



明治学院大学機関リポジトリ  
<http://repository.meijigakuin.ac.jp/>

Title	On Vowel Devoicing in Japanese
Author(s)	VARDEN, J. Kevin
Citation	明治学院大学教養研究センター紀要カルチュラル = The MGU Journal of Liberal Arts Studies KARUCHURU, 4(1): 223-235
Issue Date	2010-03
URL	<a href="http://hdl.handle.net/10723/83">http://hdl.handle.net/10723/83</a>
Rights	

# On Vowel Devoicing in Japanese

J. Kevin VARDEN

## Abstract

This paper reports on research into High Vowel Devoicing in Standard Modern Japanese. The data discussed support an analysis both phonological and phonetic: phonological devoicing involving the feature [+spread glottis] to explain devoicing at very slow speech rates, and phonetic devoicing involving gestural overlap of the vowel's voicing gesture by glottal spreading gestures to explain observed speech rate effects. In addition, the following conclusions can be drawn from the current data set: 1st mora vowels devoiced significantly more often than 2nd mora vowels. Males devoiced more often than the females. SR had a smaller but still significant effect; devoicing environment had less influence still. Many instances of consecutive devoicing were observed. Finally, the extreme individual variation observed warrants caution extending the conclusions of any small-sample study to any wider population.

## 1. Introduction

High Vowel Devoicing (HVD), whereby the high vowels [i] and [u] are pronounced without voicing, is one of the hallmarks of Standard Japanese. It is generally considered rare, although a diverse number of the world's languages display vowel devoicing – see Kondo (1997) ch.2 for further discussion, Gordon (1998) for a cross-linguistic survey.

Analysis of HVD in the literature has evolved from a context sensitive feature-changing rule operating on the feature [voice] (McCawley 1968) to a P2 rule within Lexical Phonology (Kaisse & Shaw 1985), overlap of glottal spreading gestures in Articulatory Phonology (Jun & Beckman 1993; Imaizumi et al. 1995), spreading of the feature [+spread glottis] in Feature Geometry (Varden 1998), and ranking of constraints within Optimality Theory (Tsuchida 1997, 2001; Teshigawara 2002).

While it has been characterized as a fast-speech rule (Hasegawa 1979; Kuriyagawa & Sawashima 1989; Maekawa 1990a), a fast speech rate (SR) is not a necessary condition. Sakuma (1929) reported that if a word contains only one devoicable vowel it must be devoiced; Bloch (1950) and Kawakami (1977) both maintained that high vowels in a devoicing environment are voiced only in unnaturally careful speech. Vance (1987: 55) suggested a reduced SR effect for some speakers while Beckman (1994: 57) reported devocalization (loss of vowel formants in devoicing sites) occurring even in the most carefully read speech. Kondo (1994) reported one speaker devoicing all nine vowels in non-consecutive devoicing sites at a normal SR, and voicing only one of nine vowels at the slow SR. Han (1994: 76) reported some participants devoicing many vowels at all SRs. In addition, Yuen & Hubbard (1998) found that sociological factors, in particular gender and speaking style,

played a greater role in devoicing than “other factors”, including SR.

One possible source of variation in these studies may be speaker age. Varden & Sato (1996; see also Sato & Varden 1997) found two participants aged 21 and 25 devoicing many more vowels than a participant in her mid-30s, suggesting higher devoicing rates among younger speakers. The study discussed here, Varden (1998), was therefore designed to test the effects of SR and other factors on devoicing by college-age Tokyo speakers, and to explore factors affecting non-devoiced vowel length.

The data set supports a SR effect on HVD for the participants in the study. However, moraic position of the devoicing site and participant gender played a much larger role. Crucially, devoicing was observed at even the slowest SRs. This observation combined with observed SR effects warrants analyzing HVD as both phonological (i.e. due to feature change) and phonetic (i.e. due to low-level articulator gesture overlap). The study also found evidence for both possible resolutions to a pitch accent/devoicing clash reported in the literature: accented vowel devoicing and pitch accent shift. The pervasive frication accompanying vowels in devoicing environments is discussed in Varden (2009; in press), while full Feature Geometry representations of devoiced, voiced and fricativized Japanese vowels are left for future work.

The organization of the paper is as follows. §2 introduces the methods and procedures used to generate the data set. §3 gives a brief summary of factors affecting token and non-devoiced vowel length, while §4 discusses those affecting devoicing. §5 reports on observed interactions between devoicing and

pitch accent. §6 reviews the evidence for two devoicing processes, and §7 gives indications for further research.

## 2. Methods & Procedures<sup>1</sup>

Since the primary aim of Varden (1998) was to examine HVD by younger Tokyo Japanese speakers, six female and four male participants ages 18 to 23 were recruited from among the student population of a Tokyo university. All reported no significant L2 ability; all were paid for their participation.

Stimuli consisted of 10 nominal tokens of the canonical form CVCV, C a voiceless obstruent and V a high vowel, followed by the syntactic particle *to* ‘and’ or *ka* ‘or’. Both token vowels were therefore potential devoicing sites since devoicing is a postlexical (P2) rule (Kaisse & Shaw 1985, Kaisse 1990). Tokens were placed in sentence-initial coordinate nominal structures of the form “A and B” or “A or B”. Sentences were standardized at 9 or 10 mora since Japanese is considered a fundamentally mora-timed language (Warner & Arai 2001; Warner & Ogasawara 2006; Brady et al. 2006) and were rendered in Chinese characters and Japanese syllabary. Romanization and translations are given below, with target tokens underlined and standard pitch accent as per NHK (1991) marked by apostrophe; *kuchi* ‘mouth’ is unaccented.

