DO BEAT GESTURES INFLUENCE AUDIOVISUAL LEXICAL TONE PERCEPTION IN MANDARIN?

Yitian Hong (The Hong Kong Polytechnic University, Hong Kong, SAR China; Donders Institute for Brain, Cognition and Behavior, Radboud University, Nijmegen, The Netherlands), Patrick Louis Rohrer (Donders Institute for Brain, Cognition and Behavior, Radboud University, Nijmegen, The Netherlands) & Hans Rutger Bosker (Donders Institute for Brain, Cognition and Behavior, Radboud University, Nijmegen, The Netherlands)
ytian.hong@connect.polyu.hk

The classic McGurk effect, in which the presence of incongruent lip movements alters listeners' speech perception, highlights the importance of visual information in speech processing. Beyond articulatory visual signals, hand gestures also shape speech perception. For example, the timing of a simple downward beat gesture guides the perception of lexical stress in Dutch (Bosker & Peeters, 2021). However, no study to our knowledge has tested the effect of beat gestures on perception in tone languages. Kinematic studies showed that moving one's hands caused fluctuations in f0 through biomechanical coupling, with f0 peaking around moments of maximum hand deceleration (Pouw et al., 2020). Perhaps listeners use this information to guide audiovisual tone perception. We know Mandarin listeners use contextual information to compensate for expected f0 rises in a *contrastive* manner (Huang & Holt, 2009). That is, the same f0 contour is perceived as relatively low-pitched when it was expected to be high, but as relatively high-pitched when listeners expected it to be low (Huang & Holt, 2009). Likewise, listeners may compensate for f0 rising due to co-speech beat gestures.

In this study, we test whether beat gestures influence lexical tone perception in Mandarin. Forty-eight participants will be presented with seven-step lexical tone continua on nine Mandarin word pairs, ranging from tone 1 (high level tone; T1) to tone 2 (rising tone; T2). These continua will be combined with videos of a speaker either performing a simple downward beat gesture (gesture condition) or with hands resting on both sides of her thighs (no gesture condition). The speaker's face will be masked to eliminate any visual articulatory information. Then, the participants will be asked to indicate which Mandarin word they perceive, the word bearing T1 or the word bearing T2.

If listeners indeed expect the speaker to produce higher f0 when producing a beat gesture, they may compensate for this visual context by giving more T2 responses (compared to no gesture condition). Data collection is underway and results will be available in summer 2023. If our hypothesis is confirmed, this would provide the first evidence of perceptual compensation for biomechanical gesture-speech coupling. Thus, we aim to contribute to cross-linguistic evidence of how the timing of simple hand gestures shapes speech perception.

Reference

Bosker, H. R., & Peeters, D. (2021). Beat gestures influence which speech sounds you hear. *Proc. R. Soc. B*, *288*(1943), 1–9. https://doi.org/10.1098/rspb.2020.2419

Huang, J., & Holt, L. L. (2009). General perceptual contributions to lexical tone normalization. *JASA*, *125*(6), 3983–3994. https://doi.org/10.1121/1.3125342

Pouw, W., Harrison, S. J., & Dixon, J. A. (2020). Gesture–speech physics: The biomechanical basis for the emergence of gesture–speech synchrony. *JEP: General*, *149*(2), 391–404. https://doi.org/10.1037/xge0000646