*ˈta/ 'tomorrow', *kikai* /ˈk*i̲*kai/ 'machine' and *kusa* /k*u̲*ˈsa/ 'grass' where all the underlined vowels in italics can be devoiced (here and henceforth, the devoiceable vowels are written underlined in italic). However, when vowels in adjacent syllables were all devoiceable (the consecutive devoicing environment), such as *kashitsuchishi* /kas*i*t*u*ˈt*i*s*i*/ 'accidental death' and *fukushikikokyuu* /f*u*k*u*s*i*k*i*ˈkokjuu/ ‘abdominal breathing’, some vowels remained voiced. All studies agree that in the consecutive devoicing environment only some devoiceable vowels undergo the devoicing process. Therefore the devoicing factors suggested above were not always effective.

The effects of most of the suggested devoicing factors were also minimal in the consecutive environment when speaking tempo was altered. If vowel devoicing is merely a consequence of articulatory undershoot or glottal gestural overlap between short high vowels and voiceless consonants, the devoicing should occur more when the speech rate increases. In Japanese, devoicing of high vowels seemed to be almost compulsory in the single devoicing environment at all tempi, even at a slow tempo (Kondo, 1997). As long as there were no devoiceable vowels in neighboring syllables, vowel devoicing seemed to be an almost compulsory process and even speaking tempo had very little effect. On the other hand, the devoicing rates in the consecutive devoicing environment varied according to the tempi. When speaking tempo was altered, the devoicing rates of the high vowels in prose text reading were not necessarily high at a comfortable speaking tempo for all types of preceding consonants. However, when the data in the consecutive devoicing environment were excluded from the statistics, the devoicing rates at all tempi significantly rose, and vowel devoicing seemed to be almost compulsory (for details, see Kondo, 1997).

The devoicing rates showed that the most important factor affecting devoicing was whether there was a devoiceable vowel in adjacent syllables: i.e. the single or consecutive environments. Almost all high vowels in the single devoicing environments became voiceless, whereas in the consecutive environment high vowels sometimes remained voiced. It was also found that vowels in the devoicing environments were often acoustically different from the same vowels in non-devoicing environments. Acoustic analyses of high vowels in the devoicing environment revealed that vowel devoicing was not always a clear-cut distinction either voiced or voiceless, but that there were also many partially voiced/devoiced vowels. In fact, phonetic realizations of vowels in the devoicing environment varied from fully voiced to completely voiceless. Despite being phonetically in the same condition, the high vowels showed different acoustic characteristics. This meant that vowel devoicing did not simply mean the presence or absence of vocal fold vibration, but involved fundamental acoustic changes of the vowels in the processes.

The devoicing processes seemed to trigger acoustic changes in the vowel process, and the degree of change was dependent on its phonological environment. Devoicing occurred mainly in the single devoicing environment. It also sometimes occurred in the presence of two consecutive syllables, but never in three at a normal speaking tempo. Devoicing conditions in single and consecutive devoicing environments are phonetically identical. This implies that phonetic conditions are not the only conditions that affect devoicing. In this paper, I will first analyze devoicing processes in terms of syllable structure. I will consider how Japanese vowel devoicing changes syllable structures, why not all vowels in devoicing environments become voiceless and also why devoicing in consecutive morae, especially in more than three consecutive morae does not occur. Secondly, the reality of devoiced vowels will be discussed in terms of the duration of devoiced vowels. Finally, the acoustic changes of vowels in the devoicing environment will be examined with respect to devoicing processes. From this, I will consider whether Japanese vowel devoicing processes are part of the vowel reduction processes that produce devoicing in other languages.

2. SYLLABLE CONSTRAINTS ON VOWEL DEVOICING

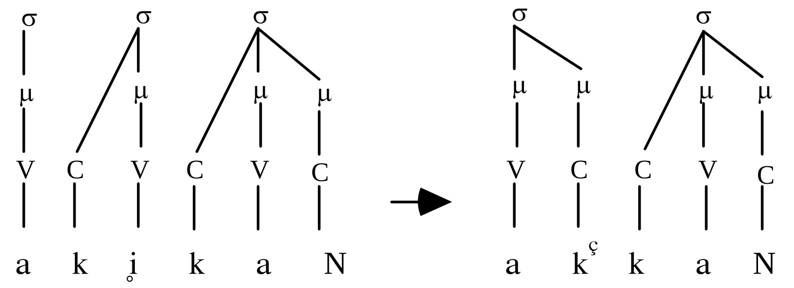
Vowel devoicing is primarily a lack of vocal fold vibration during the production of the high vowels /i, u/ between voiceless sounds. It is not very economical for the vocal folds to vibrate during the vowel production as the glottis is open for the preceding and following sounds. However, the devoicing process is actually influenced by various other phonetic and phonological factors. Phonetically ideal environments do not always trigger devoicing. For example, an experiment found that devoicing occurred in two consecutive morae in some samples, but not in three or more consecutive morae (Kondo, 1997). Devoicing factors did not always trigger devoicing, and speaking tempo did not significantly affect devoicing rates in single devoicing sites. These results indicated that phonetic analyses cannot explain all devoicing processes.

High vowels are almost always devoiced in the single devoicing environment, whereas devoicing is an optional process in the consecutive devoicing environment. The fundamental difference between the two is whether the preceding consonant can be resyllabified to an adjacent syllable after a vowel becomes voiceless, i.e. the change of syllable structure. When a vowel becomes voiceless, it is acoustically manifested either as the continuation of a preceding fricative or as a preceding stop released into a fricative, thus creating sequences of voiceless consonants. Japanese syllables are predominantly light open syllables /(C)V/. Consonant clusters do not occur within a syllable except for the rare occurrence of /-NC/ sequences in a syllable final position as in /hoNtte/ 'Books are...' and /abadi:Nkko/ 'Aberdonian', and /Cj-/ sequence in a syllable initial position if we consider a glide /j/ as a consonant, as in /tja/ 'tea' and /gjoo/ 'line'. When the voiceless vowel loses its sonority, the preceding consonant in the devoiced syllable cannot constitute a syllable on its own. The syllable structure of a word is altered as a result of vowel devoicing. In addition to the mora, the syllable is important in Japanese as an accent baring unit (Shibatani, 1991), and constrains various phonological processes, such as formation of loan words (Itô, 1990; Shinohara, 1997). Vowel devoicing processes change the syllable structure by creating consonant sequences, but maintain the mora structure as the duration of a word or phrase is maintained.