(1) Stimuli used in the study and translations.

Tsuki’ to hoshi ga kakureta.  
‘The moon and stars are hidden.’

Tsuchi’ ka hiryou ga iru.  
‘We need dirt and fertilizer.’

Chichi' to haha ga genki da.  
'Father and mother are healthy.'

Chi'ki to kazoku ga suki da.  
'I like my friends and family.'

Ki'ki to fuan ga tsunoru.  
'Crises and anxiety are mounting'

Kishi' ka kawa ga suzushi.  
'Either the riverbank or the river is refreshing.'

Kuki' ka ha ga joobu da.  
'Either a stem or a leaf is strong.'

Kushi' ka burashi ga hoshii.  
'I wish I had a comb or a brush.'

Kuchi to hana ga ookii.  
'His mouth and nose are big.'

Ki'chi ka honbu ga chikai.  
'Either the military base or main office are close by.'

After ensuring familiarity with all Chinese characters, the sentences were presented in random order on a computer monitor. Participants matched their productions to display times of 4.0s, 2.8s, or 1.7s for the slow, normal and fast SRs, respectively.<sup>2</sup> For training, three random sentences were displayed once at each SR. One training run proved sufficient to learn the paradigm; no difficulties completing the task were observed or reported. Productions were recorded at 44.1 kHz on a Sony DAT recorder and downsampled onto an Apple computer at 22.05 kHz. A total of 180 stimuli were presented to each participant (10 sentences x 3 SRs x 6 repetitions at each SR)

in two blocks with a short break between, resulting in 3600 devoicing sites (10 P x 180 repetitions x 2 devoicing sites in each token).

Initial signal analysis was performed with the software Signalyze™; follow-up measurements were made with Signalyze™ and Praat™ (Boersma & Weenink 2009). Devoicing judgements and vowel voicing durations were taken from waveforms filtered at 1 kHz in conjunction with wide- and narrow-band spectrograms. The start and end of voicing was judged by the onset and offset of the cyclic activity using the sinusoidal envelope of the waveform as a guide, with incidences of irregular voicing ignored (see Varden 1998 online ch.4 for details). Statistical analysis was performed with StatView™ and Super Anova™. Logistic regression ANOVA and ANCOVA models were constructed by University of Washington statistical consultants, one model of each including the factor PARTICIPANT and the other the factor GENDER.

### 3. Factors affecting token and voiced vowel duration

As SR increases, one would expect a decrease in overall token duration and a concomitant decrease in vowel voicing duration. In addition, since vowel devoicing and deletion can drastically affect word durations (Warner & Arai 2001; Warner & Ogasawara 2006), tokens containing two devoiced vowels should be shorter than those containing one, which should be shorter than those containing only voiced vowels. While a full reanalysis of the interplay between SR increase and token/vowel voicing durations for the current data set is still pending, this section provides a brief summary of factors affecting token and

voiced vowel duration.

Two ANCOVA models were used to determine effects of factors on token duration. The ranking of significant factors is given below (#DEV is the number of devoiced vowels in a token; REPETITION is the ordinal number of a given repetition).<sup>3</sup>

(2) ANCOVA factor ranking for effects on token durations (GENDER model).

SR ( $F=1700$ ) > #DEV ( $F=210$ ) > GENDER ( $F=37$ ) > TOKEN ( $F=7.6$ ) > REPETITION ( $F=5.6$ )

(3) ANCOVA factor ranking for effects on token durations (PARTICIPANT model).

SR ( $F=2500$ ) > #DEV ( $F=100$ ) > PARTICIPANT ( $F=97$ ) > TOKEN ( $F=8.8$ )

Combining results from both models gives the following simplistic ranking of effects.

(4) Simplistic factor ranking for effects on token durations.

SR >> #DEV > PARTICIPANT > GENDER > TOKEN > REPETITION

As expected, both SR and the number of devoiced vowels in the token affected token durations. SR in particular had a large effect; recalling that all tokens are 2 mora long, the average token duration across all tokens and participants was 273 ms at the fast SR, 356 ms at the normal SR, and 472 ms at the slow SR. Although there was significant overlap in token durations between the three SRs due to natural variation in speech production, the program used to generate the data was judged to be successful. The number of devoiced vowels in a token also had a large effect; i.e. across SRs and participants tokens containing two devoiced vowels averaged 288 ms, those containing one devoiced vowel averaged 360 ms, while those containing none averaged 448 ms. There was more variation attributable to individual differences than gender, while the

effects due to devoicing environment (segmental context and pitch accent) had much less effect. There was also a very small effect due to variation between different repetitions of the same token by a given speaker, readily attributable to normal speech variation.

The following gives the ranking of effects generated by the two ANOVA models used for effects of factors on voicing durations.

(5) ANOVA factor ranking for effects on vowel voicing durations (GENDER model).

MORA ( $F=290$ ) → SR ( $F=260$ ) → GENDER ( $F=14$ ) → TOKEN ( $F=5.3$ )

(6) ANOVA factor ranking for effects on vowel voicing durations (PARTICIPANT model).

