

Contextual factors driving co-speech gesture production: Insights from the ECOLANG corpus

Ed Donnellan¹, Yan Gu^{1,2}, Beata Grzyb¹ & Gabriella Vigliocco¹

¹University College London, UK

²University of Essex, UK

ed.donnellan@ucl.ac.uk, yan.gu@ucl.ac.uk, b.grzyb@ucl.ac.uk, g.vigliocco@ucl.ac.uk

In face-to-face interactions people convey information across modalities. Recently, a growing body of work has demonstrated that interlocutors integrate visual input provided by a speaker's gestures in their speech processing (see Özyürek, 2014; Trujillo & Holler, 2023 for reviews). In particular, co-speech iconic gestures that imagistically represent properties of referents, and deictic gestures (like pointing to or communicatively manipulating referents) influence language processing, as N400 effects occur when gestures are incongruous with accompanying speech (Kelly et al., 2010), and are reduced for less predictable speech accompanied by gestures (congruous in meaning; Zhang et al., 2021). Furthermore, speakers are more likely to produce co-speech iconic gestures when producing less predictable words (Grzyb et al., 2022). While this may support speaker's language production processes (e.g., gestures facilitating thought/speech about concepts; Kita et al., 2017), Grzyb et al. (2022) demonstrated an audience effect (i.e., whether the interlocutor was a child or adult), suggesting that these adjustments are also flexibly deployed by speakers to facilitate interlocutor comprehension. In contrast to these referential gestures, non-referential gestures (e.g., beat or pragmatic gestures) do not clearly support language production processes (Zhang et al., 2021).

The current study investigates adult speakers' use of co-speech gestures when in conversation with other adults. We use the ECOLANG corpus (Gu et al., in prep), a novel semi-naturalistic corpus consisting of $N=33$ adult-adult dyads involved in conversation for 30-40 minutes. In the corpus interaction, one person (designated the speaker) was asked to discuss 24 objects (from 4 groups: animals, food items, musical instruments and tools) with the other. Half of the objects were familiar (e.g., kangaroo, mango, recorder, compass) and half were unfamiliar (e.g., saiga, rambutan, Hulusi, strigil), with the speaker given information on unfamiliar objects to learn one day before the interaction. The same objects were present for half the interaction, and absent for the other half. Taking advantage of these unique manipulations in the ECOLANG corpus, we sought to determine how two interacting contextual factors drive production of both referential (e.g., iconic, pointing and communicative object manipulations) and non-referential co-speech gestures (e.g., pragmatic gestures that manage the interaction by regulating turn-taking or holding the floor; Bavelas et al., 1992). Firstly, co-speech gestures may particularly facilitate learning about novel concepts, with deictic gestures highlighting visually accessible referents (either by pointing at or by manipulating objects), and iconic gestures mapping perceptual and action properties onto referents. Thus, we compared co-speech gesture production between when speakers were talking about familiar versus unfamiliar objects. We predicted that referential gestures would be more likely to accompany utterances about unfamiliar objects, while non-referential gestures should not differ dependent on object familiarity. Secondly, as iconic gestures can refer to properties of displaced objects, we compared co-speech gestures between when speakers spoke about present or absent objects. We predicted that iconic gestures would be more common when referents were absent than present, and that pragmatic gestures (as not meaningfully related to semantic content of speech) would not differ. We transcribed all utterances (defined as a unit that expresses a single event; Berman & Slobin, 1994) and coded iconic gestures, pointing and communicative object manipulations as well as non-referential pragmatic gestures using ELAN (Sloetjes & Wittenburg, 2008). For each utterance ($N=19528$, $Med = 583$, range 432-938 per speaker, excluding filler-only utterances and those that were not about the objects), we determined whether any gestures co-occurred (e.g., overlapped in time). For iconic and pragmatic gestures, we ran logistic mixed effects models, with co-occurrence of gesture (1=co-occurrence, 0=none) predicted by object familiarity (familiar/unfamiliar), interacting with presence (present/absent), with random slopes for each speaker. For pointing and object manipulations, we only considered utterances when objects were present, so models did not include presence as a predictor.