The phonological structure of the word *akikan* /a.ki.kaN/ [aki̥kaN] 'empty can' with devoiced [i̥] is represented as (2a)[1]. When the high vowel /i/ becomes voiceless, the devoiced mora /$k^{ç}$/ cannot be attatched to the syllable node because the syllable has lost its core element. Therefore, the second syllable cannot sustain its status as a syllable. As the remaining preceding consonant /k/ cannot constitute a syllable on its own, it has to be syllabified with an adjacent syllable as (2b). However, it becomes moraic since the mora keeps the duration of the devoiced vowel and the quality of the devoiced vowel /i/ is reflected in the palatalisation of the preceding consonant as /$k^{ç}$/. The process creates a bimoraic heavy syllable /VC/ (which is permitted in Japanese), and reduces the number of syllables, but maintains the moraic structure of the word.

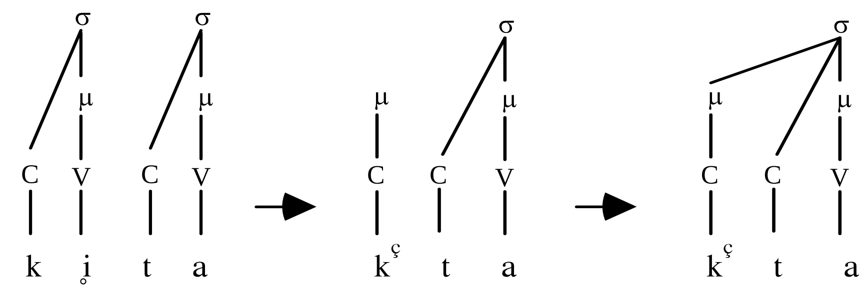
("σ" denotes a syllable, "μ" denotes a mora.)

(2a) *akikan* /akikaN/ [aki̥kaN][2] (2b)



When devoicing occurs at the beginning of a word, such as *kita* /ki.ta/ [ki̥ta], 'north' and *hikari* /hikari/ [çi̥ka'ɾi] 'light', the consonant in the devoiced syllable is syllabified with the following syllable as the syllable onset because this is the only possible place it can move to. As shown in (3), when the vowel /i/ becomes voicelss in the word *kita* /ki.'ta/ (3a), the preceding consonant /s/ becomes moraic (3b) and is syllabified with the following syllable /ta/ (3c).

(3a) *kita* /kita/ [kita] (3b) [ki̥ta] (3c) [$k^{ç}$ta]



---

[1] The mora tier usually represents an alternative rather than an addition to the CV tier, and onset consonants are attatched directly to the syllable node as they are nonmoraic (Hayes, 1989; Kenstowicz, 1994). However, I use a separate CV tier in order to present clearly the formation of moraic consonant and resulting syllable structures.

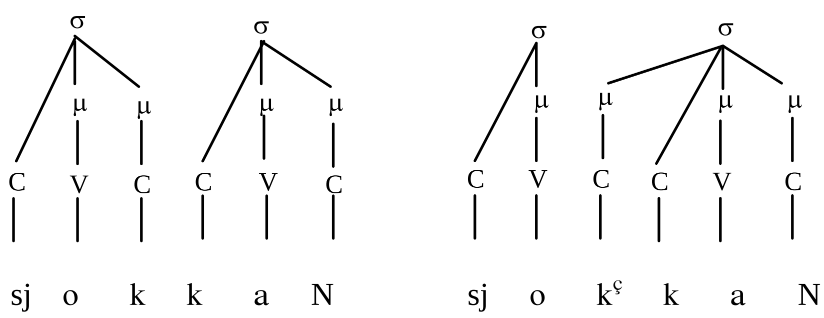
[2] Pseudo-phonemic transcriptions are used to describe devoiced vowels and resulting moraic consonants for convenience. The examples are the devoiced vowels /i̥, u̥/, the allophones of consonants /ɕ/ instead of /s/ and /sju/ (in /si/ and /sju/), /ç/ and /ɸ/ instead of /h/ (in /hi/ and /hu/), /tɕ/ instead of /t/ (in /ti/) and /ts/ instead of /t/ (in /tu/). /$k^{ç}$/ was also used to indicate palatalisation of /k/ and its release into a palatal fricative [$k^{ç}$] in /ki/, and [k^{x}] for /k/ in /ku/ to indicate backness of the /k/ and its release into a velar fricative [x].

Application of this process must be restricted to certain positions because the resulting syllable structure is awkward as consonant clusters do not occur in syllable initial positions in Japanese, except for the /Cj-/ sequence. In these processes, the number of syllables is reduced.

Devoicing occurs even when a devoiceable vowel is in an accented syllable. Devoicing of an accented vowel can be blocked or avoided by shifting the accent to another syllable (McCawley, 1977; Sakurai, 1985; Vance, 1987). However, vowel devoicing in accented syllables can also occur in normal speech (Hattori, 1989). In Japanese, a syllable carries the lexical accent and not a mora. When the vowel in an accented syllable is devoiced, and its preceding consonant cannot sustain its syllabic status, it can no longer carry an accent. The preceding consonant becomes moraic, but the mora is not an accent baring unit. For instance, the underlined /i/ in the word *shokikan* /sjo.'ki̲.kaN/ 'cabinet secretary' is devoiceable. When it is devoiced and loses its syllabicity, it cannot simply be syllabified to the preceding syllable /sjo/ as (4a). This is because it was the syllable /ki/ that carried the accent in the word and the accent baring syllable has now been lost.