SR ( $F=320$ ) → MORA ( $F=300$ ) → PARTICIPANT ( $F=53$ ) → BLOCK ( $F=18$ ) → TOKEN ( $F=7.4$ )

Combining results from both models give the following simplistic ranking of effects.

(7) Simplistic factor ranking for effects on vowel voicing durations.

MORA, SR >> PARTICIPANT > GENDER > TOKEN

As SR increased, vowel durations in general decreased, but first mora vowels were significantly shorter than those in the second mora across SRs, as can be seen in the following table.

**Table 1 Average voiced vowel durations at each SR.**

SR	Average voiced vowel durations	
	1st mora	2nd mora
fast	56 ms	62 ms
normal	59 ms	78 ms
slow	72 ms	99 ms

As for minor effects, the average vowel voicing durations produced by the females

increased slightly during the course of the experiment while it decreased slightly for the males. Finally, vowel durations decreased more for some tokens than others, although a thorough examination of the factors involved is still pending.

Factors affecting the number of devoiced vowels in a given token will now be examined.

#### 4. Factors affecting number of devoiced vowels

Two ANOVA models were used to determine effects of factors on the number of devoiced vowels in a token, again one model containing the factor GENDER, the other containing the factor PARTICIPANT. The following ranking of main effects was observed for each model.

(8) ANOVA factor ranking for number of devoiced vowels (GENDER model)

MORA ( $F=570$ ) > GENDER ( $F=380$ ) > SR ( $F=89$ ) > TOKEN ( $F=9.2$ )

(9) ANOVA factor ranking for number of devoiced vowels (PARTICIPANT model)

MORA ( $F=630$ ) > SR ( $F=98$ ) > PARTICIPANT ( $F=86$ ) > TOKEN ( $F=10$ )

Collapsing the two results gives the following simplistic ranking of factors affecting the number of devoiced vowels in a token:

(10) Simplistic ANOVA factor ranking for effects on number of devoiced vowels

MORA > GENDER >> SR > PARTICIPANT >> TOKEN

The most significant factor affecting vowel devoicing in the current data set was by far vowel location: first mora vowels (V 1) devoiced significantly more often than second mora vowels (V 2) for nine of ten of the speakers. GENDER also had a large effect on devoicing, with males overall devoicing more than females. The effects of SR and PARTICIPANT

(individual variation) were smaller but noteworthy. The effect of TOKEN (combined segmental and pitch accent influences) was statistically significant with the caveat that the much larger effect of the other factors may have inflated its significance. Compare this result with the ranking of effects given in Yuen & Hubbard's (1998) report on the effect of gender on devoicing rates: vowel length (single or double), speech style (read or conversational), gender, and others (SR, word position, and neighboring consonants). In both studies SR played a subordinate role to gender in regard the number of devoiced vowels.

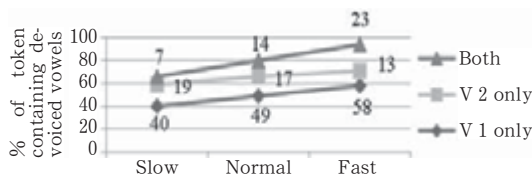
Each of the factors included in this study will now be discussed in more depth.

##### 4.1. Effects of MORA on vowel devoicing

In an earlier study, Kuriyagawa & Sawashima (1989) also noted a higher incidence of V 1 devoicing, attributing it to the larger maximal glottal opening associated with word-initial aspiration than with word-medial aspiration. In the current study the sentence- and phrase-initial placement of the tokens may also have further contributed to prosodic strengthening and hence greater aspiration and devoicing in the first mora. In addition, six of ten tokens used are accented on V 2, making it more likely for the unaccented V 1 to devoice. These two factors most likely conspired to produce the large imbalance in V 1 and V 2 devoicing seen in the data.

The effect of MORA was larger than the effect of SR by an order of magnitude in both statistical models. Across the entire data set, 49% of 1,800 tokens contained only a devoiced V 1, 16% contained only a devoiced V 2, 15% had both vowels devoiced, with 20% having both vowels voiced. There was a significant

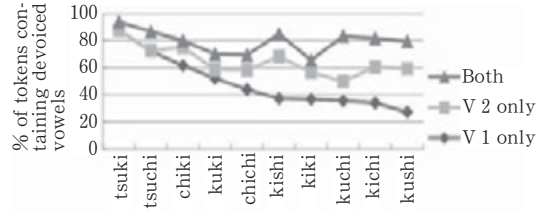
interaction between MORA and SR ( $F_{2, 3569} = 24$ ,  $p < .05$  for the model including GENDER;  $F_{2, 3553} = 26$ ,  $p < .05$  for the model including PARTICIPANT); this effect can be seen in the following figure. (The bottom line gives the percent tokens with a devoiced V 1 only at each SR; the middle line the percent tokens with a devoiced V 2 only at each SR; the top line the percent tokens with both vowels devoiced at each SR. Line positions give cumulative totals.)



**Figure 1** Percent tokens containing devoiced vowels, by mora and SR.

As can be seen, as SR increased (left to right) the percentage of tokens containing both devoiced vowels increased: 7%, 14%, and 23% at the slow, normal and fast SRs, respectively. For those tokens where both vowels did not devoice, devoicing preference shifted toward V 1, with the percentage of tokens containing only a devoiced V 2 decreasing.

