Accent-meter alignment in Japanese vocal music

Sunghye Cho, University of Pennsylvania

There is an ongoing debate about the nature of pitch accents. Even in Tokyo Japanese (TJ), which is the most widely cited pitch-accent language, the pitch accent is analyzed in two distinct ways: an accentual approach and a tonal one ([2, 3, 4, 7], among others). In the accentual approach (1a), an accent marks the position of a pitch drop from High (H) to Low (L), whereas in the tonal approach (1b), a HL melody is aligned with the mora to be accented in (1a). (In (1), accented syllables are marked with an acute diacritic and H tones are represented with capital letters. 'edge-NOM' is accentless due to lack of a pitch drop.)

(1)	a. Accentual:	há.si-ga	ha.sí-ga	ha.si-ga
	b. Tonal:	HA.si-ga	ha.SI-ga	ha.SI-GA
		'chopstick-NOM'	'bridge-NOM'	'edge-NOM'

The accentual approach (1a) predicts that an accented syllable has a phonological (abstract) prominence similar to stress in English. Our research question is whether a pitch accent in TJ marks a phonological prominence. If it does, this would suggest that the accentual analysis is on the right track. One way to answer this question is to examine how Japanese lyrics align with musical tunes. It is well known that stressed syllables are likely to align with strong beats of the meter in vocal music of stress languages, such as English and German [1, 5]. If an accent in TJ is a phonological prominence similar to stress, accented morae will tend to fall on strong beats of the meter. In this study, we examine this question with traditional Japanese songs.

The songs used in our study are all drawn from a single collection, *Nihonno Warabeuta: Sitsunai Yuugika Hen* [6]. The book contains a large number of traditional children's songs from various areas of Japan. We selected 27 songs sung in the Tokyo area. Adopting the methodology in [8], we classified every syllable in our corpus in three ways: its prosodic type, the accentual level of the note, and its syllabic weight. For prosodic types, accentedness was marked with *u* (unaccented) and *a* (accented), content words were coded as *c,* and function words were labeled as *f*, creating four prosodic types in total: *ac, af, uc,* and *uf.* For syllabic weights, heavy syllables were labeled as *h.* If a coda consonant of a heavy syllable was assigned to a different note from the nucleus, both the coda and the nucleus were coded as *l* separately*,* and light syllables were also labeled as *l.* For musical labeling, the tactus level was defined as level 2 (the quarter-note level in 2/4, 3/4, and 4/4 time signatures), and one level above the tactus was level 3 (the one-measure level in 2/4 and 3/4, and the half-measure level in 4/4). The one-measure level in 4/4 was defined as level 4, and all beats below the tactus were level 1 (Fig. 1).

For statistical analysis, we first examined whether accented words are more likely to fall on strong beats than accentless words. For this comparison, two songs were excluded, because they only contained accented words. We computed the mean accentual strength of accented and unaccented words for each song and conducted a paired *t*-test across songs. The result of a *t*-test showed that there is no significant difference in accentual strength between accented (mean = 2.22) and unaccented words (mean = 2.204), $t(24) = -0.29$, $p = 0.77$.

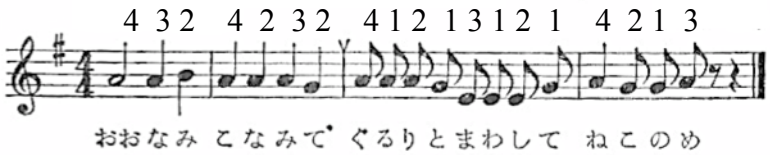

Fig. 1: *Oonami Konami* [6], showing the encoding of accentual levels.

The next question we examined was whether accented syllables are more likely to fall on strong beats than unaccented ones in accented words. For this comparison, we excluded accentless words, which have no accented syllable. The mean values of accented syllables are higher than those of unaccented ones in both content and function words (Fig.2). Also, a paired *t*-test reveals that there is a significant difference in accentual strength between accented and unaccented syllables, $t(26) = 4.3$, $p = 0.0002$, suggesting that accented syllables (mean = 2.503) are more likely to fall on strong beats than unaccented ones (mean = 1.98).

