

On measuring prominence

Emil Tanev

Graduate School of Language and Culture, Osaka University,

The present work addresses the problem of measuring intonational prominence in spontaneous Japanese speech. It has been shown that prominence is realized through pitch range expansion on the item in question, although the exact definition of pitch range differs, according to model of intonation adopted. Pierrehumbert & Beckman (1988) posit a speaker specific reference line, which the pitch range is calculated from, while Fujisaki's accent command is superimposed on a declining baseline (1981, 2004). While the former approach is fairly straightforward in theory, it encounters difficulties when applied in connected speech (Venditti 2000) due to the confounding effects of downstep and declination. The disadvantage of the latter model is that it is based on an analysis-by-synthesis procedure, which does not guarantee descriptive consistency.

It is suggested here that prominence may be realized through changes in the speed of pitch excursion (S_{pe}) as well. If this is indeed true, listeners should be sensitive to S_{pe} . That is, if the assumption that discrete gestures on the part of the speaker are recovered as such by the listener ('tHart et al. 1990) is correct, then by studying the perceptual process new insights on the production process could be gained. In order to verify the relevance of S_{pe} , a perceptual experiment using synthetic stimuli was conducted. Thirteen native speakers of Japanese listened to twelve target pairs in which the pitch range was kept constant, and the S_{pe} was varied – an early peak and a standard peak. The IPO approach to intonation was adopted in the manipulation ('tHart et al. 1990). Non-parametric statistics indicated that prominence was significantly influenced by S_{pe} , that is higher S_{pe} led to more judgments of prominence. The difference in pitch range excursion (initial rise) was below the perceptually relevant limit of 1st and thus could not have influenced the judgments. A number of tokens were difficult to disambiguate, and a possible explanation is the small difference in S_{pe} combined with the confounding effect of accentuation. Finally, it is suggested that the difference in emotionality between late rises and an early rise, suggested by Kawakami (1995), could be explained not by variation in the amount of pitch excursion but by the speed of it. The latter suggestion is yet another confirmation of the relevance of S_{pe} for prominence.

References:

- Fujisaki, H. (1981). Dynamic characteristics of voice fundamental frequency in speech and singing. Acoustical analysis and physiological interpretations. *Speech, Music and Hearing Quarterly Progress and Status Report*, 22 (1), 1-20.
- Fujisaki, H. (2004). Information, Prosody, and Modeling. *Proceedings of Speech Prosody 2004*, Nara, Japan.
- Hart, J., Collier, R. & Cohen, A. (1990). *A perceptual study of intonation: an experimental-phonetic approach to speech melody*. Cambridge: Cambridge University Press.
- Kawakami, S. (1995). "Buntoo no intoneeshon." *Nihongo akusento ronshuu*. Tokyo: Kyuuko shoin.
- Pierrehumbert, J. & Beckman, M. (1988). *Japanese Tone Structure, Linguistic Inquiry Monograph 15*.

Cambridge, MA: MIT Press.

Venditti, J. (2000). Discourse structure and attentional salience effects on Japanese intonation. Ph.D. dissertation, Ohio State University.