There was also a significant interaction between MORA and TOKEN ( $F_{1, 3569} = 48$ ,  $p < .05$  for the model including GENDER;  $F_{1, 3553} = 52$ ,  $p < .05$  for the model including PARTICIPANT). Figure 2 gives the percent vowels devoiced by MORA and TOKEN, tokens arranged in descending order of V 1 only devoiced. Line positions again give cumulative totals; percents are given in Table 2 below.



**Figure 2** Percent tokens containing devoiced vowels by mora and token; tokens in descending order of percent V 1 devoiced.

At the high end of the devoicing spectrum, across the data set 88% of the productions of *tsuki* contained only a devoiced V 1, while none contained only a devoiced V 2. 6% contained both a devoiced V 1 and V 2. At the other end of the spectrum 27% of the productions of *kushi* contained a devoiced V 1 only while 32% contained a devoiced V 2 only; 20% contained both a devoiced V 1 and V 2.

As for the effects of the preceding consonantal on vowel devoicing, Han (1962) suggested the following ranking of the previous consonant's influence: fricative > affricate > stop. Likewise, Takeda & Kuwabara (1987) stated that high vowels are more likely to devoice after fricatives and affricates than after plosives. In figure 2 it is obvious that tokens with initial [tsu] devoiced much more readily than other tokens—88% of the token *tsuki* and 72% of the token *tsuchi* contained a devoiced V 1 across SRs. In addition, [tsu]-initial tokens in this data set are accented on V 2, further contributing to the tendency for V 1 to devoice. Similarly, the segmental effects of a mora-initial [tʃ] can be seen in the rela-

**Table 2** Percent devoiced vowels by mora and token; tokens in descending order of percent V 1 devoiced.

	tsuki	tsuchi	chiki	kuki	chichi	kishi	kiki	kuchi	kichi	kushi
<b>Both</b>	6	14	5	11	11	17	9	33	21	20
<b>V 2 only</b>	0	1	13	7	14	31	20	14	27	32
<b>V 1 only</b>	88	72	62	52	44	37	37	36	34	27

tively high incidence of devoicing of V 1 in *chiki* and *chichi*.

The following figure gives the percent of devoiced vowels for each preceding segment by mora. Note that no first mora-initial fricatives were included in the study.

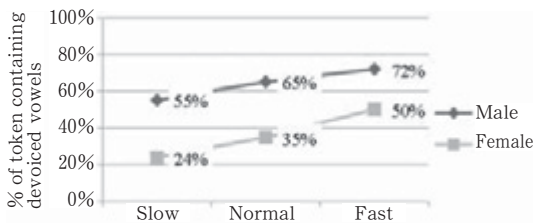


**Figure 3** Percent tokens containing devoiced vowels by mora and mora-initial segment.

The above devoicing rates are in general agreement with previous studies, especially when one considers that [ts] is generally sustainable in Japanese, unlike [tf] which is not. [ts] therefore classes with the fricatives [s] and [ʃ]. Further studies using a more balanced set of segments are called for to further tease out consonantal environment effects.

**4.2. Effects of GENDER and PARTICIPANT on vowel devoicing**

The factor having the second largest effect on vowel devoicing was GENDER. The differences in the number of vowels devoiced in the study by females (n = 6) and males (n = 4) can be seen in the figure below.

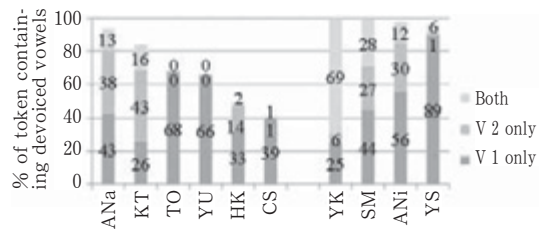


**Figure 4** Percent devoiced vowels by gender and speech rate.

While the males in general devoiced more

often than the females, as SR increased the percentage of vowels devoiced by the females increased more than the percentage devoiced by the males. This is taken to be due to the already extremely high devoicing rates for the males at the slow SR.

However, these averages mask the large individual difference among participants. While the effect of PARTICIPANT on devoicing was much smaller than that of MORA, it does signal significant individual variation. The following figure gives the percent devoiced vowels by participant. Individuals in each gender group are ranked in descending order by percent tokens containing at least one devoiced vowel, female participants are on the left (ANa to CS) and males on the right (YK to YS). Bar heights give cumulative totals.



**Figure 5** Percent devoiced vowels by gender and participant.

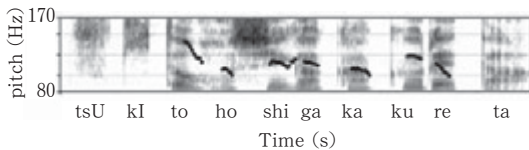
Two things are obvious here. First is the large effect of gender on devoicing – while the males produced very few tokens with no devoiced vowels (0, 1, 2, and 4, respectively, of 180 tokens each), for the females the number of tokens containing no devoiced vowels range from 6% for ANa to 58% for CS.<sup>4</sup> Second, there is large individual variation in the number of tokens with a devoiced V 1 only (26% to 68% for the females; 25% to 89% for the males) and devoiced V 2 only (0% to 43% for the females; 1% to 30% for the males) despite all having similar language back-

grounds. This is consistent with the individual variation in devoicing rate reported in other studies (e.g. Kondo 1994, 1997) and speaks to extreme caution in extending results of studies with small numbers of participants to larger populations.

