DISCRIMINATION ABILITY AND PRONUNCIATION PREFERENCE BETWEEN VOICED AND DEVOICED VOWELS BY NATIVE JAPANESE SPEAKERS

Seiya Funatsu[a], *Satoshi Imaizumi*[b], *Masako Fujimoto*[c] *& Ryoko Hayashi*[d]

[a]Science Information Center, Prefectural University of Hiroshima, Hiroshima, Japan;
[b]Faculty of Health and Welfare, Prefectural University of Hiroshima, Mihara, Japan;
[c]The National Institute for Japanese Language and Linguistics, Tachikawa, Japan;
[d]Faculty of Intercultural Studies, Kobe University, Kobe, Japan
funatsu@pu-hiroshima.ac.jp; imaizumi@pu-hiroshima.ac.jp;
mfuji@ninjal.ac.jp; rhayashi@kobe-u.ac.jp

ABSTRACT

Japanese vowel devoicing was investigated from the viewpoint of perception, an approach which has not been taken in any previous studies. It appears in some dialects of Japanese, including Tokyo dialect.

Three tests, a vowel preference test, a same/different test, and an inter-stimulus interval effect test were performed. In the vowel preference tests, 46.3% of devoiced vowel dialects speakers and 41.2% of voiced vowel dialects speakers could discriminate between devoiced and voiced vowels. These percentages were lower than expected in general. The same/different test resulted in a similar percentage. Moreover, in the short inter-stimulus interval, although we expected that the percentage of subjects who could discriminate would increase, on the contrary it decreased.

From these results, it appears that vowel devoicing might not be necessary for Japanese pronunciation.

Keywords: Japanese vowel devoicing, vowel preference test, dialects, devoiced vowel, voiced vowel

1. INTRODUCTION

There are many studies on the Japanese vowel devoicing, which frequently appears in some dialects of Japanese, including Tokyo dialect (standard pronunciation). These studies are from the viewpoint of production.

According to previous studies, the high vowels [i] and [u] tend to be pronounced as devoiced vowel [i̥] and [u̥] with a high probability when surrounded by voiceless consonants. Such allophonic variations do not affect word meaning. For instance, [kokusai] versus [koku̥sai], both mean "international", although their acoustical and perceptual quality is different. These devoiced vowels, [i̥] and [u̥], are allophones of vowels [i] and [u], respectively. Moreover, it is confirmed that speech rate affects devoicing, e.g., fast speech rate is easier for devoicing than slow rate [2-4, 6, 8]. However, there are a lot of inconsistent results. For instance, with regard to the influence of the accent upon the devoicing, unaccented [i] and [u] are more devoiced than accented [i] and [u] [2, 4]. However, Maekawa [5] reported that the probability of devoicing does not relate to accent. As for the consonantal environment, vowels preceded by fricatives are easier to devoice than vowels preceded by stops and affricates [2, 5]. On the other hand, Kuriyagawa, et al. [4] showed that the probability of devoicing does not depend upon the types of consonants. In the most recent studies using the large-scale Japanese corpus (The Corpus of Spontaneous Japanese: CSJ), Maekawa and Kikuchi [7] concluded that in the case where fricative is followed by stop or affricate, vowel devoicing occurs most frequently (devoicing rate 97.5%, 94.5%, respectively), and in the case where affricate is followed by fricative, vowel devoicing is the hardest to occur (48.1%).

Almost all the investigations about vowel devoicing are about devoiced vowel production, whereas investigations from the viewpoint of perception are few. Cutler and her colleagues [1], for instance, showed that Japanese speakers do not automatically restore or do not always perceive vowels in devoicing contexts. However, it is not clear that Japanese speakers can discriminate between devoiced and voiced vowels. The present paper investigates the perception of devoiced vowel by native Japanese. If Japanese can discern devoiced vowels from voiced vowels by listening,

vowel devoicing is then necessary for pronunciation, while if they can not, vowel devoicing is inconsequential.

The present study aimed at studying discrimination ability and pronunciation preference between devoiced and voiced vowel by native Japanese speakers. With regard to vowel devoicing, naïve Japanese speakers will be difficult to answer whether vowels are devoiced or not, or whether there are any vowels between consonants. This is because, as mentioned above, voiced and devoiced vowels are allophones, and do not affect word meaning. Accordingly, a vowel preference test, a same/different test, and an inter-stimulus interval effect test were conducted.