Therefore the preceding /k/ has to be syllabified to the following syllable /kaN/. This analysis seems appropriate since the accented syllable corresponds to the acoustic manifestation of an accent match. The acoustic cue for the lexical accent is the fall of the fundamental frequency (F0) from an accented vowel to the following syllable. Perceptual cue for accent on a devoiced vowel is the unusually high starting F0 of the following vowel that then falls very sharply (Sugito and Hirose, 1988). Resyllabification of the first /k/ to the following syllable /kaN/ as in (4b) creates the superheavy syllable /kˤkaN/ (/CCVC/). The acoustic cue of the lexical accent is manifested in that syllable.

(4a) */sjo'kikaN/ [çoki̥kaN]   (4b)



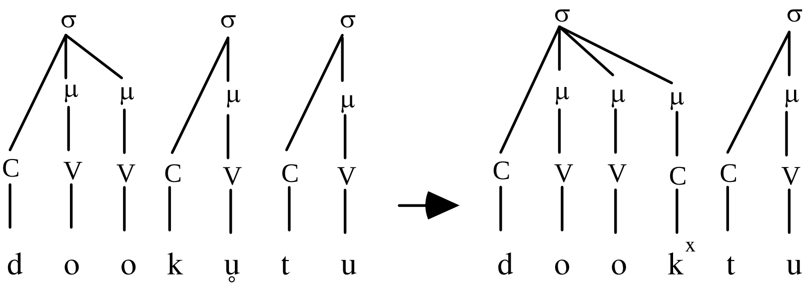
When there is only one devoiced vowel in a word, the preceding consonant in the same mora can be syllabified to its preceding syllable. In the case of word initial position or in an accented syllable, it can be syllabified to the following syllable. Therefore, vowel devoicing in the single devoicing environment is always possible[3]. However, in the consecutive devoicing environment, not all

---

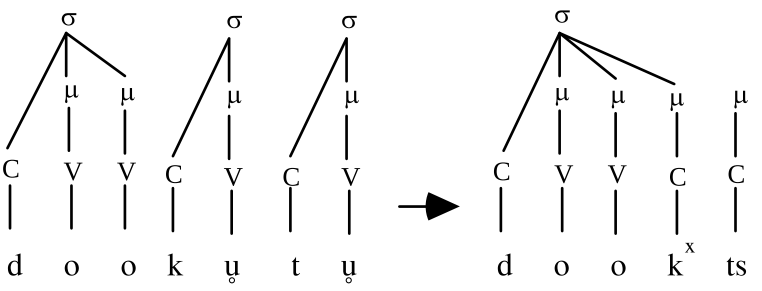
[3] For the arguments concerning vowel devoicing and the loss of syllabicity, see Kondo (2001).

devoiceable vowels become voiceless. For example, both underlined italic /u/ are devoiceable in the word *dookutsu* /dookutu/ (no accent) 'cave', and common pronunciations for the word are [do:ku̥tsɯ] with the first /u/ devoiced and [do:kɯtsɯ] with both /u/ voiced. In an earlier study [do:ku̥tsɯ] occurred in 16 out of 18 samples (Kondo, 1997). The first pronunciation is possible because /k/ in /ku/ can be syllabified to the preceding syllable as shown in (5a). However, two consecutive devoicings *[do:ku̥tsɯ̥] are not possible in normal speech. As shown in (5b), /k/ in /ku/ can be syllabified to the preceding syllable, but /t/ in /tu/ cannot because it would create sequences of */CVVCC/. This is not considered to be an acceptable superheavy syllable as the second last consonant is not a moraic nasal. The pronunciation of [do:kɯtsɯ] with both voiced vowels is more favorable than the clusters of three consonants.

(5a) [do:ku̥tsɯ]



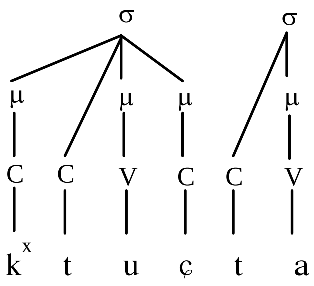
(5b) *[do:ku̥tsɯ̥]



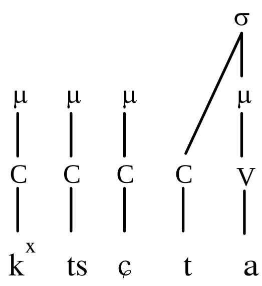
In the word *kutsushita* /ku'tusita/ 'sock(s)', where the underlined vowels are devoiceable, unaccented and preceding consonants were the plosive [k], the affricate [ts] and the fricative [ɕ]. The pronunciation [kɯ̥'tsɯɕi̥ta] with the first and third vowels devoiced and the second vowel voiced is most common. This process is also explained by the syllable structure, as shown in (6a) and (6b). When the first vowel /u/ in /ku/ becomes voiceless, the preceding /k^x/ becomes moraic and is syllabified to the beginning of the following syllable. The third vowel /i/ also becomes voiceless, and the preceding /s/ is syllabified to the end of the preceding syllable. The process creates a less common superheavy syllable /CCVC/, but it is still better than devoicing in three consecutive morae. Triple devoicing is not acceptable as shown in (6b). The first two consonants /k^x/ and /ts/ cannot be syllabified to the following syllable, because it would create triple-consonant clusters. Moreover, in the word /ku'tusita/ 'sock(s)', the syllable /tu/ carries the lexical accent. It is most logical and sensible to leave the vowel of /tu/ voiced rather than creating an awkward heavy syllable */tssi/ or

*/tsç̥ta/. Alternatively, it is also acceptable to pronounce this word with all vowels voiced [kɯtsɯçita], but never all three vowels devoiced.

