

# Vowel duration and its effect on the frequency of vowel devoicing in Japanese: A comparison between Tokyo- and Osaka dialect speakers

Masako FUJIMOTO<sup>†</sup> and Shigeru KIRITANI<sup>‡</sup>

<sup>†</sup>The National Institute for Japanese Language, 3-9-14, Nishigaoka, Kita-ku, Tokyo, 115-8620 Japan,

<sup>‡</sup>Kobe Kaisei College, 2-7-1, Aotani-cho, Nada-ku, Kobe-shi, Hyogo, 657-0805 Japan

E-mail mfuji@kokken.go.jp, kiritani@kaisei.ac.jp

## ABSTRACT

"Vowel devoicing" is a widely observed phenomenon in Japanese. It occurs more frequently in the Tokyo dialect (standard Japanese) and less frequently in the Osaka dialect. Previous studies have suggested that the frequency of devoicing is highly related to the speaking habits of the speakers and how carefully they pronounce their words. It has been said the speakers of the Tokyo dialects pronounce vowels less carefully than the speakers of the Osaka dialects. The present study examines the relation between vowel duration and the occurrence of vowel devoicing.

The results showed the relative duration of the vowel in a mora in a non-devoicing environment was significantly shorter for Tokyo speakers, who consistently showed vowel devoicing, than for the Osaka speakers who did not show vowel devoicing. However, the vowel duration of the Osaka speakers who showed devoicing in fast speech is not necessarily longer than that of Tokyo speakers. Furthermore, the duration of /i/ of Tokyo speakers was constantly short regardless of speech rate, while that of /e/ varies corresponding to speech rate. This finding indicates that Tokyo speakers categorically differentiate the pronunciation of high- and non-high vowels. Phonetic and phonological aspects of vowel devoicing are discussed based upon the present results and those of our physiological study.

## 1. INTRODUCTION

"Vowel devoicing" (hereafter "devoicing") is a widely observed phenomenon in Japanese where high vowels /i/ and /u/ are devoiced when they are preceded and followed by a voiceless consonant [1,2]. Conceptually, it is an assimilation of the vowel to the voicelessness of the surrounding consonants. It may be regarded as a simple physiological process. If so, it is expected that devoicing occurs commonly in many dialects in Japan. However, the degree of its occurrence is different among dialects. It is more observed in the dialects of eastern Japan including the standard (Tokyo) Japanese than in those of western Japan including the Osaka dialect [3,4,5], although both dialects are mora-timed (a type of syllable timed) and have pitch accents.

In qualitative descriptions of traditional studies of Japanese, the differences in the frequency of devoicing among dialects is claimed to be due to the differences in "style" of general pronunciation habits among dialects; i.e., devoicing occurs more frequently in a dialect in which vowels are pronounced less "carefully," such as the Tokyo

dialect, and less frequently in a dialect in which vowels are pronounced more "carefully," such as the Osaka dialect [4,6]. It is also claimed that consonants are long and vowels are short in the Tokyo dialect whereas consonants are short and vowels are long in the Osaka dialects [7,8]. In spite of these claims, little quantitative studies have been done to examine the relationship between the occurrence of devoicing and the style of pronunciation in dialects.

In the present study, segment durations in CV morae were compared between the Tokyo and the Osaka dialects, with reference to the occurrence of devoicing.

## 2. EXPERIMENT

### Test materials

Test words were /kite/, /kide/, /kete/ and /kede/. Note that /i/ in /kite/ has the typical devoicing environment. These test words were uttered in the carrier sentence "sore-o \_\_\_-to kae-masu" (I change it to \_\_\_). The accent type of the test words in isolation was specified as "Low-High" for the Tokyo dialect and "High-High" for the Osaka dialect, for a meaningful word such as "ika" ('squid'). However, the pitch patterns in these dialects become comparable ("High- High") in the following carrier sentence [9].

	so	re	o		___	to	ka	e	ma	su
Tokyo	L	H	H		<u>H</u>	H	H	H	L	L
Osaka	H	H	H		<u>H</u>	H	H	H	H	H

(L: Low pitch, H: High pitch)

### Subjects

Seven Tokyo speakers (all male, 20-40 yrs. old) and seven Osaka speakers (5 male and two female, 24-54 yrs. old) served as subjects. They lived at least until the age of 18 either in the Tokyo district or the Osaka districts and lived in the same or surrounding area at the time of data recording.

### Procedures

The subjects read the test words in six different random orders at the subject's comfortable normal and fast speaking rate. They were asked to speak without any noticeable break within the sentence. In order to assure this type of pronunciation, a practice session using a word like "ika" preceded the data recording.

