

Production and perception of lexical accent in Japanese

Word prosody of Tokyo Japanese is often labeled as pitch accent, characterized by a steep F_0 (= fundamental frequency) **fall** from the accented mora to the following one (e.g. McCawley, 1968). This study examined the acoustics and perception of Japanese pitch accent using twenty minimal pairs that differed only in accent. For example, /hana/ with a low-high (LH) tone sequence means ‘flower’ when the final mora is accented and ‘nose’ when there is no accent. The traditional account of Japanese pitch accent claimed that the only difference between the two accent types is the pitch on the following particle (e.g. Kindaichi, 1947). It has a low pitch after a final-accented word (thus, /hana^ˆ ga/ LH L ‘flower NOMINATIVE’, ‘^ˆ’ indicates accent on the preceding mora) whereas it has a high pitch after an unaccented word (/hana ga/ LH H ‘nose NOMINATIVE’). However, recent experimental studies suggest that final-accented and unaccented words are different even within words, with final-accented words showing a higher F_0 **peak** on the second mora and a greater F_0 **rise** from the first to second morae than unaccented words (e.g. Pierrehumbert and Beckman, 1988; Vance, 1995). But their results were based on only a couple of (near-)minimal pairs and they were inconsistent as to any difference between final-accented and unaccented words when they were produced in isolation. Additionally, past research on Japanese pitch accent studied production and perception separately. Given the communicative function of speech, it is important to understand the relation between them.

The present study was designed to overcome these problems. First, a computerized database (Amano & Kondo, 1999) was used to search for all bimoraic minimal pairs existing in Japanese. Bimoraic words were necessary because F_0 rise cannot be measured with monomoraic words and minimal pairs are virtually nonexistent for trimoraic or longer words (Kitahara, 2001). Second, because word familiarity is known to influence word recognition and production (Amano, Kondo, & Kato, 1999; Wright, 1997), only words that had a relatively high familiarity rating in the database were used, resulting in 20 minimal pairs. Ten Tokyo Japanese speakers produced the 20 pairs in isolation and sentence-medially followed by a particle.

The production study found that the two accent types differed significantly within words when they were produced sentence-medially, as evidenced by a greater F_0 peak and F_0 rise as well as by a greater F_0 fall for final-accented words than unaccented words ($p = .0002$, $p < .0001$, $p = .0001$, respectively). However, the two types of words showed no significant difference when they were produced in isolation (F_0 peak: $p = .3330$; F_0 rise: $p = .6817$). The perception experiment used recordings of one male and female collected in the production study to examine if acoustic differences found for final-accented and unaccented words produced sentence-medially can be used by listeners to identify words. The listeners ($n=23$) were not able to identify words reliably when the words produced in a sentence were excised and presented to them in isolation ($p = .2411$). Thus even though the two types of words differed in F_0 peak and F_0 rise, these acoustic properties were not sufficient for the listeners to identify words. Furthermore, word identification was only about 80 percent accurate even when the words produced in a sentence were excised with the following particle. However, a strong positive correlation was found between the accuracy and the size of F_0 fall difference maintained between the final-accented and unaccented words in each pair (male voice: $r = .750$, $p = .0035$; female voice: $r = .917$, $p < .0001$), indicating that F_0 fall played an important role in perceiving accent. Results of the two experiments show a mismatch between how talkers embed accent information and the cues that listeners use to perceive accent. In other words, while accent information was redundantly present in F_0 peak, F_0 rise, and F_0 fall differences, listeners only used F_0 fall to distinguish the accent type.

References

- Amano, S., & Kondo, K. (1999). *Nihongo-no goitokusei* [Lexical Properties of Japanese]. Tokyo: Sanseido.
- Amano, S., Kondo, T., & Kato, K. (1999). Familiarity effect on spoken word recognition in Japanese. *Proceedings of the 14th ICPhS*, 2, 873-876.
- Kindaichi, H. (1947). Tookyoogo ni okeru “hana” to “hana” no kubetsu – tookyoo akusento shinnidankan kyoochooron. *Kokugo [The Japanese language]*, Summer.
- Kitahara, M. (2001). *Category structure and function of pitch accent in Tokyo Japanese*. Unpublished doctoral dissertation, Indiana University.
- McCawley, J. (1968). *The phonological component of a grammar of Japanese*. The Hague: Mouton.
- Pierrehumbert, J., & Beckman, M. (1988). *Japanese tone structure*. Cambridge, MA: MIT Press.
- Vance, T. J. (1995). Final accent vs. no accent: utterance-final neutralization in Tokyo Japanese. *Journal of Phonetics*, 23, 487-499.
- Wright, R. (1997). *Lexical competition and reduction in speech: A preliminary report* (Progress report No. 21). Bloomington, IN: Speech Research Laboratory.