Regional variation of VOT in Ibaraki Japanese

Akira Utsugi (Nagoya Univ.), Kan Sasaki (Sapporo Gakuin Univ.),
and Yosuke Igarashi (Hiroshima Univ.)
utsugi@lang.nagoya-u.ac.jp, ksasaki@sgu.ac.jp, igarashi@hiroshima-u.ac.jp

1. Introduction

This research is part of our broader project on the phonetic and phonological variation and change in Ibaraki Prefecture, Japan. Ibaraki is an intriguing area where several phonetic and phonological phenomena are found that differ from those in Tokyo Japanese, despite being only 40–160 km from Tokyo. Interestingly, this dialect shares several phenomena with some Tohoku dialects that are spoken in northern Japan. Such phenomena include intervocalic voicing and no lexical use of pitch (so-called "accentless dialect") (e.g., Miyajima 1961). In this sense, Ibaraki is an area where two major Japanese dialects meet.

Although our project dealt with several phenomena, this paper focuses on VOT (voice onset time) in word-initial stops. VOT is a time interval between a stop burst and start of voicing (Lisker and Abramson 1964). A positive VOT value indicates that a burst precedes a voicing and a negative value indicates that a voicing precedes a burst. Thus, voiceless stops have positive VOT values. On the other hand, voiced stops can have either negative or small positive VOT values. Hereafter, we will call the former "prevoicing" and the latter "short-lag."

Recently, Takada (2006, 2011) reported that VOT in voiced stops varies regionally in Japanese spoken by older generations. According to her, speakers in Tohoku tend to have short-lag pronunciation for phonologically voiced stops, while speakers in other areas have prevoicing pronunciation. In her study, the linguistic border appeared in the Ibaraki and Tochigi Prefectures in the northern Kanto region.

Within Ibaraki, voiced stop pronunciation gradually changes from the northeast, where short-lag is dominant, to the southwest, where prevoicing is dominant (Takada 2006, 2011). Our study establishes the precise distribution of the two areas where short-lag and prevoicing are dominant. In addition, it describes the gradual change there. When we compare Takada's results with the administrative area classification of Ibaraki, we notice that speakers who frequently use prevoicing are found in the Kensei region (West Ibaraki). Thus, we start our investigation from a working hypothesis that prevoicing is dominant in Kensei and short-lag is dominant in other areas of Ibaraki.

This research is a necessary step to develop our study into one focused on the sociophonetic variation and change of stops in Ibaraki over multiple generations.

2. Methods

2.1. Participants

Fifty-two speakers, aged between 69 and 88 years, participated in this research. Among these speakers, 21 were recruited from the Kensei area (Koga, Yachiyo, Shimotsuma, and Joso) and the other 31 speakers came from other areas of Ibaraki (Kitaibaraki, Mito, Kamisu, Inashiki, and Miho).

2.2. Materials

We used 24 bimoraic words as speech materials, which were the same as those used in Takao (2012). These materials comprised of two datasets; Dataset 1 (20 words) and Dataset 2 (4 words). This paper analyzes only the first dataset. This dataset were designed to investigate VOT as well as other acoustic features such as voice quality although only the former is reported in this paper. Words in this dataset, as shown in Table 1, were selected on the basis of three factors: (i) voice of an initial consonant (voiceless or voiced), (ii) place of articulation of an initial consonant (labial or dorsal), and (iii) the following vowel (/i, e, a, o, u/). Coronal consonants (/t, d/) were not included in the second factor to avoid affricated allophones before /i/ and /u/.

These words were embedded in a carrier sentence: ___ *tte yomuyo* ("it is pronounced as ___").[1]

We prepared three lists of words, each of which comprised 24 words as mentioned above, which were then randomized. Thus, each word was read three times.

Table 1: Materials

	/a/	/i/	/u/	/e/	/o/
Labial, voiceless	/pari/	/piza/	/puro/	/peke/	/poro/
	'Paris'	'pizza'	'professionals'	'cross mark'	'polo'
Labial, voiced	/bara/	/biri/	/buta/	/bero/	/boro/
	'rose'	'tail-ender'	'pig'	'tongue'	'rag'
Dorsal, voiceless	/kari/	/kiri/	/kuri/	/keri/	/kori/
	'hunt'	'fog'	'chestnut'	'kick'	'stiffness'
Dorsal, voiced	/gaku/	/giri/	/guru/	/geta/	/goku/
	'frame'	'duty'	'accomplice'	'clogs'	'phrase'

2.3. Recording procedures

In the recordings, participants read speech materials aloud. The materials were shown as printed paper on which the carrier sentence was not shown. Recordings were conducted either at the participant's home or at public facilities such as a local community center.

