Cross-dialectal perspective on the relation between form and meaning: the case of Tone 3 sandhi in spoken Beijing and Taiwan Mandarin

Yuxin Lu[1], Yu-Hsiang Tseng[1], and R.Harald Baayen[1]
[1]Quantitative Linguistics, University of Tübingen

One of the most well-known tone sandhi rules in Mandarin is the Tone 3 sandhi. When a Tone 3 (T3, dipping tone) syllable is followed by another T3 syllable, the first T3 is typically pronounced as a Tone 2 (T2, rising tone). For example, disyllabic words with a T3-T3 tone pattern, such as 水果 (*shui3guo3*, 'fruits'), are phonetically realized as T2-T3 (*shui2guo3*). However, the extent to which this T3-T3 pattern is neutralized to T2-T3 varies across different Mandarin varieties.

This study investigates the realization of Tone 3 sandhi in spontaneous speech by comparing Beijing Mandarin and Taiwan Mandarin. We selected 26 disyllabic word types with T2-T3 or T3-T3 tone patterns that were present in both a spoken Taiwan Mandarin corpus and a spoken Beijing Mandarin corpus. We used Generalized Additive Mixed Models (GAMMs), which allowed us to model pitch contours over time as a function of predictors such as normalized time, gender, word, duration, word position in the utterance, neighbouring tones, and speaker.

The statistical analyses revealed an effect of gender when modelling the difference between T2-T3 and T3-T3. In Beijing Mandarin, the difference between T2-T3 and T3-T3 tone patterns was significant for male speakers but not for female speakers. In contrast, Taiwan Mandarin exhibited the opposite pattern: the difference was significant for female speakers, but not for male speakers. This finding contrasts with previous research based on carefully controlled laboratory speech, which has typically reported incomplete neutralization. This discrepancy may be due, in part, to differences in speech register, as our data were drawn from spontaneous speech. Another reason for dialectal differences is that previous studies did not take into account the fact that pitch contours in Mandarin are often modulated in word-specific ways. Without controlling for the effect of word, the tone sandhi in our data also appeared incomplete across all speaker groups.

To deepen our understanding of how pitch contours may be driven by semantics, we obtained semantic vectors presented by 768-dimensional contextualized embeddings from the GPT-2 large language model. We found that word pairs exhibiting distinct pitch contours in the GAMM analysis also had dissimilar semantic representations. Furthermore, we made use of the Discriminative Lexicon Model (DLM), a computational model focusing on the relationship between form and meaning. The DLM was able to predict the pitch contours of word types from their meanings with an accuracy of 82.7%. However, when we permuted the semantic vectors between Beijing and Taiwan Mandarin, the model's prediction accuracy dropped to 80.8%. This decline indicates that dialectal variation influences the mapping between tonal realization and semantics.

In summary, there are gender and dialectal effects in the realization of Tone 3 sandhi in spontaneous Mandarin, pointing to a complex and nuanced interplay between tonal realization and semantics.