

Do semiotics experiments really show the “superiority” of gesture over vocalization for iconic representation? And even if they do, does it matter?

Marcus Perlman^{*1} and Vinicius Macuch Silva¹

^{*}Corresponding Author: m.perlman@bham.ac.uk

¹English Language and Linguistics, University of Birmingham, Birmingham, United Kingdom

Sometime in the emergence of language, our ancestors faced the challenge of creating new symbols when there were none before. A version of the symbol grounding problem, this ability to establish meaningful symbols without convention is noted as a major hurdle that needed to be overcome to get language off the ground. This has led some researchers to argue for a gesture-first origin of language, based largely on the premise that gestures afford vastly more iconicity than what is assumed to be the only negligible iconicity afforded by vocalizations (e.g., Arbib, 2005; Armstrong & Wilcox, 2007; Corballis, 2003; Sandler, 2013). Thus the use of gestures, but not vocalizations, would have enabled our ancestors to create “already meaningful” signals that could, in turn, be conventionalized into symbols (Tomasello, 2008). Then, at a later stage, vocalizations, intrinsically void of iconicity, would have needed to piggy-back off meaningful gestures – perhaps even fully-fledged signed languages – to bootstrap the transition to the first spoken languages.

A major source of evidence indicating the so-called superiority of gestures over vocalizations for iconic representation comes from semiotics experiments in which participants play “charades” using either gestures or vocalizations, allowing – in theory – a semiotic comparison between them (Fay et al., 2013, 2014, 2022). These experiments have generally found that participants communicate more accurately with gestures than vocalizations, which is taken as evidence of gesture’s greater iconic potential. In this paper, we argue that these experiments, while informative, are fundamentally limited in ways that make the task of comparing gestures and vocalizations unbalanced if not impossible. We provide both methodological and theoretical reasons for why such experiments put vocalizations on unequal footing with gestures, biasing the results towards the conclusion that gestures are superior. First, we explain how these semiotics

experiments, by their design, constrain the use of convention differently in each modality. Producers must suppress iconic words in the vocal condition, whereas they are permitted to gesture freely in the gesture condition – a critical difference that confounds the comparison between the modalities. Second, we argue that these experiments, by implementing an unnatural division that equates vocalization with purely acoustic communication and gesture with purely visual communication, pits the modalities against each other in a completely unrealistic scenario that favours gesture.

By highlighting these issues, we hope to inform future semiotics experiments seeking to compare different modalities in their potential for iconic communication, and concomitantly, for grounding new symbols. We conclude by questioning whether it even matters if gestures are “superior” to vocalizations in their potential for iconicity. While guessing accuracy may be, on average, higher with gestures, semiotics experiments show that vocalizations also afford plenty of iconicity (Ćwiek et al., 2021), which may be sufficient to ground the formation of vocal symbols (Perlman et al., 2015). This point raises important questions for understanding the *complementary* roles of vocalization and gesture in a multimodal evolution of language (Macuch Silva et al., 2020).

References

- Arbib, M. A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*, 28(02), 105–124.
- Armstrong, D. F., & Wilcox, S. E. (2007). *The Gestural Origin of Language*. Oxford University Press.
- Corballis, M. C. (2003). *From hand to mouth: The origins of language*. Princeton University Press.
- Ćwiek, A., Fuchs, S., Draxler, C., Asu, E. L., Dediu, D., Hiovain, K., Kawahara, S., Koutalidis, S., Krifka, M., Lippus, P., Lupyan, G., Oh, G. E., Paul, J., Petrone, C., Ridouane, R., Reiter, S., Schümchen, N., Szalontai, Á., Ünal-Logacev, Ö., ... Perlman, M. (2021). Novel vocalizations are understood across cultures. *Scientific Reports*, 11(1), Article 1. <https://doi.org/10.1038/s41598-021-89445-4>
- Fay, N., Arbib, M., & Garrod, S. (2013). How to Bootstrap a Human Communication System. *Cognitive Science*, 37(7), 1356–1367. <https://doi.org/10.1111/cogs.12048>
- 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. <https://doi.org/10.3389/fpsyg.2014.00354>
- Fay, N., Walker, B., Ellison, T. M., Blundell, Z., De Kleine, N., Garde, M., Lister, C. J., & Goldin-Meadow, S. (2022). Gesture is the primary modality

- for language creation. *Proceedings of the Royal Society B: Biological Sciences*, 289(1970), 20220066. <https://doi.org/10.1098/rspb.2022.0066>
- 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. <https://doi.org/10.1098/rsos.182056>
- Perlman, M., Dale, R., & Lupyan, G. (2015). Iconicity can ground the creation of vocal symbols. *Royal Society Open Science*, 2(8), 150152. <https://doi.org/10.1098/rsos.150152>
- Sandler, W. (2013). Vive la différence: Sign language and spoken language in language evolution. *Language and Cognition*, 5(2–3), 189–203. <https://doi.org/10.1515/langcog-2013-0013>
- Tomasello, M. (2008). *Origins of human communication*. MIT press.