(6a) [kɯ̥'tsɯç̥ita]



(6b) *[kɯ̥'tsɯ̥ç̥ita]



For analyses of other syllable structures such as in cases including devoicing before geminate consonants and the acceptability of bimoraic syllable onset, refer to Kondo (2001).

3. Acoustic characteristics of vowels in the devoicing environment

3.1 Durations of devoiced morae

The previous section examined the syllable structures constraining vowel devoicing processes. When phonological conditions are in favor of vowel devoicing and the high vowels become voiceless, are the acoustic changes of the vowels simply the change of phonation or do the processes involve other acoustic changes?

It is well known that Japanese speech rhythm is based on the mora: there is a tendency towards equalizing duration of morae and the duration of a whole word or a phrase is proportional to the number of morae in that word or phrase (Campbell and Sagisaka, 1991; Sato, 1993; Han, 1994, etc.). Port et al. (1987) showed that the duration of a whole word was proportional to the number of morae in the word, even if 2 vowels in a 5-mora word were voiceless. However, Beckman (1982) found the duration of the morae with devoiced vowels were significantly shorter than their voiced counterparts. If vowel devoicing simply means the change of vowel phonation from voiced to voiceless, the voiceless vowel should retain its duration. On the other hand, if devoicing is part of the vowel weakening process, then durational reduction may occur. The duration of the devoiced vowels may be shorter than voiced vowels. Therefore, an experiment was carried out in order to examine whether devoiced vowels retain their durations. The durations of devoiced morae were compared with the durations of their voiced counterparts in the same phonetic environment.

In the experiment, six subjects pronounced 41 test words containing 74 devoicing sites three times each in random order. Their individual pronunciation of devoiceable vowels in the same words was not always consistent. When there was variation in voicing of the same devoiceable vowel in the same word, the word was segmented and the segmental durations measured. The durations of consonants in devoiced morae were compared with those of corresponding CV morae. In order to minimize the effects of the various factors that control segmental duration such as (a) type of phonemes, (b) neighboring phonemes, (c) mora position in a breath group, and (d) speaking rate, durational comparisons were made only of the same mora in the same word uttered by the same speaker in utterance internal positions. For example, both /i/ vowels in /hoo'ti̱ki̱/ ‘fire alarm’ are in the consecutive devoicing environment. Voicing and devoicing patterns of the both /i/ can vary from time to time. For example, the same speaker may devoice the first /i/ in one utterance and voice it in another utterance. When voicing of the same vowel in the same position in the same word by the same speaker varied, the duration of the mora with a voiceless vowel was compared with the duration of the voiced counterpart with voiced vowel. The comparison was made only between the data from the same speaker.[4]

The data were collected from 738 recorded words (41 test words x 6 subjects x 3 pronunciations) containing 1332 devoiceable vowels (74 devoicing sites x 6 subjects x 3 pronunciations). Among them, 45 devoicing sites (45 sites x 3 times = 135 high vowels) had voicing variations, excluding word-final position and pre-pausal position. All words which had voicing variation were segmented using Waves+ speech analysis on a SUN workstation.

The durational measurement found that morae with voiceless vowels were significantly shorter in duration than those with voiced vowels [t(44)=8.49, p<.001] (Figure 1). The average ratio of devoiced morae against /CV/ counterparts was 83.93% (SD 12.97). However, the durations of devoiced morae were significantly longer than their corresponding consonant's portion of /CV/ morae [t(44)=13.62, p<.001].

The closure durations of plosives and the plosive part of affricates in devoiced morae were compared with the closure durations of prevocalic plosives and the plosive part of prevocalic affricates in /CV/ morae. It is technically impossible to measure the duration of devoiced vowels after voiceless fricatives, since there is no way to tell the boundary between the voiceless vowel and preceding voiceless consonants. Therefore, the durations of morae with devoiced fricatives were excluded from the data. The average closure duration of plosives and affricates in /CV/ morae was not significantly different from that in devoiced morae as shown in Figure 2. The results suggested that when a vowel was devoiced, the vowel became shorter than its fully voiced counterpart, and as a result, the whole duration (57.16 ms) [t(31)=0.51, n.s.] of the mora was reduced.[5] The average closure duration of plosives and affricates in CV morae (56.16 ms) was not significantly different from devoiced morae (57.16 ms) [t(31)=-0.51, n.s.] (two-tailed).

[4] Word final morae were excluded from the comparison, even if there was variation in voicing of the vowel. Some test words had possibilities of consecutive devoicing and devoicing sites occurring word initially, but there were no cases of voicing variations in word-initial position.

[5] It should be noted that despite durations of devoiced morae being reduced, the durations of the whole words with devoiced morae were still proportional to the number of morae in a word. In other words, there seemed to be a word level durational adjustment despite a number of devoiced vowels in a word. See Kondo (1995) for the details.

Figure 1 ***Average durational difference between devoiced morae and consonants and vowels in CV morae of all types of consonants***