[1] We allowed participants to modify a carrier sentence to one that they could usually use. For example, some speakers modified it into "___ *tte yomudo*."

The recordings were made using a portable SD recorder (Marantz PMD 661) and a condenser microphone (AKG C520). Sounds were recorded as WAV files (44100Hz, 16bit) directly by the recorder.

2.4. Analyzing procedures

We first measured VOT for each token. These raw VOT data were subjected to discussion only for voiceless stops.

As for voiced stops, we classified tokens into two categories: “prevoicing,” which has negative values in VOT, and “short-lag,” which has zero or positive values in VOT. We then counted the number of tokens in each category. The reason why raw VOT values were not discussed is because we assume two categorical allophones, i.e., prevoicing and short-lag, in Japanese voiced stop phonemes. We have two reasons for this assumption. First, Japanese voiced stops show bimodal distribution in VOT (Takada 2011: 80ff). A gap in the bimodal distribution appears between 0 ms and −10 ms, indicating that one mode corresponds to prevoicing and the other to short-lag. Second, it has been found that speech rate has different effects on the prevoicing and short-lag categories in several languages such as French, Thai, and Spanish (Kessinger and Blumstein 1997; Magloire and Green 1999). Prevoicing changes its duration as a function of speech rate; a faster speech rate results in a shorter duration of prevoicing. On the other hand, short-lag duration remains more or less stable, regardless of the change in speech rate. These findings suggest that Japanese voiced stops have two categorical allophones, i.e., prevoicing and short-lag. It will be sociophonetically meaningful to investigate the frequency of each category rather than raw VOT values, which would be affected by other phonetic factors such as speech rate.

On the basis of the frequency of two voiced-stop allophones, we calculated “short-lag ratio,” i.e., the number of short lag tokens divided by the total number of tokens in each speaker.

3. Results

As stated previously, we calculated the short-lag ratio for voiced stops. Figure 1 shows histograms of short-lag ratios in Kensei and other areas of Ibaraki. It is clear that in both regions, most speakers showed a ratio higher than 0.8. However, data from Kensei showed a bimodal distribution, one peak above 0.8 and another below 0.4, suggesting that some speakers prefer prevoicing rather than a short-lag for phonologically voiced stops. Such a distribution was not found in other areas.

Figure 2 shows a distribution map in Kensei, in which short-lag ratio was indicated in grayscale. Dots marking a short-lag ratio less than 0.2 were located in the south of Koga, while dots with ratios between 0.2 and 0.4 were located in the south of Yachiyo and the south of Joso.

