

Acoustic cues of the stop voicing contrast in modern Tokyo Japanese *

©Jiayin Gao^{1,2,3}, Takayuki Arai¹

(¹Sophia University, ²JSPS, ³Lacito, LPP, CNRS-Paris 3)

1 Introduction

Among languages with a two-way laryngeal contrast, two types are usually distinguished [1]: (a) “true voicing” languages such as French, and (b) “aspiration” languages such as German. The laryngeal contrast is realized as prevoiced vs. voiceless unaspirated in (a), and as voiceless unaspirated vs. voiceless aspirated in (b).

Japanese is traditionally described as having a true voicing distinction. However, unlike French, voiceless stops in Japanese are slightly aspirated in word-initial position [2]. Voiced stops are reported to have negative VOTs [3], but recent studies suggest a trend toward devoicing word-initial /b, d, g/ in Tokyo Japanese [4].

This study is an acoustic investigation of the two stop series in different prosodic positions in modern Tokyo Japanese (TJP). Section 2 presents our main findings, and Section 3 draws some conclusions about the phonological implications.

2 Experiment

2.1 Method

2.1.1 Participants and speech material

Twelve native speakers (3 males and 9 females) of TJP aged 20 to 27 were recorded.

We report results from 37 (near-) minimal pairs with a voicing contrast for word-initial (31 pairs) or word-medial (6 pairs) labial, dental and velar stop onsets, plus 4 words with /m/ onset, making a total of 78 words. Two pitch-accent contexts were used: accentless (38 words) vs. initial accent (40 words), that is, low vs. high pitch on the first mora. They were elicited in isolation in the first session, and in a carrier sentence [sore o ___ to iu] ‘(I) call it ___’ in the second session. Three prosodic positions were thus compared: (1) utterance-initial (UI, 64 words in isolation); (2) utterance-medial & word-initial (UM, 64 words in a carrier sentence); and (3)

word-medial (WM, 28 words either in isolation or in a carrier sentence). Each utterance was repeated twice, making a total of 312 utterances (=78 words*2 sessions*2 repetitions).

2.1.2 Procedure and analyses

Speakers were recorded individually in a sound-proof room. The two repetitions of the females and the first repetition of the males were analyzed. We focus our analyses on VOT (voice onset time) and F0 (fundamental frequency) as main cues to voicing. Was measured the mean F0 of each of the 20 equal time intervals of the moraic vowel following the target onset.

2.2 Results

2.2.1 VOT

Consistent with previous reports, voiceless stops have medium-lag VOTs word-initially, either utterance-initially or –medially, but short-lag VOTs word-medially, for both genders. Results for female speakers are shown in Table 1. Moreover, VOTs are slightly longer in accentless than initial accent context in average.

Table 1 Mean VOTs in ms (SD) according to POA and prosodic positions (females).

	UI	UM	WM
/p/	38.0 (15.3)	33.5 (17.0)	15.3 (5.8)
/t/	35.3 (13.7)	29.8 (12.0)	14.7 (3.2)
/k/	59.1 (16.6)	49.4 (25.3)	22.8 (7.3)

Concerning voiced stops, the devoiced pattern is mostly observed for female speakers in utterance-initial position. As shown in Fig 1, while all three male speakers (M01-3) produced prevoiced stops more than 60% of the time, this was the case for only one of the nine females. More than half of them produced unvoiced stops more than 75% of the time. As the prosodic hierarchy lowers, the prevoiced proportion increases for females (76% in UM, 99% in WM).

Fig 2 further shows the VOT distributions for utterance-initial stops for female speakers.

* 現代日本語東京方言における閉鎖音の有声性に関する音響的手がかり。

高佳音^{1,2,3}, 荒井隆行¹ (¹上智大学, ²日本学術振興会, ³Lacito, LPP, CNRS-Paris 3)。

Limited to unvoiced realizations, VOTs of /b,d,g/ (6-63 ms) and /p,t,k/ (9-107 ms) overlap, but differ significantly, $\chi^2(1)=54.0, p<.0001$.