### 4.3. Consecutive devoicing

One of the supposed constraints on vowel devoicing in Standard Japanese is a prohibition against consecutive devoicing, the devoicing of vowels in consecutive mora. However, both Maekawa (1990a, b) and Maekawa & Kikuchi (2005) reported consecutive devoicing; in the latter's analysis of the Corpus of Spoken Japanese, 84 of 318 (26.4%) consecutive devoicing environments contained consecutively devoiced vowels. They also noted when both vowels are preceded by affricates consecutive devoicing approached 75%.

In the current data set all of the token stimuli consist of consecutive devoicing environments – both vowels of all tokens are devoicable. And indeed, across the current data set 14.7% of 1,800 tokens contained consecutively devoiced vowels. A representative example is given below (U=devoiced [u]; I=devoiced [i]).



**Figure 6** Consecutive devoicing of vowels in *tsuki to* 'moon and', participant YK.

The lack of voicing for both vowels of *tsuki* in this production is evident by the lack of voice bars and pitch trace across [tsUkI]. The range of tokens containing consecutively devoiced vowels for each participant can be seen in Fig-

ure 5 above (the counts associated with 'Both'). Individual rates for consecutive devoicing ranged from 0% for participants TO and YU to 69.4% for participant YK. Eight of the ten participants produced consecutively devoiced vowels; whatever the constraints on its occurrence, consecutive devoicing is certainly not prohibited for these speakers.

We will now turn to the interaction between vowel devoicing and pitch accent placement.

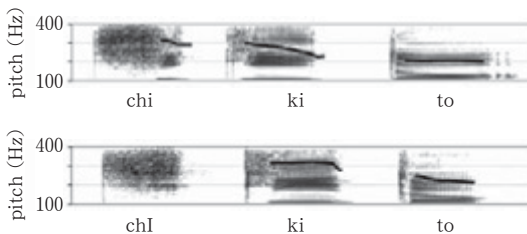
## 5. Interaction between devoicing and pitch accent

There are two common strategies for dealing with a pitch accent/devoicing conflict in Japanese: i) shift the pitch accent to another vowel; or ii) devoice the accented vowel. Works citing pitch accent shift away from a canonically accented vowel in a devoicing environment include Bloch (1950), McCawley (1968, 1977), and Haraguchi (1977). Works reporting devoicing of accented vowels include Sakuma (1929) and Sugito (1966, 1969), a position upheld by experimental studies (Sugito & Hirose 1988; Kondo 1993, 1997; Tsuchida 1997; Kitahara 1997; Varden 1997, 1998). Accented vowel devoicing is also reflected in NHK (1991), an accent dictionary of Standard Japanese compiled by broadcasters, which lists both voiced and devoiced variants for many accented vowels. Both strategies can be seen in the current data set. The loss of phrasal accent accompanying devoicing proposed by Kitahara (1998) will not be discussed here other than to note that the examples reported in Varden (1998) as accent deletion upon closer inspection turn out to be instances of pitch range flattening, where the jump from

the low pitch at the end of one phrase to the high pitch at the beginning of the next phrase was greatly reduced but still present. Discerning the nature of these and other unusual cases of pitch realization seen in the data set are left for future work.

### 5.1. Pitch accent shift

One strategy for avoiding accented vowel devoicing is to shift the pitch accent to another vowel. While this pitch shift strategy may be disappearing for younger speakers (Kondo 1997; Tsuchida 1997), it can still be seen in the current data set. The following figure gives two productions of *chiki to* ‘friends and’ by participant KT. The first is at the normal SR; the second at the slow SR.



**Figure 7** Productions of *chiki* showing accent shift to second mora.

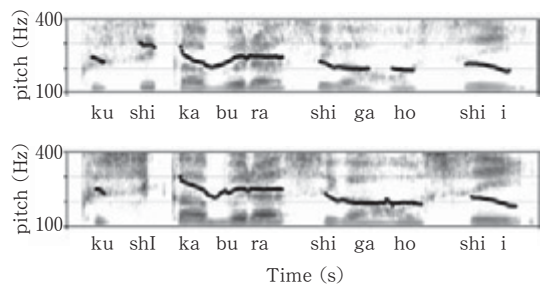
NHK (1981) lists two pronunciations for this word, both with V 1 devoiced and V 2 voiced. The first pronunciation has accent on the devoiced V 1; the second has accent on the voiced V 2. In the first production above (upper panel) both vowels of the token *chiki* are perhaps unnaturally voiced but clearly show the prototypical pitch pattern for a word accented on V 1; the pitch falls from the high level associated with the accent to the low pitch level associated with the end of the nominal phrase (see Pierrehumbert & Beckman 1988 for extensive discussion of accent assignment in the Tokyo dialect). Although not a standard

pronunciation according to NHK (1981), in the second production (bottom panel) KT chose to shift the accent to V 2 as illustrated by the sharp drop from the sustained high level pitch to the low level on the following mora.

The fact that only 3 of 18 repetitions of *chiki* by KT were accented on V 2 speaks to this being a lesser-used option. Indeed, while accent shift is observable in the data set, devoicing of accented vowels was the more commonly applied strategy.

### 5.2. Accented vowel devoicing

In addition to the accent shift discussed above, participants in this study also produced many accented devoiced vowels. This is illustrated in the following productions of *Kushi ka burashi ga hoshii* ‘I wish I had a comb or a brush’ by participant HK.