The third question we addressed was whether syllable weight plays a role in the meter alignment. Fig. 3 shows the mean values of all syllables by their syllable weight. The result suggests that heavy syllables (mean = 2.68) are more likely to fall on strong beats than light syllables (mean = 2.15). We conducted a paired *t*-test with the mean values of heavy and light syllables across songs, excluding four songs which have no heavy syllables. The result reveals that the difference between heavy and light syllables is significant ($t(22) = 3.84$, $p = 0.0008$).

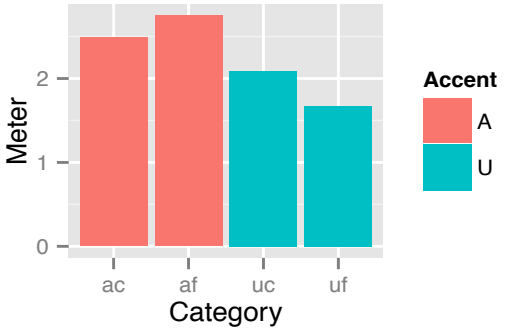

Fig. 2: Mean accentual strength by prosodic type

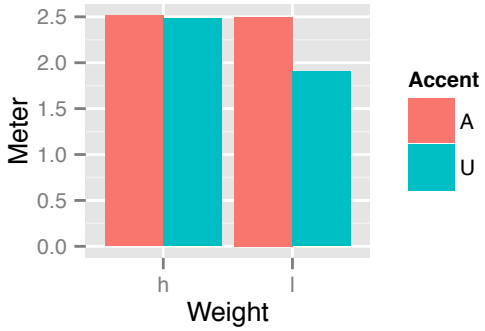

Fig. 3: Mean accentual strength by syllabic weight

Lastly, we built a linear mixed-effects model to examine the effect of the accentual levels and syllable weights on the musical meter with accented light syllables as the reference category and songs as a random effect (Fig. 3). The model shows a significant effect of Accent ($p < 0.0001$), indicating that the accentual strength of unaccented light syllables is 0.61 lower than that of accented light syllables. However, syllable weight was not a significant factor among accented syllables ($p = 0.9$). That is, whether heavy or not, accented syllables tend to fall on strong beats. The interaction of the two factors is significant, meaning that the estimated accentual strength of accented heavy syllables is 0.71 higher than that of unaccented light syllables ($p = 0.004$).

The implication of our study is that it provides evidence in favor of the accentual approach to the tonal approach in TJ. We find that a pitch accent seems to be a phonological prominence in TJ. Accented syllables, regardless of their syllabic weight, tend to fall on strong beats, whereas for unaccented syllables, syllabic weight plays a role in the meter alignment. Our full paper also compares our results of Japanese meter alignment to those of English and French reported in [8].

References: [1] Halle, J. & F. Lerdahl. 1993. A generative textsetting model. *Current Musicology* 55, 3-23. [2] Hirayama, T. 1960. *Zenkoku Akusento Ziten*. Tokyo: Meiji Shoin. [3] Hyman, L. 2009. How (not) to do phonological typology: the case of pitch-accent. *Language Science* 31, 213-238. [4] Kubozono, H. 2012. Varieties of pitch accent systems in Japanese. *Lingua* 122, 1395-1414. [5] Palmer, C. & M. Kelly. 1992. Linguistic prosody and musical meter in song. *Journal of Memory and Language* 31, 525-542. [6] Obara, A. 1932. *Nihonno Warabeuta: Sitsunai Yuugika Hen.* Tokyo: Shohan. [7] Poser, W. J. 1984. The phonetics and phonology of tone and intonation in Japanese. Doctoral Dissertation, MIT. [8] Temperley, N. & D. Temperley. 2013. Stress-meter alignment in French vocal music. *Journal of Acoustical Society of America* 134, 520-527.