Modelling morphological variation in the realization of pitch contours in spontaneous spoken Taiwan Mandarin

Yuxin Lu, Xiaoyun Jin & R. Harald Baayen
University of Tübingen
yuxin.lu@uni-tuebingen.de, xiaoyun.jin@ uni-tuebingen.de,
harald.baayen@uni-tuebingen.de

A growing body of studies has pointed out that fine phonetic detail in how words are actually spoken reflects differences in meaning (Gahl &Baayen 2024). In Mandarin, Chuang et al. (2024) found that the tonal realization of disyllabic words with T2-T4 tone pattern is in part driven by their context-sensitive meaningss. We investigated the realization of pitch contours of Mandarin disyllabic compounds in a corpus of conversational Taiwan Mandarin (Fon 2004). We made use of Generalized Additive Mixed Models (GAMs) to model f0 as a function of normalized time and a range of other predictors, including word duration, speech rate, segmental properties, tone sandhi, word position, and word sense.

In GAMs analysis, for monosyllabic words, the effect of tone sandhi on pitch contours is similar in magnitude to the effect of word meaning. By contrast, for the compounds, the effect of word meaning on pitch contours is substantially larger than the effect of tone sandhi. Besides, for simple words, the pitch contours revealed by GAMs for the different tone are level tones, whereas the pitch contours for tone patterns of compounds vary substantially with time. What's more, we fitted an Linear Discriminative Learning model to the compounds, mapping contextualized embeddings onto pitch contours. The mean accuracy was 23.5% on training dataset and 15.1% on testing dataset (permutation baseline 0.004). To obtain the pitch contours predicted for the tone pattern, we gave the centorids of the contextutalized embeddings for each of 20 tone patterns as input to the linear mapping. The resulting predicted contours are remarkably similar to the GAM-derived contours (average Eculidean distance 0.15).

These results indicate that the pitch contours of Mandarin tone patterns reflect the prototypical meaning of the compound words sharing these tone patterns. Apparently, average morphological form (pitch) reflects average morphological meaning (embedding centroids).

References. Chuang, Y.-Y., M. J. Bell, Y.-H. Tseng & R. H. Baayen (2024). Word-specific tonal realizations in Mandarin. *arXiv* preprint arXiv:2405.07006. • Fon, J. (2004). A preliminary construction of Taiwan Southern Min spontaneous speech corpus. *Technical report, Tech. Rep.* NSC-92-2411-H-003-050, National Science Council, Taiwan. • Gahl, S. & R. H. Baayen (2024). Time and thyme again: connecting English spoken word duration to models of the mental lexicon. *Language* (accepted).