Acoustic covariants of the Japanese geminate/singleton contrast and their relative perception

Kaori Idemaru, University of Oregon

The temporal acoustics of the speech signal are strongly influenced by speaking rate (e.g., Miller 1981). As speaking rate varies, segmental durations may be stretched or reduced in time. The temporal acoustic correlates of voicing, short versus long vowels, and singleton versus geminate stops are influenced by speaking rate such that the acoustics of contrasting categories may overlap substantially across different speaking rates (e.g., Miller & Liberman 1979). Therefore, in natural speech where rate changes both within and across speakers, absolute duration alone may not provide reliable information to categorize speech sounds distinguished on the basis of temporal differences. There is no invariant acoustic cue with which to determine speech category membership.

Japanese singleton vs. geminate stop length contrast is an example where category overlap has been reported when variability due to speech rate is considered (Fujisaki 1979, Hirata 1990). Thus Japanese listeners may have some means of adjusting perception for rate-induced variability to correctly categorize singleton and geminate stops. In general, characterization of perceptual adjustment bas been pursued by work on durational context effects, rate normalization and relative timing, all focusing on the effect of context (e.g., Newman & Sawusch 1996, Wade & Holt 2005, Port & Dalby 1982). Finding from these studies converge on the point that the immediately adjacent segment(s) exerts an important influence in the judgment of target segmental length, and studies on Japanese singleton vs. geminate stop length contrast has provided insights into this problem from different angles (Ofuka et al. 2005, Kingston et al. 2009, Hirata & Whiton 2005, Amano & Hirata 2010).

Whereas the findings by the prior work indicate the importance of durational context effect and relative perception of Japanese stop length, Ham (2001) reports an intriguing cross-linguistic difference in the degree with which singleton versus geminate stops are differentiated by stop duration, which may have a bearing on the role of relative perception. According to Ham, Japanese shows a relatively robust differentiation of singleton and geminate categories by stop duration. Given such robust differentiation, a question arises as to what degree Japanese listeners rely on relative perception versus absolute perception in making judgment about stop length. The current study adds to the accumulating knowledge on the topic of perception of Japanese stop length by exploring potential secondary cues among acoustic features that covary with the stop length distinction and further investigating perceptual strategies Japanese listeners employ in judging singleton versus geminate length of stop consonants.

This work presents four experiments. Study 1 investigates potential secondary acoustic and perceptual cues to the singleton vs. geminate stop length. Study 2 explores various relative timing by conducting a perception experiment. Study 3 further investigates perceptual strategies by examining listeners' reliance on relative timing and absolute stop duration. Study 4 examines the link between perceptual and production strategies in distinguishing singleton versus geminate stops. The results indicate that (1) for Japanese stop length contrast, multiple acoustic cues are available; (2) although these secondary cues are robust, listeners rely exclusively on the stop duration cue; and (3) Japanese listeners exploit the characteristics of the signal utilizing both relative timing and absolute duration cues. Therefore, despite the theoretical advantage of relative timing, it is not necessarily utilized by all Japanese listeners. These results indicate that the role of relative versus absolute acoustic dimensions in speech perception must be examined in relation to the distributional regularities of the acoustic dimensions in the language.

References

Amano, S., & Hirata, Y. (2010). Perception and production boundaries between single and geminate stops in Japanese. *JASA*, *128*(4), 2049-2058.

Fujisaki, H. (1979). On the Modes and Mechanisms of Speech Perception: Analysis and Interpretation of Categorical Effects in Discrimination. In B. Lindblom & S. E. G. Ohman (Eds.), *Frontiers of Speech Communication Research* (pp. xxv, 393): Academic, New York Pagination: 177- 189.

Hirata, Y. (1990). Perception of geminated stops in Japanese word and sentence levels by English speaking learners of Japanese language. *Onseigaku Kaihoo, 195*, 4-10.

Hirata, Y., & Whiton, J. (2005). Relational acoustic invariance in the single/geminate stop distinction in Japanese. *JASA, 117*(4), 2569.

Idemaru, K. & Guion, S. G. (2008). Acoustic covariants of length contrast in Japanese stops. *JIPA, 38*(2), 167-186.

Idemaru, K. & Guion-Anderson, S. G. (2010). Relational timing in the production of singleton and geminate stops in Japanese. *Phonetica*, 67(1-2), 25-46.

Idemaru, K., Holt, L., & Seltman, H. (under review). Cue weighting in the perception and production of Japanese singleton and geminate stops.

Kingston, J., Kawahara, S., Chambless, D., Mash, D., & Brenner-Alsop, E. (2009). Contextual effects on the perception of duration. J. Phonetics, 37(3), 297–320.

Miller, J. L. (1981). Effects of Speaking Rates. In P. D. Eimas & J. L. Miller (Eds.), *Perspectives on the study of speech* (pp. 39-74). Hillsdale, NJ: Lawrence Erlbaum.

Miller, J. L., & Liberman, A. M. (1979). "Some effects of later-occurring information on the perception of stop consonant and semivowel," Percept. Psychophys., 25(6), 457-65.

Newman, R. S., & Sawusch, J. R. (1996). Perceptual normalization for speaking rate: Effects of temporal distance. Percept. Psychophys., 58(4), 540–560.

Ofuka, E., Mori, Y., & Kiritani, S. (2005). Sokuon no chikakuni kansuru senkoo kouzoku boin cho no eikyoo [The effects of the duration of preceding and following vowels on the perception of geminates]. J. Phonet. Soc. Japan, 9, 59-65.

Port, R. F., & Dalby, J. (1982). Consonant/vowel ratio as a cue for voicing in English. *Perception & Psychophysics, 32*(2), 141-152.

Wade, T., & Holt, L. L. (2005). Perceptual effects of preceding nonspeech rate on temporal properties of speech categories. Percept. Psychophys., *67*(6), 939–950.