

Multimodality in Emerging Communication Systems: a Virtual Reality Approach

Kotryna Motiekaitytė¹, Oxana Grosseck¹, Luisa Wolf¹, Hans Rutger Bosker^{1,2}, Gerardo Ortega³, Marcus Perlman³, David Peeters^{1,4}, Limor Raviv^{1,5}

¹Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

²Donders Institute for Brain, Cognition and Behavior, Radboud University, Nijmegen, The Netherlands

³English Language and Applied Linguistics, University of Birmingham, Birmingham, UK

⁴Department of Communication and Cognition, TiCC, Tilburg University, Tilburg, The Netherlands

⁵School of Psychology and Neuroscience, University of Glasgow, Glasgow, UK

kotryna.motiekaityte@mpi.nl, oxana.grosseck@mpi.nl, luisa.wolf@mpi.nl, hansrutger.bosker@donders.ru.nl, g.ortega@bham.ac.uk, m.perlman@bham.ac.uk, d.g.t.peeters@tilburguniversity.edu, limor.raviv@mpi.nl

A prominent question in the field of language evolution relates to the role of modality in the origins of language with opposing positions of gesture- vs. speech-first have been proposed to account for early human communication (e.g., Ackermann et al., 2014 for vocalizations-first, Gentilucci & Corballis, 2006 for gestures-first positions). However, recent research increasingly points to a complex interaction of the two modalities in modern languages (Perlman, 2017; Waciewicz & Zywczyński, 2017 among others). Moreover, non-human apes and other primates seem to make use of both gestures and vocalizations to communicate (e.g. Pollick, & De Waal, 2007). Therefore, it is worth considering that multimodality may also date back to the first languages. Since languages don't fossilize, the debate between the gesture-first and speech-first dichotomy or the ways these modalities may interact in the early stages of language evolution is still ongoing. To shed light on these issues, researchers are investigating naturalistic indirect sources of evidence such as creoles and emerging sign languages, as well as using experimental communication paradigms that simulate language emergence whereby participants need to create novel communication systems in the lab. However, many of these experiments use written or pictorial means of communication (e.g., Raviv et al., 2019; Garrod et al., 2007) that do not represent natural communicative situations (and especially those of early humans), with only a handful of studies comparing gestural and vocal language emergence (e.g., Macuch Silva et al., 2020; Fay et al., 2014; Fay et al., 2013).

Here we introduce an experimental study that addresses these issues by using the Cave Automatic Virtual Environment (CAVE), which allows for greater ecological validity and immersion while maintaining tight experimental control (Nölle & Spranger, 2022; Peeters, 2019). Pairs of participants played a communication game in a VR forest, taking turns in describing and matching stimuli (i.e., unfamiliar animals moving in different ways) in three different conditions: vocal, gestural, and multimodal (combining gestures and vocalizations). We present results from a first cycle of data collection and preliminary analyses using accuracy scores, video, and audio data. We make comparisons between conditions based on communicative success, dominance between modalities and how different stimuli features are encoded on acoustic (e.g., vocalization duration, intensity) and gestural (e.g., use of space) measurements. The results help to shed light on the communicative advantages and disadvantages of each modality in the formation of novel communication systems, as well as on the way gestures and vocalizations can intertwine.

Index Terms: language evolution, gesture, vocalization, communication game, virtual reality

References

- Ackermann, H., Hage, S. R., & Ziegler, W. (2014). Brain mechanisms of acoustic communication in humans and nonhuman primates: An evolutionary perspective. *Behavioral and Brain Sciences*, 37(6), 529-546.
- Fay, N., Arbib, M., & Garrod, S. (2013). How to bootstrap a human communication system. *Cognitive science*, 37(7), 1356-1367.
- Fay, N., Lister, C. J., Ellison, T. M., & Goldin-Meadow, S. (2014). Creating a communication system from scratch: gesture beats vocalization hands down. *Frontiers in Psychology*, 5, 354.
- Garrod, S., Fay, N., Lee, J., Oberlander, J., & MacLeod, T. (2007). Foundations of representation: where might graphical symbol systems come from?. *Cognitive science*, 31(6), 961-987.
- Gentilucci, M., & Corballis, M. C. (2006). From manual gesture to speech: A gradual transition. *Neuroscience & Biobehavioral Reviews*, 30(7), 949-960.
- Macuch Silva, V., Holler, J., Ozyurek, A., & Roberts, S. G. (2020). Multimodality and the origin of a novel communication system in face-to-face interaction. *Royal Society open science*, 7(1), 182056.
- Nölle, J., & Spranger, M. (2022). From the field into the lab: Causal approaches to the evolution of spatial language. *Linguistics Vanguard*, 8(s1), 191-203.
- Peeters, D. (2019). Virtual reality: A game-changing method for the language sciences. *Psychonomic bulletin & review*, 26, 894-900.
- Perlman, M. (2017). Debunking two myths against vocal origins of language: language is iconic and multimodal to the core. *Interaction Studies*, 18(3), 376-401.
- Pollick, A. S., & De Waal, F. B. (2007). Ape gestures and language evolution. *Proceedings of the National Academy of Sciences*, 104(19), 8184-8189.
- Raviv, L., Meyer, A., & Lev-Ari, S. (2019). Compositional structure can emerge without generational transmission. *Cognition*, 182, 151-164.
- Waciewicz, S., & Zywczyński, P. (2017). The multimodal origins of linguistic communication. *Language & Communication*, 54, 1-8.