All the utterances were recorded with a DAT at 44kHz sampling rate. Data were quantized into a computer at a 16-bit rate and analyzed using the program Multi speech. T-tests were done to statistically evaluate the results.

### 3 RESULTS

#### 3.1 Vowel devoicing in /kite/

Table 1 shows the voicing pattern of the vowel /i/ in /kite/. At normal speech rate, /i/ in /kite/ was devoiced for all the Tokyo speakers and was voiced for all the Osaka speakers. At the fast speech rate, it was also devoiced without exception for Tokyo speakers. As for Osaka speakers, it was voiced for three subjects and devoiced for the other three. The remaining one subject produced 4 devoiced /i/ among 6 tokens. Hence, in the analysis of fast speech, Osaka speakers are divided into two groups excluding the subject whose voicing type is not consistent in fast speech. Hereafter, the Osaka speakers whose /i/ in /kite/ were voiced regardless of speech rate are referred to as the "Osaka voiced group", and those whose /i/ in /kite/ at fast speech were devoiced are referred to as the "Osaka semi-devoiced group." No devoicing was observed except with the test word /kite/.

Table 1. Voicing pattern of the vowel /i/ in /kite/

Speakers		Speech rate	
		Normal speech	Fast speech
Tokyo n=7		Devoiced	Devoiced
Osaka	n=3	Voiced	Devoiced
	n=1	Voiced	Voiced/Devoiced
	n=3	Voiced	Voiced

#### 3.2 Segment duration of /kV/ in /kide/, /kede/ and /kete/

Speech tempi varied from subject to subject and on the average were slower for Osaka speakers. In order to make comparison among speakers without the effect of speech tempi, the durations of the consonants and the vowels were normalized by the mora duration.

Table 2 shows the relative duration of the segments of /kV/ at normal speech rate. The segments in the column in /ki/ are for /kide/ and in /ke/, for /kede/ and /kete/. As for /ki/, the duration of the vowel was significantly longer, hence that of the consonant was shorter for the Osaka speakers than for the Tokyo speakers. Within the consonant /k/, the aspiration period was significantly longer for Tokyo speakers than for Osaka speakers, while closure period was significantly longer for Osaka speakers than for Tokyo speakers. Thus, the consonant /k/ is longer for Tokyo speakers due to the longer aspiration period. The same was true for the syllable /ke/.

Table 3 shows the relative duration of the segments of /kV/ in fast speech. When comparing the Tokyo and Osaka voiced groups, the duration of the vowels was again significantly longer, hence that of the consonants was shorter for the Osaka voiced group than for the Tokyo speakers. Within the consonant /k/, the aspiration period was significantly longer for the Tokyo speakers than for the Osaka voiced group, while the closure period was significantly longer for the Osaka voiced group than for the Tokyo speakers. These results were in common for /ki/ and /ke/.

In contrast, when the Tokyo and the Osaka semi-devoiced group are compared, the duration of the vowels was similar for /ke/, and even significantly shorter, for /ki/, for the Osaka semi-devoiced group than for Tokyo speakers. Consequently, the duration of the consonant was similar or longer for the Osaka semi-devoiced group.

However, the aspiration period of /k/ was longer for the Tokyo speakers than for the Osaka semi-devoiced group for both /ki/ and /ke/. This suggests that the relative duration of /k/ for the Osaka semi-devoiced group becomes longer than for the Tokyo speakers in fast speech and they do this by increasing the closure period.

Table 2. Comparison of the relative duration of segments in normal speech. Values were normalized by the mora duration (100).

\*: p<0.05, \*\*: p<0.01

Segments			Speaker groups	
			Tokyo (n=7)	Osaka (n=7)
/ki/	/k/	Closure	42	45 *
		Aspiration	41	33 **
		Total	83	78 **
	/i/	17	22 **	
/ke/	/k/	Closure	41	47 **
		Aspiration	30	22 **
		Total	71	68 **
	/e/	29	32 **	

Table 3. Comparison of the segment duration in fast speech. Normalized by the mora duration (100).

\*: p<0.05, \*\*: p<0.01

Segments			Tokyo (n=7)	Osaka voiced (n=3)	Osaka semi-devoiced (n=3)
			/ki/	/k/	Closure
Aspiration	39	32 **			35
Total	81	77 *			86 *
/i/	19	23 *		14 *	
/ke/	/k/	Closure	41	46 **	50 **
		Aspiration	31	21 **	23 **
		Total	72	67 **	74
	/e/	28	33 **	26	

Comparison was further made for the duration of the vowel and the mora length independent of the speech rate. Figure 1 shows the relationship between the duration of the vowel and the mora. Each data point represents the duration averaged over the 6 tokens of the test word by a speaker at a given speaking rate. The regression lines for each speaker-group are shown as reference.