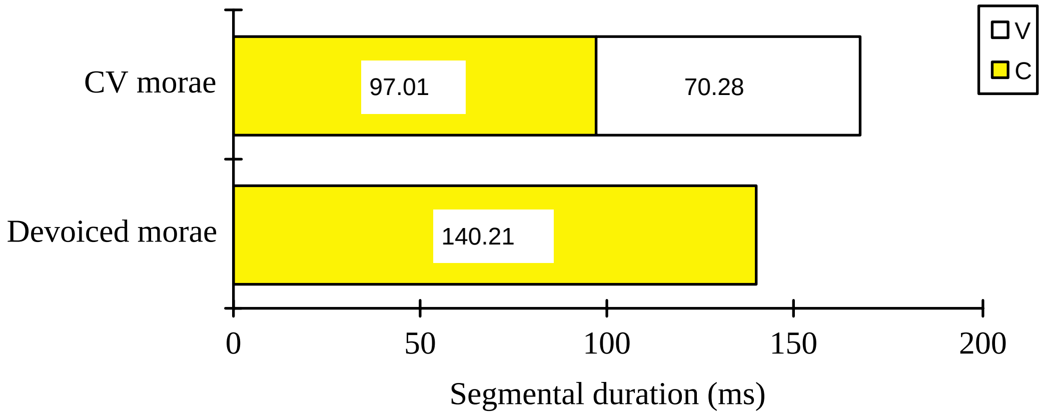
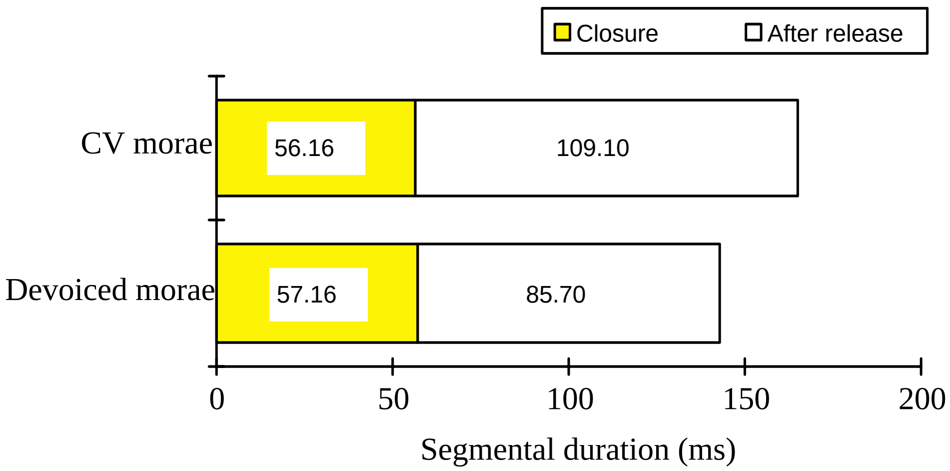


Figure 2 ***Average closure duration and the duration after release of plosives and plosive part of affricates in CV morae and devoiced morae***



On the other hand, the durations of the section after the release of the stop closure and fricative part of affricates in devoiced morae were very different from the total duration of this fricative part and the following vowel. The average duration of plosive and affricates in devoiced morae excluding closure duration (i.e. after release of stop closure) was 85.70 ms (SD = 24.36), whereas that of CV morae excluding closure duration of the consonants was 109.10 ms (SD = 25.16). The statistical analysis by T-test (one-tailed) found the duration of plosive and affricates after release in devoiced morae was significantly longer than that of /CV/ morae [$t(31)=7.12$, $p <.005$]. The results suggested that although the durations of whole morae differed significantly depending on whether the vowel was voiced or devoiced, the closure durations of plosives and affricates in devoiced morae, and those of non-moraic counterparts, did not show a significant difference. The results suggested that when vowels were devoiced, their durations were reduced[6].

---

[6] The presence of devoiced vowels did not affect the duration of whole words. Despite shorter duration of devoiced morae, the whole durations of words with up to three devoiced vowels did not show significant difference from words

3.2 INTENSITIES OF VOWELS IN THE DEVOICING ENVIRONMENTS

The previous section showed that there was durational reduction when vowels were devoiced. In addition, voiced devoiceable vowels in the devoicing environments showed differences in quality in terms of intensity and duration depending on whether they were in the single or consecutive devoicing sites. Devoicing rates showed that vowels in the single devoicing sites were almost always devoiced, whereas only some vowels in the consecutive devoicing sites became voiceless. If devoicing in the single site is a natural process, voiced high vowels in the single sites are unnatural and therefore they may be acoustically different from the same vowels in the non-devoicing environment. On the other hand, if only some high vowels undergo the devoicing process in the consecutive sites, then it must be natural for the voiced vowels in the consecutive devoicing sites to retain the same acoustic qualities as the same vowels in the non-devoicing environment. Moreover, if devoicing is part of vowel a weakening process, voiced vowels in the devoicing environment may show weakening of their intensities as well as durational reduction.

An experiment was carried out in order to measure intensities of vowels in the devoicing environments at three different speaking tempi and to compare them with the intensities of voiced vowels in the non-devoicing environment in the same word. Three subjects pronounced 6 test words listed in (7a) and (7b) with single and consecutive devoicing sites at slow, comfortable and fast tempi (devoiceable vowels are underlined in italic).

(7)	(a)	/ta.i.sjo.k*u*. 'te.a.te/	[taiɕokɯteate]	‘retiring allowance’
		/ka.mo.t*u*. 'se.N.pa.k*u*/	[kamotsɯsempakɯ]	‘cargo boats’
		/ta.ka.sa.k*i*. 'si.mi.N/	[takasakiɕimĩĩ]	‘the Takasaki citizens’
	(b)	/h*u*.k*u*.sjo.ku. 'ke.N.sa/	[ɸɯkɯɕokɯkensa]	‘dress inspection’
		/sjo.k*u*.h*i*.se.tu.ja.k*u*/	[ɕokɯçisetsɯjakɯ]	‘ a cut in food expenses’
		/ha.i.sj*u*.t*u*.ki.dju.N/	[haiɕɯtsɯkidʑɯɯ̃]	‘exhaust limit’

The total number of devoiceable vowels in single devoicing sites in the test words was 6 (as some of the test words contain more than 1 single site or both single and consecutive sites), and that in consecutive devoicing sties was also 6, yielding 324 devoiceable vowels ([6 vowels + 6 vowels] x 3 rates x 3 repetitions x 3 speakers = 324 devoiceable vowels).

The intensities were analyzed by subject and speaking rate, and were compared using a T-test. The average intensities of vowels in non-devoicing environments, excluding word-initial and word-final morae, were calculated by three tempi and by individual subjects. The results showed that the intensities of voiced devoiceable vowels were significantly lower than those of ordinary vowels in the non-devoicing environment at all tempi, except for one of the subjects at normal and slow tempi. The result was predicted because an earlier experiment found that phonetic realization of vowels in

---

of the same number of morae without devoiced vowels. See Kondo (1997) for details.

the devoicing environment varied from fully voiced to completely voiceless. When a vowel happened to be voiced in the devoicing environment, it was often partially voiced, i.e. the duration of the vowel tended to be shorter and its intensity was lower.

