Accent-meter/tune 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), tones are directly aligned with morae without the notion of accent. (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.si-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, whereas the tonal approach expects that the position claimed to be accented is not prominent. My 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. If not, the tonal approach is a better analysis for TJ. 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 occur at strong beats of the meter in vocal music of stress languages, such as English [1, 5]. Also, previous studies show that in tonal languages there is a mapping between tonal transitions and musical note transitions [9, 10]. If an accent in TJ is a phonological prominence similar to stress, accented morae will tend to occur at strong beats of the meter and a falling transition is aligned with the accented mora. On the other hand, if TJ is a tonal language with a restricted tonal system, accented morae won’t be necessarily aligned with strong beats. In this study, I examine this question with traditional Japanese songs.

The songs in this study are all drawn from a single collection, Nihonno Warabeta: Sitsunai Yuugika Hen [6]. The book contains a large number of traditional children’s songs from various areas of Japan. I selected 27 songs sung in the Tokyo area. Adopting the methods in [8], I classified every syllable in our corpus in four ways: its accent type, metrical strength, syllable weight, and musical note transition (from the given mora to the next one). For accent 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 accent types in total: \( ac, af, uc, \) and \( uf \). For syllabic weights, heavy syllables were labeled as \( h \) and light syllables were also labeled as \( l \). For metrical strengths, the tactus level was defined as accentual strength 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. For musical note transitions, we classified the direction of musical note transitions into three groups: rising, falling, and level.

For statistical analysis, I first examined whether accented syllables are more likely to fall on strong beats than unaccented ones in accented words. The mean values of accented syllables are higher than those of unaccented ones in both content and function words (Fig. 1). I computed the mean accentual strength of accented and unaccented words for each song and conducted a paired \( t \)-test across songs. The result reveals that there is a significant difference in metrical strength between accented and unaccented syllables, \( t(26) = 4.3, p = 0.0002 \), suggesting that accented syllables (Mean (M) = 2.5) occur at stronger beats than unaccented ones (M = 1.98).
The next question I addressed was whether syllable weight plays a role in the meter alignment. The result shows that heavy syllables \((M = 2.68)\) occur at stronger beats than light syllables \((M = 2.15)\). I built a linear mixed-effects model to examine the effect of accent and syllable weights on the metrical level with accented light syllables as the reference category and songs as a random effect (Fig. 2). The model shows a significant effect of Accent \((p < 0.001)\), indicating that the metrical 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 third question I examined was whether accented morae are likely to be aligned with a falling transition. The result shows that accented morae in both content and function words are aligned with a falling transition about 60\% of the time, whereas unaccented ones are aligned with a falling transition about 20\% of the time (Fig. 3). I conducted a Wilcoxon test across songs, and the test reveals that the difference between accented and unaccented morae is significant \((p < 0.001)\).

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 prominence in TJ. Accented syllables, regardless of their syllabic weight, tend to fall on strong beats, and accented syllables are more likely to be aligned with a falling transition than unaccented ones.