As for the vowel /i/, the regression line for the Tokyo speakers was lower than that for the Osaka voiced group. This indicates that the vowel is shorter for Tokyo speakers at a given mora duration. The regression line for the Osaka semi-devoiced group was lower than that for the Osaka voiced group. This indicates that the vowel is shorter for the Osaka semi-devoiced group at a given mora duration. The regression line for the Osaka semi-devoiced group is higher than that for the Tokyo speakers at longer mora duration and lower at shorter mora duration. This indicates that the vowel /i/ for the Osaka semi-devoiced group is longer than that of Tokyo speakers at slower speech and shorter at faster speech. It should be noted that the regression line for Tokyo speakers is nearly horizontal for /i/. This indicates that Tokyo speakers keep short duration for /i/ regardless of the speech rate. In other words, they control the mora duration of /ki/ by controlling mainly the consonant duration.

In the case of the vowel /e/ in the lower panel, the regression line for Tokyo speakers is lower than that for the Osaka voiced group. It is nearly identical to that for the Osaka semi-devoiced group. This indicates that, regardless of dialect, the speakers who showed devoicing produced shorter /e/ than the speakers who did not show devoicing. Interestingly, unlike the case of /i/, the regression line for Tokyo speakers declines as speech rate increases, indicating that the vowel duration varied corresponding to the mora duration.

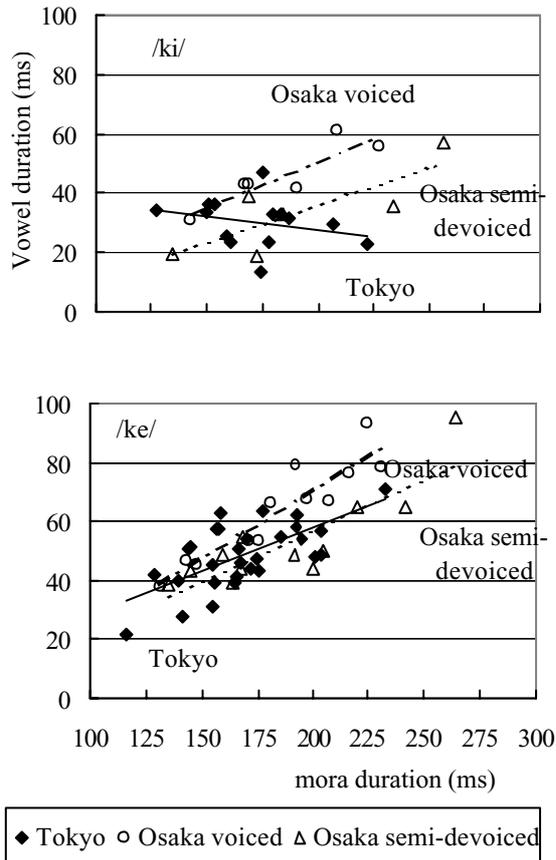


Figure 1. Vowel duration as a function of mora duration in /ki/, upper panel, and /ke/, lower panel. Each data point represents the average of six repetitions at normal and fast speech. Regression lines for each speaker group are shown.

### 3.3 Comparison of segmental duration in /ki/ and /ke/

It is generally known that high vowels are more easily devoiced than non-high vowels. Segment durations of /ki/ and /ke/ shown in tables 2 and 3 were further examined. Table 4 shows the statistical results of the comparison of the segment duration for normal speech rate, and table 5, for fast rate.

As for Tokyo speakers in table 4 the consonant was longer and the vowel was significantly shorter for /ki/ than for /ke/. Within the consonant /k/, the aspiration period was significantly longer for /ki/ than /ke/. The same is true for the Osaka speakers. Also, the same is true for all three groups in table 5.

Table 4. Comparison of the consonant and the vowel duration in the mora /ki/ and /ke/. Results of significance values are shown. Speech rate: Normal.