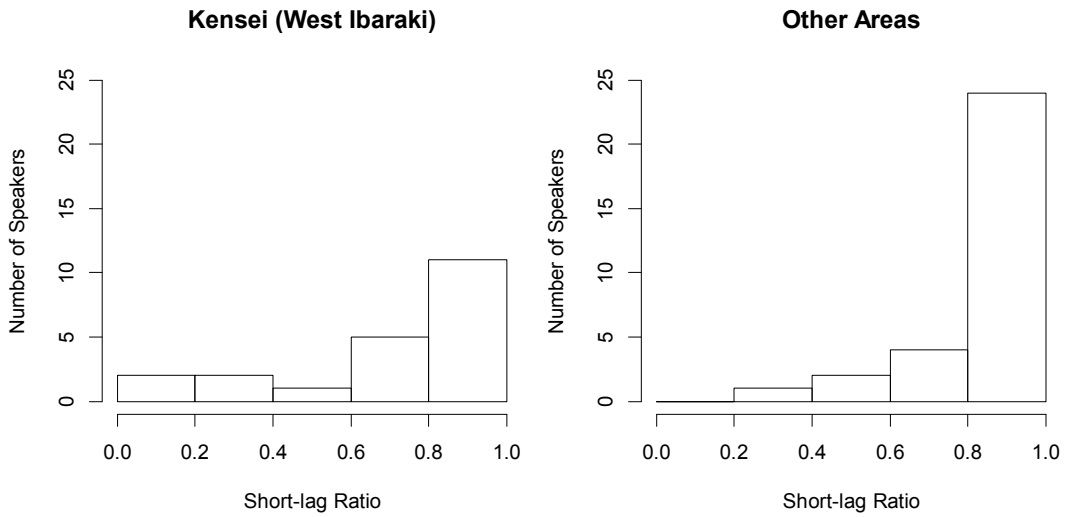

Figure 1: Histograms of short-lag ratios in Kensei (West Ibaraki, left panel) and other areas (right panel).

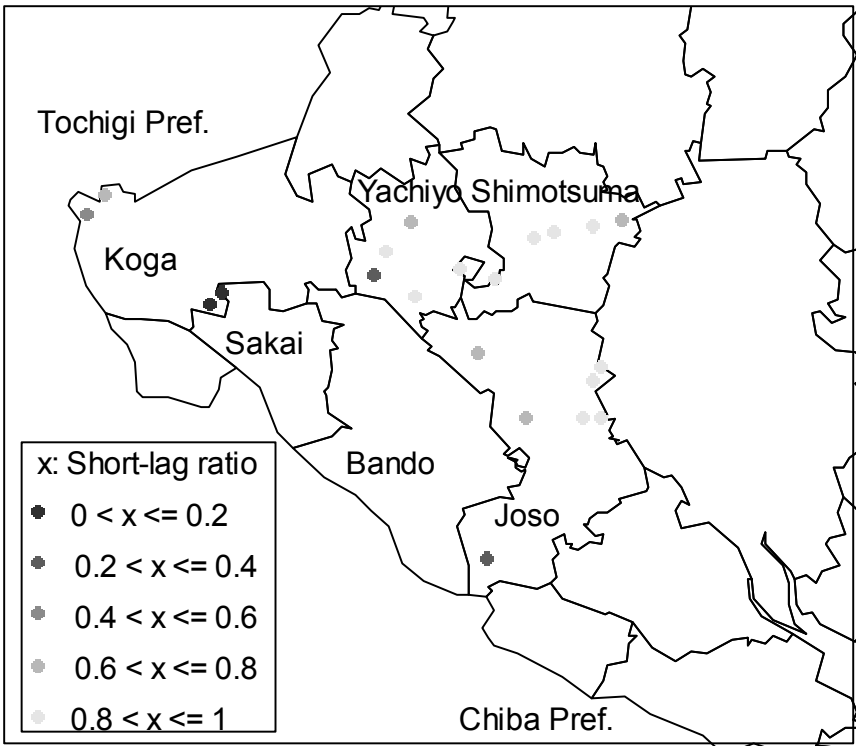

Figure 2: Dot distribution map of short-lag ratios in Kensei (West Ibaraki). Locations of each dot correspond to a speaker's home village. The short-lag ratios are shown in variations of grayscale.

Figure 3 shows the histograms of VOT for voiceless stops. According to ANOVA with one between-subject factor ([Area]: Kensei or other areas) and one within-subject factor ([Place], i.e., place of articulation: labial or dorsal), the main effect of [Area] and the interaction between [Area] and [Place] were not statistically significant ([Area]: F (1, 48) = 1.430, p = 0.238, [Area] × [Place]: F (1,1503) = 0.003, p = 0.958. The main factor of [Place] was statistically significant (F (1,1503) = 570.877, $p < 0.05$).