2. VOWEL PREFERENCE TEST

Naïve Japanese speakers, as mentioned above, may not be able to answer the question whether vowels are devoiced or not. Thus, we carried out the vowel preference test, in which subjects hear stimuli pairs (voiced-devoiced vowel pairs) and select preferable pronunciation from the pair. We supposed that if subjects are able to discriminate between devoiced and voiced vowels, their responses will be deflected to one of two.

2.1. Methods

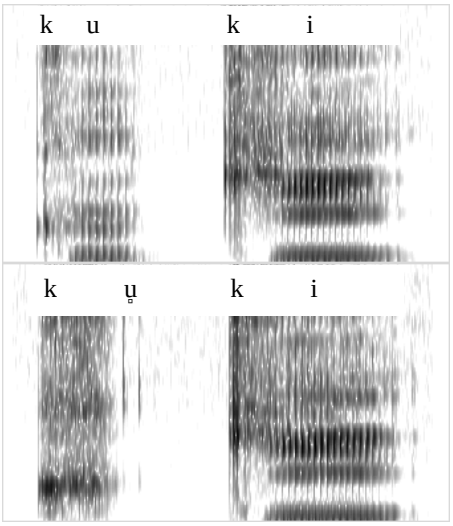
2.1.1. Stimuli

Stimuli were 12 bi-syllabic words which had /tsu/ or /ku/ in word-initial position. The word-initial syllables were followed by voiceless stops or affricates. These words were uttered with a plain accent at normal speed in isolation by one Japanese male and female speaker. They were speakers of so-called Common Japanese. They pronounced initial /tsu/ and /ku/ as voiced [tsu] and [ku] or devoiced [ts̥ɯ̥] and [kɯ̥], respectively. Vowel devoicing was confirmed by visual inspection for spectrogram in these stimuli. Fig. 1 shows examples of spectrograms of the stimuli. Thus, the following 12 pairs were used for the discrimination test.

[tsuta]/[ts̥ɯ̥ta](ivy), [tsutɕi]/[ts̥ɯ̥tɕi](dirt),
[tsutsu]/[ts̥ɯ̥tsu](tube), [tsute]/[ts̥ɯ̥te](connection),
[tsuto]/[ts̥ɯ̥to](suddenly), [tsuka]/[ts̥ɯ̥ka](mound),
[tsuki]/[ts̥ɯ̥ki](moon), [kuki]/[kɯ̥ki](stalk),
[kuku]/[kɯ̥ku](multiplication table), [kusa]/[kɯ̥sa](grass),
[kuɕi]/[kɯ̥ɕi](comb), [kutɕi]/[kɯ̥tɕi](mouth)

Voiced and devoiced word order in each pair was counterbalanced. Inter-stimulus interval was 1 second and inter pair interval was 2 seconds. Thus the total 144 pairs (12 pairs x 2 speakers x 6 times) were presented to subjects.

Figure 1: Spectrograms of /kuki/ (top) and /kɯ̥ki/ (bottom).



2.1.2. Subjects

The subjects were 41 university students from the Kanto area (20 from Tokyo, 5 Chiba, 5 Saitama, 4 Kanagawa, 4 Ibaraki, 2 Tochigi, 1 Gumma) where vowel devoicing is easy to occur, and 34 university students from the Chugoku area (27 from Hiroshima, 3 Okayama, 2 Tottori, 1 Yamaguchi, 1 Shimane) where vowel devoicing is hard to occur. They were instructed to hear stimuli pairs carefully, and to mark 1 or 2 on answer sheets to indicate which pronunciation they preferred, first or second.

2.2. Results

We supposed that subjects who selected devoiced vowels more than 70% of the time had a devoiced vowel preference, and those who selected voiced vowels more than 70% of the time had a voiced vowel preference. Moreover, we supposed that the subjects who had devoiced or voiced vowel preference are able to discriminate between devoiced and voiced vowels. Results are shown in Table 1.

For the Kanto subjects 39.0% preferred devoiced to voiced vowels, whereas 29.4% of those from the Chugoku area preferred the opposite. The difference in the preference between Kanto and Chugoku subjects was tested. We performed tests for independence and there was a

significant difference between Kanto and Chugoku subjects ($\chi 2(2)=10.50$, $p<0.01$). Namely, Kanto subjects (devoiced vowel dialect speakers) preferred devoiced vowels, while Chugoku subjects (voiced vowel dialect speakers) preferred voiced vowels.