\*: p<0.05, \*\*: p<0.01, NS: non-significant

Speaker groups Segments		Osaka	Tokyo
		(n=7)	(n=7)
/k/	Closure	NS	NS
	Aspiration	/ki/>/ke/ **	/ki/>/ke/ **
	Total	/ki/>/ke/ **	/ki/>/ke/ **
Vowel		/ki/</ke/ **	/ki/</ke/ **

Table 5. Comparison of the consonant and vowel duration in the mora /ki/ and /ke/. Results of significance values are shown. Speech rate: Fast

\*: p<0.05, \*\*: p<0.01, NS: non-significant

Speaker groups Segments		Osaka	Tokyo	Osaka
		voiced (n=3)	(n=7)	semi- devoiced (n=3)
/k/	Closure	NS	NS	NS
	Aspiration	/ki/>/ke/ **	/ki/>/ke/ **	/ki/>/ke/ **
	Total	/ki/>/ke/ **	/ki/>/ke/ **	/ki/>/ke/ **
Vowel		/ki/</ke/ **	/ki/</ke/ **	/ki/</ke/ **

## 4. DISCUSSION

Vowels are easily devoiced when they are short. Therefore, devoicing and deletion of high vowel or schwa are observed in many languages [10]. Vowels may be also easily devoiced for those speakers who pronounce them shorter than other speakers. In the present results, the vowels for Tokyo speakers and the Osaka speakers who showed devoicing were shorter than those for the Osaka speakers who did not show devoicing. These results were common for both high vowel /i/ and non-high vowel /e/. The tendency to produce vowels shorter in general is considered to be a phonetic factor that induces devoicing for these speakers.

In addition, Tokyo speakers alter the control of the vowel duration between /i/ and /e/, and keep /i/ consistently short regardless of speech rate. It indicates that Tokyo speakers categorically distinguish the pronunciation of high- and non-high vowels. Devoicing for Tokyo speakers is regarded as a categorical and, hence, phonological phenomenon rather than a byproduct induced by phonetic factors.

On the other hand, the duration of /i/ for the Osaka semi-devoiced group varied corresponding to mora duration, and was as short as that for Tokyo speakers at fast speech. This is in good accordance with the result that devoicing occurred for the Osaka semi-devoiced group only in fast speech. Devoicing for these Osaka speakers is regarded as a phonetic phenomenon.

The aspiration period of /k/ was consistently longer for Tokyo speakers than for Osaka speakers. The result suggests that it is aspiration period, but not closure period, that characterizes the stop consonants of Tokyo speakers. Our study on glottal control showed that the longer aspiration period in /k/ for Tokyo speakers is achieved by a longer glottal opening and earlier occurrence of plosion than that for the Osaka speakers. Moreover, Tokyo speakers showed plosion of /k/ at the time of peak glottal opening [11]. Qualitative studies on Japanese claim that consonants in the Tokyo dialect are strong and long [7]. A

longer aspiration period is considered to be a cue for consonants in the Tokyo dialect in order to be perceived strong and long.

Furthermore, it is shown by our physiological study that the glottal opening for /kit/ with a devoiced vowel is significantly longer and larger than that for /k/ in /kid/ with a voiced vowel produced by the same speakers [11]. This indicates that glottal control for /CVC/ with devoiced vowels is especially controlled for Tokyo speakers, as previous studies reported [12,13]. A previous EMG study suggested that upper-level control in speech production is involved for such glottal control [14].

Taking these results into account, devoicing for Tokyo speakers is regarded as a categorical and phonological phenomenon, while that for Osaka speakers is considered a byproduct of a phonetic tendency. Why, then, did this difference arise? One reasonable interpretation is as follows. The tendency to produce long and strong aspiration leads to longer consonants, which, in turn, reduces vowel duration in CV mora. This habit to produce shorter vowels often results in devoicing of high vowels, since their duration is inherently short. Frequent occurrence of devoicing makes Tokyo speakers insensitive to voicing of high vowels. This habit, on one hand, led to altering the pronunciation between high and non-high vowels, and the duration of high vowels in the atypical condition of devoicing becomes short regardless of speech rate. In a typical condition, on the other hand, frequent occurrence of devoicing led Tokyo speakers to produce devoiced syllables automatically by phonological control at the upper level of speech production.

When /ki/ and /ke/ are compared, the vowels were shorter and the aspiration period and consonants were longer in /ki/ than /ke/ regardless of speech rate and of speaker groups. This is considered to be a phonetic factor that induces devoicing in /i/. Our laryngeal observations showed that, unlike the difference between Tokyo and Osaka speakers, shorter duration of /i/ is an outcome of the aerodynamic condition [11].

## 6. CONCLUSION

Comparison of segment duration in /kV/ mora showed that the duration of the vowel for Tokyo speakers was shorter in both /i/ and /e/ than that for the Osaka speakers who did not show devoicing. The tendency to produce shorter vowels is considered to be a factor that induces devoicing in the Tokyo dialect. The vowel duration of /i/ in /kide/ is constant in time for Tokyo speakers, whereas that of /e/ varied corresponding to speech rate. Devoicing in the Tokyo dialect is considered to be a categorical and phonological phenomenon.