Since the effect of speaking tempo on devoicing rates was significant in the single devoicing environment but not significant in the consecutive devoicing environment, the intensities of vowels were compared by their devoicing environments using a T-test. When there was only one devoicing site in a word (the single devoicing condition), and the devoiceable vowels remained voiced, the intensities of the devoiceable vowels were significantly lower than those of non-devoiceable vowels at all speaking tempi for all subjects (Table 1). One of the subjects (A) devoiced all devoiceable vowels at the normal tempo and only once voiced the devoiceable vowel (/hukusjok*u*keNsa/) at the slow rate.

In the consecutive devoicing environments, however, when devoiceable vowels were voiced their intensities were not necessarily lower. At the fast tempo, the average intensity of devoiceable vowels was higher than those of non-devoiceable vowels for Subject C. The intensity of voiced devoiceable vowels and the intensity of non-devoiceable vowels were significantly different at all tempi for Subject A (Table 2). There was also a significant difference at the normal tempo for Subject B, but not at fast or slow tempi, and not at any tempo for Subject C. Since Subject A voiced very few devoiceable vowels, Subject A's data were excluded, and the results of Subjects B and C are presented in Figure 3.

Table 1 **T-test results of intensity differences between vowels in single devoicing sites and non-devoicing environments (one-tailed)**

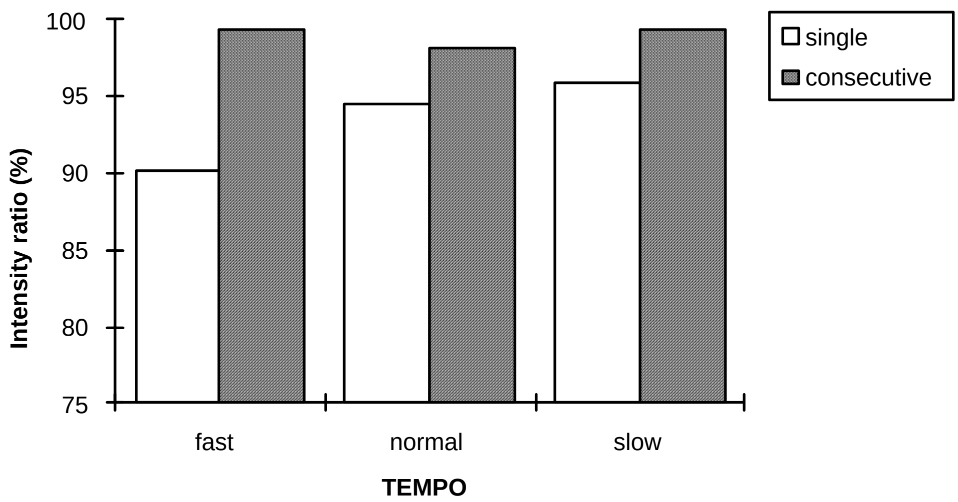
subject	tempo	average intensity of devoiceable vowels (dB)	average intensity of non-devoiceable vowels (dB)	df	p-value
A	fast	67.07	77.34	3	$p < 0.005$
	normal [7]	N/A	N/A	N/A	N/A
	slow [8]	(59.57)	(76.56)	N/A	N/A
B	fast	68.69	75.14	2	$p < 0.025$
	normal	72.63	77.64	2	$p < 0.05$
	slow	69.10	72.45	11	$p < 0.001$
C	fast	69.69	76.27	4	$p < 0.025$
	normal	71.10	74.85	5	$p < 0.05$
	slow	70.32	72.93	7	$p < 0.001$

[7] Subject A devoiced all devoiceable vowels.
[8] There was only one instance of a voiced vowel in the devoicing site.

Table 2 T-test results of intensity differences between vowels in consecutive devoicing environments and in non-devoicing environments (one-tailed)

subject	tempo	average intensity of devoiceable vowels (dB)	average intensity of non-devoiceable vowels (dB)	df	p-value
A	fast	70.20	78.25	5	$p < 0.025$
	normal	72.37	75.60	6	$p < 0.005$
	slow	70.15	76.12	9	$p < 0.005$
B	fast	71.57	73.83	3	n.s.
	normal	76.12	78.47	13	$p < 0.001$
	slow	72.03	72.63	11	n.s.
C	fast	74.71	74.63	10	n.s.
	normal	74.46	74.79	11	n.s.
	slow	71.76	72.26	12	n.s.

Figure 3 ***The intensity ratios between voiced devoiceable vowels and non-devoiceable vowels in single and consecutive devoicing environments at three tempi by subjects B and C***



Intensity of sound is very sensitive and is influenced by various factors, such as neighboring sounds, pitch, stress and accentuation. Under the same conditions, the same vowel has higher intensity in higher pitch than in lower pitch, and also higher intensity in a stressed position than in an unstressed position. Moreover, different vowels have their own intrinsic intensities even when spoken with equal effort. Vowels made with a wider vocal tract have a higher intensity level than the close vowels. For the same degree of opening, vowels with closer F1 and F2 have a higher intensity than vowels with F1 and F2 far apart; i.e. back vowels are a little more intense than front vowels. In

