

Multimodal prosodic phrasing in infant-directed speech: testing the cumulative-cue hypothesis with gesture restriction

Roos Ledebøer¹, Victoria Reshetnikova², Roy Hessels³, Aoju Chen²

¹Donders Centre for Cognition, Donders Institute for Brain, Cognition, and Behaviour, Radboud University, ²Institute for Language Sciences, Utrecht University, ³Experimental Psychology, Helmholtz Institute, Utrecht University

Human communication is inherently multimodal (McNeill, 1992); the relation between speech and co-speech gestures is well established and facilitates face-to-face interaction (e.g., Church et al., 2017; Cravotta et al., 2019; Esteve-Gibert & Prieto, 2013). However, the underlying cognitive mechanisms remain unclear. Important insights emerge from research on modality disruption and the Cumulative-Cue Hypothesis proposed for prosodic prominence (Ambrazaitis & House, 2023). This hypothesis entails that when facing restrictions in one of the two modalities, speakers redistribute communicative effort within the same modality and only enhance the cues in the other modality in cases of additional effort. By investigating the effects of hand gesture restriction on acoustic and visual prosodic phrasing cues in infant-directed speech, the Hypothesis can be extended, offering a broader framework for understanding the relation between speech and co-speech gestures. Given its engaging and effortful nature, infant-directed speech may be particularly susceptible to modality disruptions, providing a suitable test context. Interactions between three German tutors and 12 Dutch infants (aged 4–5 months and 8–9 months) were recorded in two conditions: Hands Free and Hands Restricted. A total of 732 segmented intonational phrases and 61,151 corresponding video frames were analyzed to compare the tutors' acoustic cues (pitch maximum, pitch minimum, final syllable duration, pause duration) and visual cues (eyebrow movement frequency and intensity) at final intonational phrase boundaries between conditions. Mixed-effects modelling showed that under hand restriction, eyebrows were raised more frequently, but only during interactions with younger infants. Acoustic cues remained unaffected by hand restriction for both age groups. The finding for the younger group supports the Cumulative-Cue Hypothesis and suggests it extends to prosodic phrasing. The age-related differences may be related to changes in communicative intent and speakers' adaptability to infants' developmental needs, yet they complicate the generalization of the findings and warrant further research.

References

- Ambrazaitis, G., & House, D. (2023). The multimodal nature of prominence: Some directions for the study of the relation between gestures and pitch accents. In O. Niebuhr & M. Svensson Lundmark (Eds.), *Proceedings of the 13th International Conference of Nordic Prosody* (pp. 262–273). Sciendo. <https://doi.org/10.2478/9788366675728>
- Church, R. B., Alibali, M. W., & Kelly, S. D. (Eds.). (2017). *Why gesture?: How the hands function in speaking, thinking and communicating*. John Benjamins Publishing Company.
- Cravotta, A., Bus, M. G., & Prieto, P. (2019). Effects of encouraging the use of gestures on speech. *Journal of Speech, Language, and Hearing Research*, 62(9), 3204–3219. https://doi.org/10.1044/2019_JSLHR-S-18-0493
- Esteve-Gibert, N., & Prieto, P. (2013). Prosodic structure shapes the temporal realization of intonation and manual gesture movements. *Journal of Speech, Language, and Hearing Research*, 56(3), 850–864. [https://doi.org/10.1044/1092-4388\(2012/12-0049\)](https://doi.org/10.1044/1092-4388(2012/12-0049))
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press.