**Figure 8** Productions of *kushi* showing both devoiced (upper panel) and devoiced (lower panel) accented vowel.

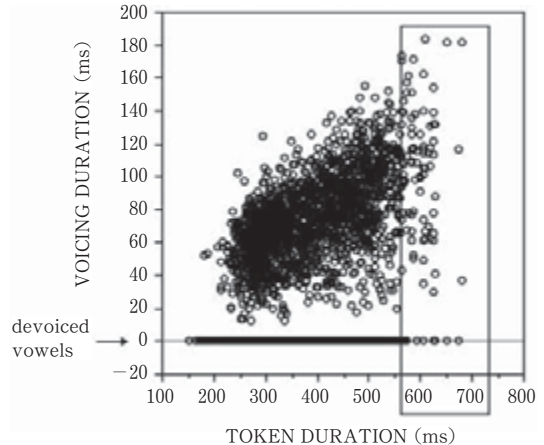
In the upper panel both vowels of *kushi* are devoiced; in the bottom panel the second vowel has been devoiced. Other than the missing pitch across [ʃi] the pitch contours are very much the same. Note in particular the initially low pitch on [ku] and the large pitch excursion across [ka]. The cues to the V 2 accent location in the second production (lower panel) have not been lost because of the devoicing but instead the location is recoverable

from the overall pitch contour as per Sugito (1966, 1969). It is clear from the initially high pitch across [ka] that the vocal cords have been tensed during the production of [ʃi] in the bottom panel even though no voicing was produced.

Further analysis is required to determine the exact interplay of pitch accent placement and vowel devoicing for the entire data set.

## 6. Two devoicing processes

Perhaps the most significant contribution of Varden (1998) was to confirm for the continued existence of a phonological process as one source of vowel devoicing. As discussed in §2, a gestural overlap model of devoicing, whereby the glottal spreading gestures of surrounding obstruents overlap the vowel voicing gesture, makes the prediction that as SR increases the consonantal glottal spreading gestures will increasingly overlap the vowel voicing gesture until complete devoicing occurs. On the other hand, as SR slows the vowel voicing gestures will be less and less overlapped; at a sufficiently slow SR no overlap and hence no devoicing should occur. However, HVD can be observed even in very slow, careful speech (Kondo 1994; Han 1994). The following plot of voicing duration versus token duration, modified from Varden (1998), shows this clearly.



**Figure 8** Scattergram plot of vowel VOICING DURATION vs. TOKEN DURATION, all SRs (total vowels,  $N=3,598$ ; devoiced vowels,  $N=1,696$ ; voiced vowels,  $N=1,902$ ).

Devoiced vowels are represented by the marks on the line at VOICING DURATION=0 ms; voiced vowels are represented by the marks above this line. Note that devoiced vowels appear in tokens over 680 ms long, a fairly slow SR. This devoicing can not be due to gestural overlap – there is ample time for the vowel voicing gesture to be realized, attested to by the fact that many tokens at this slower SR contain voiced vowels (marks within the rectangle above the line). At the other end of the TOKEN DURATION spectrum 65 tokens with durations of shorter than 250 ms contain voiced vowels. If all devoicing were due to gestural overlap, both conditions would not occur; shorter tokens produced at the faster SRs would contain all devoiced vowels, and longer tokens produced at the slower SRs would contain all voiced vowels. This is clearly not the case. Phonological devoicing must still be at work in HVD, and has a place beside phonetic devoicing.

Taking into account the above discussion and other results reported in the literature, it

is clear that the term HVD is, at least in Japanese, an inaccurate umbrella term for a variety of processes resulting in the same phonetic output—a set of acoustic cues that allows the listener to recover the underlying vowel and pitch accent placement.

## 7. Summary and future directions

This paper presents the major findings of Varden (1998). In particular, it provides strong support for the continued existence of a phonological devoicing at work in the speech of younger Tokyo Japanese speakers, in addition to the phonetic devoicing discussed in Yoshioka (1981), Tsuchida et al. (1997). In this data set the factor having the largest effect on token durations was, unsurprisingly, SR, with the number of devoiced vowels in a token playing a secondary role. SR similarly affected vowel voicing durations, as did the moraic position of the vowel. In regard to the number of devoiced vowels, the most significant factor was moraic position; first mora vowels devoiced more often across the data set than second mora vowels. Keeping in mind the significant individual variation seen in the data, the males as a group devoiced significantly more vowels than the females.

Simplistic rankings for effects for each of the measures made are repeated below.

(11) Simplistic ranking of factors affecting mean token durations.

MORA > #DEV > PARTICIPANT > GENDER > TOKEN  
> REPETITION

(12) Simplistic factor ranking for effects on vowel voicing durations.

MORA, SR >> PARTICIPANT > GENDER > TOKEN

(13) Simplistic ranking of factors affecting

number of devoiced vowels.

MORA > GENDER >> SR > PARTICIPANT >> TOKEN

As for the interaction between pitch accent placement and vowel devoicing, while apparent pitch accent shift can be found in the current data set the devoicing of accented vowels is the more commonly used strategy for these participants. Finally, consecutive devoicing was also strongly present.