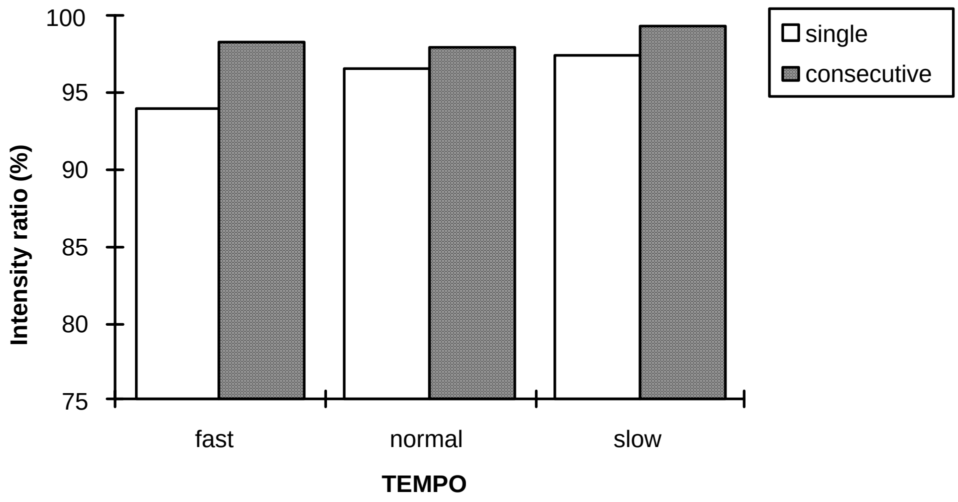
Japanese, the F1 and F2 of the vowels [i], [e] and [ɯ] are relatively far apart while [a] and [o] have relatively close F1 and F2. In other words, the intensities of [i], [e] and [ɯ] are generally less than those of [a] and [o]. In this experiment, all devoiceable vowels were either [i] or [ɯ] with an inherently weak intensity. This may have lowered the average intensity ratios of devoiceable vowels against non-devoiceable vowels that are inherently greater in intensity.

It was extremely difficult to find ideal test words for comparing intensities in both devoicing and non-devoicing environments, and therefore the type of vowel tested was not always the same. Although there were differences between the intensities of voiced vowels in the devoicing and non-devoicing environments, this might simply have been due to the different types of vowels in the two environments. Under equal conditions, high vowels have intrinsically lower intensities than non-high vowels, and all vowels in the devoicing environment are high vowels. Therefore a subset containing test words with the same types of vowels in the two devoicing environments was selected for comparison of intensity ratios. The subset comprised three out of the six original test words: namely, /ta.ka.sa.ki.si˺.mi.N/ 'the Takasaki citizen', /sjo.ku.hi.se.tu.ja.ku/ 'a cut in food expenses' and /ha.i.sju.tu.ki.dju.N/ 'exhaust limit'. In all test words, the devoiceable vowels are underlined. /ta.ka.sa.ki.si.mi.N/ has a devoiceable /i/ in a single devoicing environment and other /i/ in non-devoicing enviroments. /sjo.ku.hi.se.tu.ja.ku/ has two /u/ in devoicing environments, the first devoiceable /u/ in a consecutive devoicing environment and the second devoiceable /u/ in a single devoicing environment, as well as another /u/ in /tu/ in a non-devoicing environment. However, all three subjects devoiced all instances of the word final /u/ in the single devoicing environment. Thus the comparison of the intensities will be made only between the first devoiceable /u/ in a consecutive environment and the other /u/ in a non-devoicing enviornment. /ha.i.sju.tu.ki.dju.N/ has two devoiceable /u/ in consecutive devoicing environments and another /u/ in /dju/ in a non-devoicing environment. Accentuation and neighboring sounds still differed and the number of samples was small, especially in the single devoicing environment. This was because devoiceable vowels in the single environment were usually devoiced and the number of voiced samples of devoiceable vowels in the single environment was limited. However, intensity measurements of the subset may provide more realistic results.

The intensity ratios between the same vowels voiced and devoiced were compared in both single and consecutive devoicing environments. This analysis was conducted for each vowel in each subset word uttered at each tempo and by each subject. The results for the subset were similar to the whole set of test words shown in Figure 3. The intensities of voiced devoiceable vowels in the single devoicing environment were generally weaker than those in the consecutive devoicing environment. Subject A did not voice voiced devoiceable vowels at the normal and slow tempi in the single devoicing environment, and Subject B devoiced only one example of devoiceable vowel at the fast tempo in the single devoicing environment.

Since subject A again showed different voicing patterns from Subjects B and C, Subject A's data were eliminated from the statistical analyses. Figure 4 shows the average intensity ratios of similar voiced devoiceable and non-devoiceable vowels in single and consecutive devoicing environments at three tempi for subjects B and C.

Figure 4 ***The intensity ratios between voiced devoiceable vowels and non-devoiceable vowels of the same kind of the subset in single and consecutive devoicing environments at three tempi by subjects B and C***



There was no significant difference in intensity ratio with devoicing condition ([F(1,2) = 4.823, n.s]) nor speaking tempo (F(2,2) = 2.056, n.s.)), nor was there any significant interaction between the two factors ([F(2,2) = 10.261, n.s]). However, the results for the subset were very similar to those of the whole set shown in Figure 3. The lack of significant results may have been due to the low number of replicates in the subset. The following patterns were also noted for the subset: (a) more intensity weakening at all tempi in the single devoicing environment than in the consecutive environment, (b) greatest intensity weakening at the fast tempo and least intensity weakening at the slow tempo in the single devoicing environment, and (c) there seemed to be no tempo effect on intensity in the consecutive devoicing environment.

Vowels in the devoicing environments were not only shorter but also had less intensity than non-devoiceable vowels. In other words, vowels in the devoicing environments are first reduced in duration and intensity, then further devoiced. In extreme cases, the vowels become deleted.

4. Conclusion

The analysis of vowel devoicing patterns found that Japanese vowel devoicing processes were controlled by the syllable structure. Vowels in typical devoicing sites became voiceless only when resyllabification of the remaining consonants was permitted. Therefore, high vowels in the single devoicing sites were almost always devoiced but only some devoiceable vowels became voiceless in the consecutive devoicing sites. Moreover, the experimental results showed that voiced high vowels in a typical devoicing environment were often not actually fully voiced and were reduced in duration. They also demonstrated that these voiced devoiceable vowels were not only shorter but also have smaller intensity. In other words, it seemed to be more natural for high vowels to undergo the devoicing process between voiceless sounds; so when they did remained voiced, they were acoustically shorter and weaker than when they occur in the non-devoicing environment.

