

Iconicity and compositionality in emerging vocal communication systems: a Virtual Reality approach

Kotryna Motiekaitytė^{*1}, Oxana Grosbeck^{1,2}, Luisa Wolf¹, Hans Rutger Bosker², David Peeters^{3,4}, Marcus Perlman⁵, Gerardo Ortega⁵, and Limor Raviv^{1,6}

^{*}Corresponding Author: kotryna.motiekaityte@mpi.nl

¹LEADS group, Max Planck Institute for Psycholinguistics, Nijmegen, NL

²Donders Institute, Radboud University, Nijmegen, NL

³Dept. of Communication and Cognition, Tilburg University, Tilburg, NL

⁴Multimodal language dept., Max Planck Institute for Psycholinguistics, Nijmegen, NL

⁵Dept. of English Language and Linguistics, University of Birmingham, Birmingham, UK

⁶cSCAN, University of Glasgow, Glasgow, UK

A central discussion within the field of language evolution is whether human communication at its origins was spoken, gestured, or multimodal as we see today (Perniss, 2018). Since motivated signals can bootstrap language emergence and learning (Perniss & Vigliocco, 2014), the great affordance for iconicity in gesture and sign languages was seen as supportive of a gesture-first origin (Tomasello, 2010; Fay et al., 2014). However, recent typological and experimental studies suggest that the vocal modality can afford more iconicity than previously thought (Perlman, 2017). Other work suggests that while some meanings can be easily expressed iconically in both modalities (e.g., size, shape, speed), some meanings may be more readily expressed iconically in the gestural modality (e.g., spatial relations) or in the vocal modality, (e.g., qualities of sound) (for overview: Dingemanse et al., 2015; Perlman & Cain, 2014). At the same time, highly iconic communication systems may reduce the pressure to develop compositional structure or systematic form-to-meaning mapping (Verhoef et al., 2016).

Here we examine the development of iconicity and compositionality in novel communication systems across different modalities (vocal, gestural, and multimodal) and across different semantic features. We conducted a dyadic communication game in an immersive Virtual Reality environment, which allows for ecological validity while maintaining high experimental control (Nölle et al.,

2020; Peeters, 2019). Namely, participants interacted face-to-face and without computer interference to refer to novel stimuli around them in a virtual forest, while not being allowed to use any existing language. We compared the creation of new communication systems in 18 pairs across three experimental conditions (vocalization-only, gesture-only, or multimodal; 6 pairs per condition). The stimuli consisted of 32 fantasy creatures that varied by four semantic features: shape (4 types of creatures, Figure 1), size (small vs. big), movement (walk vs. jump), and speed (fast vs. slow). Our results show that communicative success was always highest in the gesture condition, with over 90% accuracy throughout (replicating Fay et al., 2014; Macuch Silva et al., 2020). For vocal and multimodal conditions, accuracy started lower yet increased over time, reaching 73% and 80%, respectively, by the end.

Since annotations of the gesture condition are still ongoing, here we focus on the vocal modality. Specifically, we analyzed the degree of iconicity (measured as effects between semantic and acoustic features) and compositionality (measured as the pair-wise correlation between meanings and orthographic annotations) in the vocal condition. Each vocalization was coded for duration, pitch, loudness, harmonics-to-noise ratio, number of syllables, and speech rate using PRAAT, and data was analyzed using mixed effect regression models in R. Concerning iconicity, we found that multiple acoustic features were significantly related to semantic features. For example, compared to small creatures, big creatures were typically described using lower pitch or louder vocalizations – in line with work on sound symbolism (Nygaard et al., 2009; Perlman & Cain, 2014). Turning to compositionality, novel vocal systems showed varying degrees of compositional structure (Figure 1), with some pairs developing highly structured vocalizations. Compositional structure increased significantly over time and led to increasing communicative success. Finally, pairs that developed the most compositional systems did not necessarily rely less on iconicity for successful communication and vice versa, i.e., the least accurate pairs struggled to employ either iconicity or compositionality. Together, our results suggest that both iconicity and compositionality can develop and co-exist in newly emerging vocal communication systems.

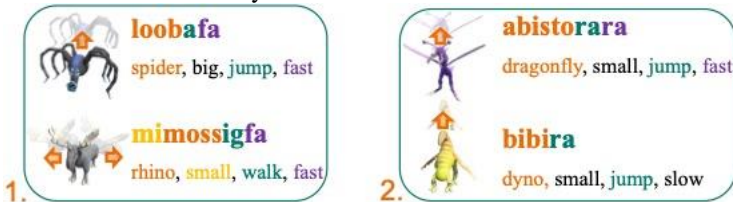


Figure 1. Examples of compositional structure in the annotated vocalizations of two different pairs.

References

- Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in cognitive sciences*, 19(10), 603-615.
- 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.
- 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., Hartmann, S., & Tinitis, P. (2020). Language evolution research in the year 2020: A survey of new directions. *Language Dynamics and Change*, 10(1), 3-26.
- Nygaard, L. C., Herold, D. S., & Namy, L. L. (2009). The semantics of prosody: Acoustic and perceptual evidence of prosodic correlates to word meaning. *Cognitive science*, 33(1), 127-146.
- 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.
- Perlman, M., & Cain, A. A. (2014). Iconicity in vocalization, comparisons with gesture, and implications for theories on the evolution of language. *Gesture*, 14(3), 320-350.
- Perniss, P. (2018). Why we should study multimodal language. *Frontiers in psychology*, 9, 1109.
- Perniss, P., & Vigliocco, G. (2014). The bridge of iconicity: from a world of experience to the experience of language. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1651), 20130300.
- reading. *Cognitive Science*, 39(6), 1348-1368.
- Tomasello, M. (2010). *Origins of human communication*. MIT press.
- Verhoef, T., Kirby, S., & De Boer, B. (2016). Iconicity and the emergence of combinatorial structure in language. *Cognitive science*, 40(8), 1969-1994.