Let's Focus on Gesture and Prosody –
Are they better together and deficient apart?

Natascha Schuldes, Alina Gregori & Frank Kügler
Goethe Universität Frankfurt am Main
{schuldes, gregori}@lingua.uni-frankfurt.de, kuegler@em.uni-frankfurt.de

Recent research highlights the interplay between gesture and speech with pioneering work by Kendon (1980) and McNeill (1992) emphazising the integral relationship between them and the need to consider both for a comprehensive understanding of a speaker's intention. This study focuses on iconic and non-referential co-speech gestures and explores the mismatch between gesture and pitch accent. Numerous production studies confirm the temporal integration of gesture and speech (e.g., Loehr 2012; Esteve-Gibert & Prieto 2013), but research on the perception of this phenomenon is rather scarce (but see Ebert et al. 2022). Our perception study investigates the temporal relationship between gesture and prosody studying whether a temporal mismatch between co-speech gesture and focus accent influences the perceived naturalness of sentences in dependence of the gesture type. We expect that incongruence between the gestural stroke and the pitch accent decreases naturalness with non-referential gestures exhibiting a greater impact due to their discourse structuring function. The study adopts a 3x2 factorial design with the factors 'Gesture type' (iconic vs. non-referential vs. none) and 'Alignment' (aligned vs. unaligned); the none gesture conditions serve as a baseline for the naturalness of the prosody manipulation. Videos of a professional speaker uttering a context sentence and the target sentence with an accompanying co-speech gesture were recorded in the aligned conditions. Mismatched conditions were created by manipulating the focus accent in the target sentence from object to subject position using Praat (Boersma & Weenink 2023). The gesture remains on the object of the target sentence, resulting in a total of six versions per item. In total, the study comprises 18 experimental items. The naturalness of the presented videos is rated by participants on a scale from 1 (absolutely unnatural) to 7 (absolutely natural). The analysis aims to uncover differences in the naturalness ratings between gesture-prosody alignment and gesture types.

References: • Boersma, P. & D. Weenink. 2023. „Praat: doing phonetics by computer". http://www.praat.org/. • Ebert, C., G. Pirillo & S. Walter. 2022. „The Role of Gesture-Speech Alignment for Gesture Interpretation". • Esteve-Gibert, N. & P. Prieto. 2013. „Prosodic structure shapes the temporal realization of intonation and manual gesture movements". • Kendon, A. 1980. „Gesticulation and speech: Two aspects of the process of utterance". *The relationship of verbal and nonverbal communication* 25 (1980): 207–27. • Loehr, D. P. 2012. „Temporal, structural, and pragmatic synchrony between intonation and gesture". *Laboratory Phonology* 3 (1). • McNeill, D. 1992. *Hand and mind: what gestures reveal about thought*. Chicago: University of Chicago Press.