The vowel duration for the Osaka semi-devoiced group, who showed devoicing at fast speech, was generally shorter than for the Osaka voiced group, who never showed devoicing. However, unlike Tokyo speakers, the vowel duration for the Osaka semi-devoiced group varied in time corresponding to speech rate for both /i/ and /e/. As a result, the duration of /i/ for the Osaka semi-devoiced group was shorter than that for Tokyo speakers when speech rate was fast. This agrees with the observation of devoicing in fast speech for these Osaka speakers. Consequently, devoicing for these Osaka speakers is considered as a phonetic phenomenon.

The duration of /k/ was generally longer for Tokyo speakers than for Osaka speakers with the exception of the Osaka semi-devoiced group at fast speech. Aspiration period of /k/ was consistently longer for Tokyo speakers than for Osaka speakers without exception. A longer

aspiration period is considered to strengthen the impression of consonant produced by Tokyo speakers.

Comparing /ki/ and /ke/, we see that the vowels were shorter and aspiration period and consonants were longer in /ki/ than in /ke/ regardless of speech rate and of speaker groups. This is considered to be a phonetic factor that induces devoicing in /i/.

## REFERENCES

- [1] T. Hirayama, "Zennihon no hatsuon to akusento. (The pronunciation and the accent in Japan)," Nihon Hoso Kyokai (Eds.), *Nihongo hatsuon akusento jiten*, Tokyo: NHK, pp. 37-69 (In Japanese) 1985.
- [2] M. S. Han, "Unvoicing of vowels in Japanese," *Study of Sounds 10*, Phonetic Society of Japan, pp.81-100, 1962
- [3] M. Sugito, "Boin no museika (Devoicing of Vowel)," *Nihongo no oto*, Osaka:Izumi-shoin, pp. 28-42, (In Japanese) 1996.
- [4] M. Umegaki, "Kinki hogen gaisetsu (The summary of Kinki dialect)," F. Inoue et al (eds.), *Nihon retto hogen sosyo 13*, Tokyo: Yumani shobo, pp. 3-59, (In Japanese) 1996.
- [5] M. Fujimoto and S. Kiritani, "Comparison of vowel devoicing for speakers of Tokyo- and Kinki dialects," *J. of the phonetic society of Japan* (In Japanese), (in press).
- [6] Y. Mase, "Tozai hogen no tairitsu (The contrast between eastern and western dialects)," *Iwanami koza nihongo 11 hougouen* Tokyo: Iwanami, pp. 235-289, (In Japanese) 1977.
- [7] S. Ohno, "Nihon no tobu to sebu. (The eastern and western Japan)," T. Shibata, M. Kato and M. Tokugawa (eds.) *Nihon no Gengogaku 6 Hougouen*. Tokyo: Taishukan, (In Japanese), pp. 100-119, 1978.
- [8] R. Horii, "Kinki hogen no gaisetsu. (The summary of Kinki dialect)," Y. Ino, J. Hibino, and R. Sato (eds.) *Koza hogengaku 7: Kinkichiho no hogen*, Tokyo: Kokusho kankokai, (In Japanese), pp. 1-26, 1982.
- [9] Uwano, Z., "Nihongo no akusento. (The accent of Japanese)," M. Sugito (ed.) *Koza Nihongo to Nihongo kyoiku 2: Nihongo onsei/on'in (jou)*, Tokyo: Meijishoin, pp. 178-205, (In Japanese) 1989.
- [10] M. Beckman, "When is a syllable not a syllable?" T. Otake and A. Cutler (eds.) *Phonological structure and language processing Cross-linguistic studies*, Mouton de Gruyter, Berlin, pp. 95-123, 1996
- [11] M. Fujimoto, E. Murano, S. Niimi and S. Kiritani, "Difference in glottal opening pattern between Tokyo and The Osaka dialect speakers: Factors contributing to vowel devoicing," *Folia Phoniatica et Logopaedica* 2002-54, Karger:Basel, pp. 133-143, 2002.
- [12] M. Sawashima, "Devoiced syllables in Japanese -A preliminary study by photoelectric glottography," *Ann. Bull. of Research Institute of Logopedics and Phoniatics* 3, Univ. of Tokyo, pp. 35-41, 1969.
- [13] H. Yoshioka, A. Löfqvist and H. Hirose, "Laryngeal adjustment in Japanese voiceless sound production," *Journal of Phonetics* 10, pp. 1-10, 1982.
- [14] H. Yoshioka, "Laryngeal adjustment in the production of the fricative consonants and devoiced vowels in Japanese," *Phonetica*38, pp. 236-351, 1981.