The subjects who can discriminate between devoiced and voiced vowel were 46.3% in Kanto subjects, and 41.2% in Chugoku subjects. The test for independence did not show a significant difference between Kanto and Chugoku subjects ($\chi 2(1)=0.201$). Thus, a mean 43.8% of subjects was able to discriminate between devoiced and voiced vowels, irrespective of their dialects.

Table 1: Percentages of vowel preference in Kanto (speakers of devoiced vowel dialects) and Chugoku (speakers of voiced vowel dialects) subjects.

Subjects	Voiced	Devoiced	No preference
Kanto	7.3	39.0	53.7
Chugoku	29.4	11.8	58.8
Total	17.3	26.7	56.0

3. SAME/DIFFERENT TEST

We carried out same/different test to evaluate the results of vowel preference test.

3.1. Methods

Stimuli were the same as used in the vowel preference test. The aforementioned 12 pairs and devoiced-devoiced pairs and voiced-voiced pairs were used for the discrimination test. Same as the vowel preference test, inter-stimulus interval was 1 second and inter-pair interval was 2 seconds. Thus a total 192 pairs were presented to subjects.

Subjects were 21 university students from the Chugoku area (they were different subjects from the vowel preference test, 19 from Hiroshima, 1 Yamaguchi, 1 Shimane). Subjects were instructed to hear stimuli pairs carefully and to judge pair words as being either the same or different pronunciation, and to mark whether they were same or different on answer sheets. The percentage of correct answers was calculated for each subject.

3.2. Results

As shown in Table 2, the number of subjects above 70% correct percentage was 9 (42.9%). These subjects were able to discriminate between devoiced and voiced vowels. Comparing this result with the vowel preference test (Chugoku subjects), the percentage of subjects who could discriminate was nearly equal (42.9 vs. 41.2%). The results of vowel preference test and same/different test were approximately the same. Namely, only almost half of subjects were able to discriminate between devoiced and voiced vowels.

Table 2: Results of same/different test.

	Ability to Discriminate	Inability to Discriminate
number	9	14
percentage	42.9	57.1

4. INTER-STIMULUS INTERVAL EFFECT TEST

According to Werker and Logan [9], speech perception is constructed by three factors, auditory, phonetic and phonemic factors, and discrimination performance changes depending on inter-stimulus interval (ISI). Namely, in the short ISI (i. e. 100 ms), subjects show auditory level discrimination, in the middle ISI, subjects exhibit phonetic level discrimination, and in the long ISI, subjects exhibit phonemic level discrimination. Thus, we tested dependence of discrimination performance on ISI.

4.1. Stimuli and subjects

Stimuli were the same as used in the vowel preference test, i.e. 12 voiced and devoiced pairs of bi-syllabic words. ISI was 100 and 500 milliseconds, and inter-pair interval was same as the vowel preference test, i.e. 2 seconds.

Subjects were 30 (ISI 100 ms) and 28 (ISI 500 ms) university students from the Chugoku area (subjects were completely different in all experiments, 100 ms: 29 from Hiroshima, 1 Yamaguchi, 500 ms: 25 from Hiroshima, 3 Shimane). They were instructed to hear stimuli pairs carefully, and to mark 1 or 2 on answer sheets to indicate which pronunciation they preferred, first or second.

4.2. Results

Table 3 and 4 shows the results of ISI effect test. In 100 ms ISI condition, 6.7% of subjects preferred the devoiced vowel and 16.7% of subjects preferred the voiced vowel. In 500 ms ISI condition, 28.6% of subjects preferred the devoiced vowel and 21.4% of subjects preferred the voiced vowel. The subjects who could discriminate between devoiced and voiced vowel were 23.3% in ISI 100 ms condition, and 50.0% in ISI 500 ms condition. In

ISI 500 ms condition, the percentage of subjects who could discriminate was nearly equal to 1000 ms condition. In ISI 100 ms condition, that was almost half of 1000 ms condition.