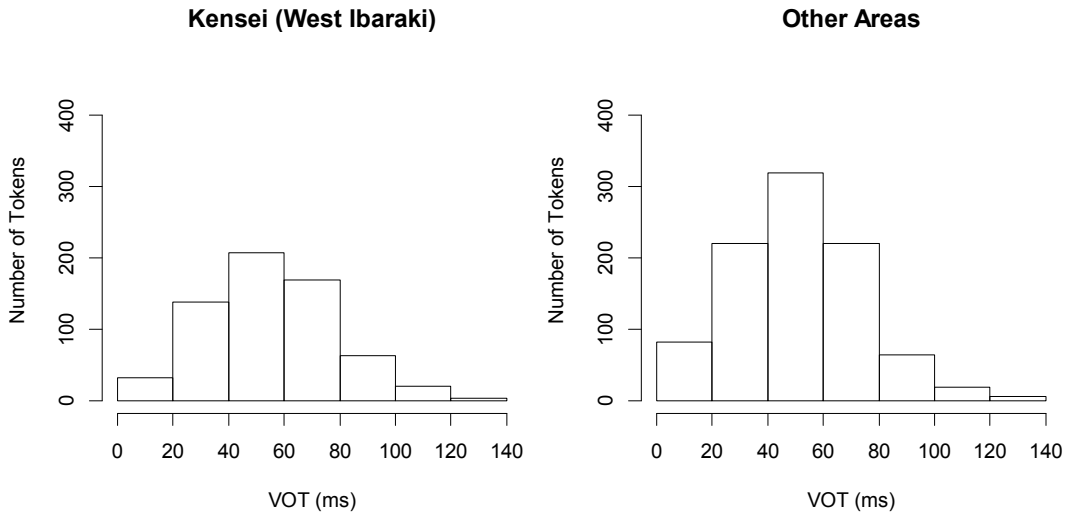

Figure 3: Histograms of VOT for voiceless stops. Left panel: Kensei (West Ibaraki); right panel: other areas.

4. Discussion

Since no difference in voiceless stops was identified (Figure 3), the following discussion focuses only on voiced stops.

Our working hypothesis was that prevoicing is dominant in Kensei, but our results reject this hypothesis. As shown in Figure 1, short-lag is dominant even in Kensei. However, this does not mean that all speakers in Kensei prefer short-lag. The bimodal distribution suggests instead that some speakers here prefer prevoicing.

Figure 2 suggests that prevoicing is dominant in a limited section of Kensei including parts of Koga, Yachiyo, and Joso. We will need more data to determine the exact areas, particularly in cities and towns in south-central Kensei that we have not yet investigated, such as Sakai and Bando.

Some speakers showed intermediate short-lag ratio. Speakers having a ratio between 0.4 and 0.8 were identified not only in Kensei but also in other areas of Ibaraki. It is unclear where the

intermediate speakers are located and whether geography is the only factor accounting for their distribution.

More data from elderly speakers will be required before we develop our study into one focused on generational differences.

Acknowledgements

We would like to thank all participants in this study. We are also grateful to the Kitaibaraki City Council of Social Welfare, Mito Federation of Senior Citizens' Clubs, Miho Village Council of Social Welfare, Senior Citizens Welfare Division of Inashiki City, Shimotsuma City Council of Social Welfare, Yachiyo Town Council of Social Welfare, and Koga City Council of Social Welfare (Sowa Office) for assisting with the recruitment of participants, and Takahide Kawabe, Takashi Otsuka, and Astghik Movsisyan for assisting with data analysis. This work was supported by KAKENHI (No. 24520418 and No. 24320087).

References

Kessinger, R.H. and S.E. Blumstein (1997) Effects of speaking rate on voice-onset time in Thai, French, and English. *Journal of Phonetics* 25, 143-168.

Lisker, L. and A.S. Abramson (1964) A cross-language study of voicing in initial stops: acoustical measurements. *Word* 20, 384-422.

Magloire, J. and K.P. Green (1999) A cross-language comparison of speaking rate effects on the production of voice onset time in English and Spanish. *Phonetica* 56, 158-185.

Miyajima, T. (1961) Hogen no jittai to kyotsugoka no mondaiten 6: Fukushima, Tochigi, Ibaraki. In M. Tojo (ed.) *Hogengaku Koza Vol.2: tobu hogen*, 236-263. Tokyo: Tokyodo Shuppan.

Takada, M. (2006) Goto no yusei haretsuon ni okeru VOT no chiikisa to sedaisa: tohoku~kanto ni tsuite. *Nihongo no Kenkyu* 2 (2), 34-45.

Takada, M. (2011) *Nihongo goto heisaon no kenkyu: VOT no kyojiteki bunpu to tsujiteki henka*. Tokyo: Kurosio.

Takao, M. (2012) VOT no hogensa ni kansuru kosatsu: VOT to ta no onseitokucho tono sokan no kanten kara. BA thesis, Hiroshima University.