This suggests that vowel devoicing is part of a vowel weakening process and the final state of the process is completely voiceless vowels. The data from consecutive devoicing cases suggest that the vowel weakening process may be influenced by Japanese syllable structure. Two different mechanisms, namely phonetic and phonological processes, seem to control Japanese vowel devoicing depending on the environment. When vowel devoicing is not favored because of the constraint on the syllable structure, vowel weakening does not occur, and the intensity of a voiced devoiceable vowel remains as high as that of non-devoiceable vowels. It should also be noted that vowel weakening in Japanese primarily affects vowel intensity, then duration, and the quality of the vowels remained relatively unchanged regardless of the intensity level of the vowel. The vowel weakening process in Japanese seems to work differently from the similar process in other languages.

References

Beckman, Mary E. and Atsuko Shoji. 1984. Spectral and Perceptual Evidence for CV Coarticulation in Devoiced /si/ and /syu/ in Japanese. *Phonetica* 4. 61-71.

Bell, A. (1978) Syllabic consonants, in Greenberg, J.H. (Ed.*) Universals of human language Vol.2 Phonology*, Stanford University Press, California.

Dauer, R.M. (1980) The reduction of unstressed high vowels in modern Greek, Journal of International Phonetic Association, 10(1-2): 17-27.

Campbell, Nick and Yoshinori Sagisaka. 1991. Moraic and Syllable-level Effects on Speech Timing. *Journal of Electronic Information Communication Engineering* SP 90-107. 35-40.

Cutler, Anne and Takeshi Otake. 1997. Contrastive Studies of Spoken-language Perception. *Journal of the Phonetic Society of Japan* 1:3. 4-13.

Han, Mieko S. 1994. Acoustic Manifestations of Mora Timing in Japanese. *Journal of the Acoustic Society of America* 96:1. 73-82.

Hattori, Noriko. 1989. *Mechanisms of Word Accent Change: Innovations in Standard Japanese*. Doctral dissertation. University College London.

Haye, Bruce. 1989. Compensatory Lengthening in Moraic Phonology. *Linguistic Inquiry* 20. 253-306.

Itô, Junko. 1990. Prosodic Minimality in Japaneses, in K. Deaton, M. Noske and M. Ziolkowski, eds., *CLS 26-II: Papers from the Parasession on the Syllable in Phonetics and Phonology*. 213-239.

Jun, Sun-Ah and Mary E. Beckman. 1993. A Gestural-overlap Analysis of Vowel Devoicing in Japanese and Korean. Paper presented at the 1993 Annual Meeting of the LSA, Los Angeles, 7-10 January, 1993

Kenstowicz, Michael. 1994. *Phonology in Generative Grammar*. MA: Blackwell.

Kondo, Mariko. 1995. Temporal Adjustment of Devoiced Morae in Japanese, *Proceedings of the 13th International Congress of Phonetic Sciences* 3. 238-241.

——— 1997. *Mechanisms of Vowel Devoicing in Japanese*. Doctral dissertation. University of Edinburgh.

——— 2001, Vowel Devoicing and Syllable Structure in Japanese, *Japanse/Korean Linguistics*, Vol. 9, CSIL, Stanford.

Kubozono, Haruo. 1989. The Mora and Syllable Structure in Japanese: Evidence from Speech Errors. *Language and Speech* 32: 3. 249-278.

McCawley, James. 1977. Accent in Japanese, in M.L. Hyman, ed. *Studies in Stress and Accent*, *Southern California Occasional Papers in Linguistics No. 4*. University of Southern California. 261-302

Maekawa, Kikuo. 1989. Boin no musei-ka. in M. Sugito, ed. *Nihon-go no Onsei-On'in (1)*. Tokyo: Meiji Shoin. 135-153.

Port, Robert F., Jonathan Dalby and Michael O'Dell. 1987. Evidence for Mora Timing in Japanese. *Journal of the Acoustic Society of America*, 81: 5. 1574-1585.

Poser, William. 1990. Evidence for Foot Structure in Japanese. *Language* 66. 78-105.

Sakurai, Shigeharu. 1985. Kyootsuu-go no hatsuon de chuui subeki kotogara. Appendix to NHK, ed., *Japanese Pronunciation and Accent Dictionary*. Tokyo: NHK Publication. 128-143.

Sato, Yumiko. 1993. The Durations of Syllable-final Nasals and the Mora Hypothesis in Japanese. *Phonetica* 50. 44-67.

Shibatani, Masayoshi. 1990. *The Languages of Japan*, Cambridge: Cambridge University Press.

Shinohara, Shigeko. 1997 The Roles of the Syllable and the Mora in Japanese Adaptations of French Words. *Cahiers de Linguistique Asie Orientale* 25: 1. 87-112. Paris: CRLAO EHESS.

Sugito, Miyoko and Hajime Hirose. 1988. Production and Perception of Accented Devoiced Vowels in Japanese. *Annual Bulletin of Research Institute of Logopedics and Phoniatrics* 22. 21-39. University of Tokyo.

Takeda, Kazuya and Hisao Kuwabara. 1987. Analysis and Prediction of Devocalizing Phenomena (in Japanese). *Proceedings of Acoustic Society of Japan*. 105-106.

Vance, Timothy J. 1987. *An Introduction to Japanese Phonology*. Albany. NY: State University of New York Press.

Yoshida, Natsuya and Yoshinori Sagisaka. 1990. Factor Analysis of Vowel Devoicing in Japanese (in Japanese), *ATR Technical Report TR-I-0159*. Kyoto: ATR Interpreting Telephony Research Laboratories.