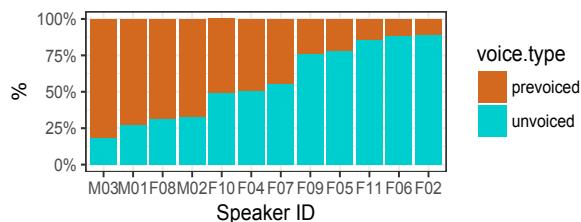


Fig. 1 Proportion of utterance-initial prevoiced vs. unvoiced /b,d,g/ of each speaker.

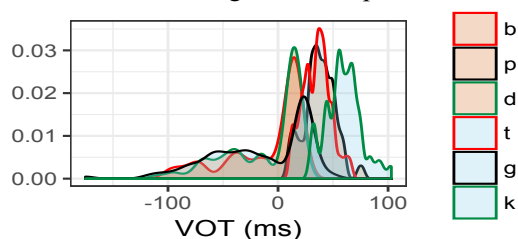


Fig. 2 Kernel density estimation of VOT distributions of utterance-initial stops (females).

2.2.2 F0

The F0 results show a large effect of pitch perturbation on the moraic vowel following word-initial onsets, either utterance-initially (Fig. 3 for females) or –medially. F0 is higher following /p,t,k/ than /b,d,g/ (prevoiced or short-lag) and /m/, not only at vowel’s onset, but lasting till its final part. This effect is larger in initial accent than accentless context, and seems more consistent for female than male speakers.

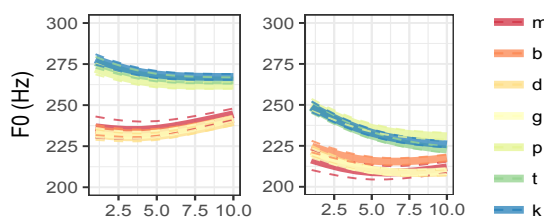


Fig. 3 Time-normalized F0 curves in Hz over the moraic vowel following utterance-initial onsets, in initial accent (left) and accentless (right) contexts (females), computed by SSANOVA [5].

3 Conclusions

In line with [4], our results showed a trend toward devoicing utterance-initial voiced stops in TJP, and female speakers again take the leading role in this sound change. More speakers need to be recorded to confirm this. However, this sound

change has not lead to a merger because the two stop series are distinguishable by VOT. Whether the VOT cue is gradually replaced with an F0 cue on the following vowel needs more investigation. Limited to VOT, [6] argues that in “aspiration” languages, utterance-initial “devoicing” suggests [voice] is not an active feature, and that voicing in utterance-medial position is *passive voicing*. Adopting this view, modern TJP is not a “true voicing” language. On the contrary, phonological rules, such as *Rendaku*, are in favor of a “true voicing” account for Japanese [7]. This might suggest that phonetic cues and phonologically active features do not necessarily coincide.

Acknowledgements

We want to thank Prof. Mafuyu Kitahara for his help in the selection of the speech material. This ongoing project is funded by JSPS (Grant-in-Aid Number: 17F17006).

References

- [1] G. K. Iverson and J. C. Salmons. “Aspiration and Laryngeal Representation in Germanic,” *Phonology*, **12**(3), 369-396 (1995).
- [2] T. J. Riney, N. Takagi, K. Ota, and Y. Uchida. “The intermediate degree of VOT in Japanese initial voiceless stops,” *JPhon*, **35**(3), 439-443 (2007).
- [3] K. Shimizu. A cross-language study of voicing contrasts of stop consonants in six Asian languages. Tokyo: Seibido, (1996).
- [4] M. Takada, E. J. Kong, K. Yoneyama, and M. E. Beckman. “Loss of prevoicing in modern Japanese /g, d, b/,” *Proc. ICPHS 2015, paper N. 873*, 1-5 (2015).
- [5] C. Gu. *Smoothing Spline ANOVA Models*, New York: Springer-Verlag, (2002).
- [6] J. Beckman, M. Jessen, and C. Ringen. “Empirical evidence for laryngeal features: aspirating vs. true voice languages,” *J. Linguistics*, **49**, 295-284 (2013).
- [7] K. Nasukawa. “The representation of laryngeal-source contrasts in Japanese,” de Weijer et al. (eds), *Voicing in Japanese*, 71-87 (2005).