Masahiko Komatsu* and Takayuki Arai°
*Sophia University, Japan, &
University of Alberta; °Sophia University

Friday
9:20

Acoustic Realization of Prosodic Types: Constructing Average Syllables

Languages are categorized in terms of accent (stress accent, pitch accent, and tone accent) and rhythm (stress-timed, syllable-timed, and mora-timed). Although there have been comparative studies of specific languages, they have not offered the whole picture of prosody including both accent and rhythm in terms of purely acoustic parameters, or the "acoustic map of prosody."

In the present study, we seek for acoustic measures quantitatively representing the accent and rhythm types, investigating samples from English, Japanese, Chinese, and Spanish. We use F0 (fundamental frequency) and intensity as in previous studies, and propose the amplitude contours of harmonics and noise as new measurements, which represent "vocalic" and "consonantal" properties of sound, respectively. While we have confirmed some global characteristics predicted from previous studies, we also emphasize the importance of local prosodic features.

We examined speech and words in the four languages, including English and Japanese speech from the same speakers. We measured the contours of F0, amplitude, and HNR (Harmonics-to-Noise Ratio) and obtained the amounts of harmonics and noise from amplitude and HNR. From these values, we calculated several parameters for the comparison among languages.

In the long-term analysis of speech, Chinese showed a greater rate of F0 change than the other languages, and English showed more frequent F0 fluctuations than Japanese. This is thought to be a reflection of different accent types, and is consistent with the comparison of English and Chinese by Eady (1982). The mean HNR was larger in Japanese than in English. We regard this as the indication that the proportion of vowel intervals in Japanese speech is larger than in English, which creates the difference in rhythm (Ramus, et al., 1999).

Some local characteristics, however, seem to have been smeared in the long-term analysis. For example, some speakers showed interesting patterns of the F0-amplitude interaction in the word pronunciation, but such characteristics did not appear in the results of long-term analyses.

We calculated an "average syllable" of each language, which has the averaged contours of amplitude, harmonics, noise, and F0. Average syllables bear the local prosodic characteristics, showing some cross-linguistic difference, as well as global features common to human speech.

We extended Eady's F0 analysis on stress and tone languages to other prosodic types, and proposed a more acoustic measure than the rhythm analysis based on phonological segmentation by Ramus et al. We also introduced an analysis method of local prosodic features, the importance of which is suggested by Thymé-Gobbel and Hutchins (1996).

Eady, Stephen J. 1982. Differences in the F_0 patterns of speech: Tone language versus stress language. *Language and Speech* 25: 29-42.

Ramus, Franck, Marina Nespor & Jacques Mehler. 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 73: 265-92.

Thymé-Gobbel, Ann E. & Sandra E. Hutchins. 1996. On using prosodic cues in automatic language identification. *Proc. of International Conference on Spoken Language Processing 96*, 1768-71.