A more detailed analysis of the current data set is warranted for the following: the interplay between elicited SR and token durations, voicing durations, and the number of devoiced vowels in a token; the interplay between pitch accent placement and devoicing; the frication accompanying devoicing sites (see Varden in press); and the effects of SR on aspiration and its relation to devoicing.

### Notes

- 1 This section provides a succinct overview of the methods and procedures of Varden (1998; available online); see that work for further discussion.
- 2 Compare these durations to the 3.3s, 2.5s and 1.7s durations utilized in Kondo (1994) for 15-mora carrier sentences, which effectively imposed a faster SR on participants.
- 3 *F*-values calculated at the .05 level of confidence and reported to the nearest two significant digits. See Varden (1998) online Appendix for a full listing of *F*-ratios for all effects and interactions.
- 4 Of note is that many of CS' "voiced" vowels were of very short duration (e.g. 20 ms). Also, the experimental paradigm may well have affected her productions in unknown ways—her vocal quality at especially the fast SR betrays a great deal of stress. HK did not display a similar level of stress.

### References

- Beckman, Mary E. (1994) When is a syllable not a syllable? *Ohio State Working Papers in*

- Linguistics* No. 44, Papers From the Linguistics Laboratory, pp. 50–69. Columbus, OH: The Ohio State University.
- Bloch, Bernard (1950) Studies in Colloquial Japanese. Part IV: Phonemics. *Language* 26: 86–125.
- Boersma, Paul & Weenink, David (2009). Praat: doing phonetics by computer (Version 5.1.12) [Computer program]. Retrieved August 4, 2009, from <http://www.praat.org/>
- Brady, Michael Connolly, Robert F. Port & Kyoko Nagao (2006). Effects of speaking style on the regularity of mora timing in Japanese. *JASA* 120. 5: 3208.
- Gordon, Matthew (1998). The phonetics and phonology of non-modal vowels: a cross-linguistic perspective. *Berkeley Linguistics Society* 24: 93–105.
- Han, Mieko S. (1962) Unvoicing of vowels in Japanese. *Onsei no Kenkyuu (Phonetic Research)* 10: 81–100.
- Han, Mieko S. (1994) Acoustic manifestations of mora timing in Japanese. *Journal of the Acoustical Society of America* 96. 1: 73–82.
- Haraguchi, Shosuke (1977) *The Tone Pattern of Japanese: An Autosegmental Theory of Tonology*. Tokyo: Kenkyuusha.
- Hasegawa, Nobuko (1979) Fast Speech vs. Casual Speech. In P. R. Clyne, W. F. Hanks & C. L. Hofbauer (Eds.) *Papers from the Fifteenth Regional Meeting of the Chicago Linguistic Society* pp. 126–137. Chicago: Chicago Linguistic Society.
- Imaizumi, Satoshi, Akiko Hayashi & Toshisada Deguchi (1995). Listener adaptive characteristics of vowel devoicing in Japanese dialogue. *JASA* 98. 2: 768–778.
- Jun, Sun-Ah & Mary E. Beckman (1993). A gestural-overlap analysis of vowel devoicing in Japanese and Korean. Paper presented at the 1993 Annual Meeting of the Linguistic Society of America January 1–10, 1993, Los Angeles, CA.
- Kaisse, Ellen M., & Patricia Shaw (1985) On the theory of Lexical Phonology. *Phonology Yearbook* 2: 1–30.
- Kaisse, E. M. (1990). Toward a typology of Post-lexical rules. In Sharon Inkelas & Draga Zec, Eds. *The Phonology-Syntax Connection*. Chicago, University of Chicago Press: 127–143.
- Kawakami, Shin (1977) *Nihongo Onsei Gaisetsu (An Overview of Japanese Phonetics)*. Tokyo: Oofuusha.
- Kitahara, Mafuyu (1997) Pitch accent, phrase tones, and vowel devoicing in Tokyo Japanese. *Journal of the Acoustic Society of America* 101.5: 3195. Poster presented at the 133rd meeting of the ASA, State College, PA, June 20–26.
- Kitahara, Mafuyu (1998). Vowel devoicing and the loss of lexical accent in Tokyo Japanese. *The Journal of the Acoustical Society of America* 103 (5): 2889–2890. Poster presented at the 135th annual meeting of the ASA, Seattle, WA, June 20–26.
- Kondo, Mariko (1993) The effects of blocking factors and constraints on consecutive vowel devoicing in Standard Japanese. Poster presented at the Fourth Conference on Laboratory Phonology, Oxford, England, August 11–14, 1993.
- Kondo, Mariko (1994) Mechanisms of vowel devoicing in Japanese. In *Proceedings of the 1994 International Conference on Spoken Language Processing*. Acoustical Society of Japan, pp. 61–64.
- Kondo, Mariko (1997) *Mechanisms of vowel devoicing in Japanese*. Ph.D. dissertation, University of Edinburgh.
- Kuriyagawa, Fukuko & Masayuki Sawashima (1989) Word accent, devoicing and duration of vowels in Japanese. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics* 23: 85–108; University of Tokyo.
- Maekawa, Kikuo & Hideaki Kikuchi (2005). Corpus-based analysis of vowel devoicing in spontaneous Japanese: an interim report. In Jeroen van de Weijer, Kensuke Nanjo & Tetsuo Nishihara, Eds, *Voicing in Japanese*. Berlin: Mouton de Gruyter: 205–228.
- Maekawa, Kikuo (1990a) *Hatsuwa sokudo niyoru yuusei kukan no hendoo* (Speech Rate effects on Japanese vowel devoicing) Technical Report SP 89–148, *Denshi Jooohoo Tsuusin Gakkai*, pp. 47–53.
- Maekawa, Kikuo (1990b). Production and perception of the accent in the consecutively devoiced syllables in Tokyo Japanese. In *Proceedings of the International Conference on Spoken Language Processing* 2: 517–520 Kobe, Japan.
- McCawley, James D. (1968) *The Phonological Component of a Grammar of Japanese*. Mouton: The Hague.
- McCawley, James D. (1977) Accent in Japanese.