The percentage of subjects who could discriminate in 100 ms decreased compared with that in 1000 ms, while that in 500 and 1000 ms did not change. In the present case, however, we inferred that the 100 ms case (auditory discrimination) shows better performance than the 1000 ms case (phonemic discrimination). We are, therefore, difficult to conclude that subjects' discrimination level would change from auditory or phonetic level to phonemic level. In 100 ms, discrimination processing would not finish during short stimulus interval. Accordingly, discrimination percentage was lowered in the case of 100 ms.

Table 3: Percentages of vowel preference in Chugoku subjects in 100 ms and 500 ms ISI conditions.

ISI (ms)	Devoiced	Voiced	No preference
100	6.7	16.7	76.7
500	28.6	21.4	50.0

Table 4: Percentages of subjects who could discriminate stimuli.

ISI (ms)	Discriminate
100	23.3
500	50.0
1000	41.2

5. DISCUSSION

In the vowel preference test and the same/different test, we set the discrimination threshold level at above 70%. If the discrimination threshold level was set at 80%, the percentage of the subjects who could discriminate was 23.5% in the preference test and 19.0% in the same/different test. In both tests, the percentage of the subjects who could discriminate was lower in 80% threshold than in 70% threshold. In either case, the percentage of the subjects who could discriminate was below 50%.

Each stimulus in a pair is represented by same *kana* characters. It is, therefore, only possible for Japanese speakers to perceive the difference between voiced and devoiced vowels through careful listening. Consequently, they might be difficult to perceive the difference between voiced and devoiced vowels.

In general, vowel devoicing is thought to be necessary for Common Japanese. In the present results, however, only approximately 40% subjects could discriminate between devoiced and voiced vowels, and the rest of subjects could not discriminate them. Moreover, in the short ISI, although we expected that the percentage of subjects who could discriminate would increase, on the contrary it decreased. Considering these results, vowel devoicing may not be necessary for Japanese pronunciation, but may inconsequential for Japanese speakers. Since they cannot perceive the difference between devoiced and voiced vowels. In any case, further studies are needed for this problem.

6. CONCLUDING REMARKS

In this study, a vowel preference test, a same/different test, and an inter-stimulus interval effect test were performed. Approximately 40% of native Japanese speakers could discriminate between devoiced and voiced vowels. This percentage was lower than our expectation.

7. ACKNOWLEDGEMENTS

This study was supported by Grant-in-Aid for Scientific Research of the Japan Society for the Promotion of Science (No. 20520394, 23520539). Our special thanks are due to Prof. Ayako Shirose (Tokyo Gakugei University) for her cooperation of vowel preference test in Kanto subjects.

8. REFERENCES

[1] Cutler, A., Otake, T., McQueen, J.M. 2009. Vowel devoicing and the perception of spoken Japanese words. *J. Acoust. Soc. Am.* 125(3), 1693-1703.

[2] Han, M.S. 1962, Unvoicing of vowels in Japanese. *The Study of Sounds* 10, 81-100.

[3] Imaizumi, S., Hayashi, A., Deguchi, T. 1995. Listener adaptive characteristics of vowel devoicing in Japanese dialogue. *J. Acoust. Soc. Am.* 98(2), 768-778.

[4] Kuriyagawa, F., Sawashima, M. 1989. Word accent, devoicing and duration of vowels in Japanese. *Annual Bulletin Research Institute of Logopedics and Phoniatrics.* Faculty of Medicine, University of Tokyo 23, 85-108.

[5] Maekawa, K. 1983. Kyotsugo ni okeru boin no museika no kakuritsu ni tsuite, *Gengo no Sekai* 1, 69-81.

[6] Maekawa, K. 1990. Effects of speaking rate on the voicing variation in Japanese. *IEICE Technical Report*, SP-89-148, 47-53.

[7] Maekawa, K., Kikuchi, H. 2005. Corpus-based analysis of vowel devoicing in spontaneous Japanese: an interim report. In van de Weijer, J., Nanjo, K., Nishihara, T. (eds), *Voicing in Japanese*. Mouton de Gruyter, 205-228.

[8] Sugito, M. 2005. The effect of speech rate on devoiced accented vowels in Osaka Japanese. In van de Weijier, J., Nanjo, K., Nishihara T. (eds.), *Voicing in Japanese.* Mouton de Gruyter, 247-260.

[9] Werker J.F., Logan, J. 1985. Cross-language evidence for three in speech perception factors. *Perception & Psychophysics* 37(1), 35-44.