- Studies in Stress and Accent (Southern California Occasional Papers in Linguistics 4)* LA: Dept. of Linguistics, UCLA.
- NHK (Nippon Hoosoo Kyoukai [Japanese Broadcasting Company]) (1991) *Nihongo Hatsuo Akusento Jiten (Japanese Pronunciation Accent Dictionary)*. Tokyo: Nihon Hoosoo.
- Pierrehumbert, Janet & Mary Beckman (1988). *Japanese Tone Structure*. Cambridge: MIT Press.
- Sakuma, Kanae (1929) *Nihon Onseigaku. (Japan Phonetics.)* Tokyo: Hokuseido.
- Sato, Tsutomu & J. Kevin Varden (1997). Taiwanjin ryugakusei ni yoru nihongo boin museika no shuutoku ni tuite. (On the acquisition of Japanese vowel devoicing by Taiwanese exchange students.) *Meiji Gakuin Ronsou* 587: 41–51.
- Sugito, Miyoko & Hajime Hirose (1988) Production and perception of accented devoiced vowels in Japanese. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics* 22: 21–39.
- Sugito, Miyoko (1966) Musei boin to akusento. (Voiceless vowels and accent.) *Nihon Onsegakkai (Japan Acoustic Society)* 123.
- Sugito, Miyoko (1969) Kusakou akusento no aru museika boin. (Reflections on the accent of voiceless vowels.) *Nihon Onsegakkai (Japan Acoustic Society)* 132.
- Takeda, Kazuya & Hisao Kuwabara (1987) Boin museika no yoin bunseki to yosoku syuhoo no kentoo. (Analysis and prediction of devocalizing phenomena.) *Nihon Onkyoo Gakkai Kooen Ronbunshu (Proceedings of the Acoustical Society of Japan)* 3–3–8: 105–106.
- Teshigawara, Mihoko (2002). Vowel Devoicing in Tokyo Japanese. In *Proceedings of the North West Linguistics Conference 2002*: 49–65. Burnaby, BC, Canada: Simon Fraser University Linguistics Graduate Student Association.
- Tsuchida, Ayako, Shigeru Kiritani, & Seiji Niimi (1997) Two types of vowel devoicing in Japanese: Evidence from articulatory data. *Journal of the Acoustic Society of America* 101.5: 3177.
- Tsuchida, Ayako (1997) *The Phonetics and Phonology of Japanese Vowel Devoicing*. Ph. D. dissertation, Cornell University.
- Tsuchida, Ayako (2001). Japanese Vowel Devoicing: Cases of Consecutive Devoicing Environments. *Journal of East Asian Linguistics* 10.3: 225–245.
- Vance, Timothy J. (1987) *An Introduction to Japanese Phonology*. Albany, NY: SUNY Press.
- Varden, J. Kevin & Tsutomu Sato (1996) Devoicing of Japanese vowels by Taiwanese learners of Japanese. *Proceedings of International Conference on Spoken Language Processing* 96. 2: 618–621. Philadelphia, PA; October 3–6, 1996.
- Varden, J. Kevin (1997). High Vowel Devoicing in Standard Modern Japanese: A 'slow-speech' rule in the making? *Meiji Gakuin Ronsou* 98: 53–80.
- Varden, J. Kevin (1998). *On High Vowel Devoicing in Standard Modern Japanese: Implications for Current Phonological Theory*. Ph.D. dissertation, Linguistics, Seattle: University of Washington. <<http://www.meijigakuin.ac.jp/~varden/devoicing/dissertation/>>
- Varden, J. Kevin (2009). Acoustic correlates of devoiced Japanese vowels. *The Journal of the Acoustical Society of America* 126 (4): 2182. Poster presented at the 158th Meeting of the ASA, San Antonio, TX, USA, Oct. 26–30, 2009.
- Varden, J. Kevin (in press). Acoustic correlates of devoiced Japanese vowels: Velar context. *The Journal of English & American Literature and Linguistics* No. 125. Tokyo, Japan: Society for Literature and Linguistics, Meiji Gakuin University.
- Warner, N. & T. Arai (2001). Japanese Mora-Timing: A Review. *Phonetica* 58. 1–2: 1–25.
- Warner, Natasha & Naomi Ogasawara (2006). Japanese mora-timing and processing: The case of devoiced vowels. *JASA* 120. 5: 3208.
- Yoshioka, Hirohide (1981). Laryngeal adjustments in the production of the fricative consonants and devoiced vowels in Japanese. *Phonetica* 38. 4: 236–51.
- Yuen, Chris L-K & Kathleen Hubbard (1998) Vowel devoicing and gender in Japanese. Paper presented at the 1998 Annual Meeting of the LSA; NY, NY; January 8–11, 1998.