As predicted, co-speech iconic gestures were significantly more common when talking about unfamiliar objects (Figure 1), as were co-speech object manipulations and pointing (Figure 2), but not pragmatic gestures. Iconic gestures were more common when objects were absent than present. However, so too were pragmatic gestures (contrary to our predictions). To some extent, this may be driven by the large number of object manipulations when objects are present (e.g., the hands are less free to gesture) or because they particularly serve a function of acquiring or maintaining another's attention to the interaction in the absence of a shared attentional focus (i.e., the objects). Interestingly however, the absence of an interaction effect between presence and familiarity on the likelihood of producing a co-speech iconic gesture, suggests that they are similarly likely to be deployed when describing unfamiliar objects, even when they are not displaced.

This work demonstrates that multimodal production differs depending on learning context, with referential (but not non-referential) gestures particularly deployed when talking about unfamiliar objects, and iconic gestures particularly used when referents are displaced. This is consistent with both theories that co-speech referential gestures can aid speakers bring properties to mind to facilitate speech about unfamiliar or displaced objects (Murgiano et al., 2021), and also theories that gestures can facilitate interlocutor comprehension and more effective learning (Grzyb et al., 2022; Valenzeno et al., 2003).

Index Terms: co-speech gestures, iconicity, multimodality

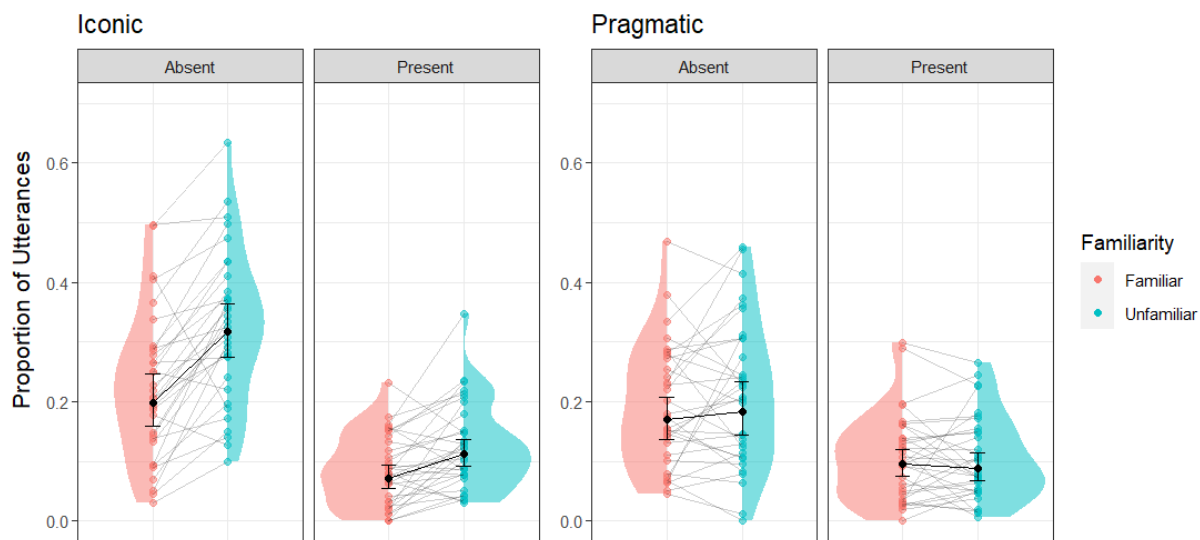


Figure 1: *Proportion of utterances with co-speech iconic and pragmatic gestures when talked about objects are present/absent (facets) and familiar/unfamiliar (red=familiar, blue=unfamiliar). Coloured data points indicate proportions per participant, with half violin plots depicting density. Black points and error bars depict model predictions ($\pm 95\%$ CIs), with the slope depicted by black lines*

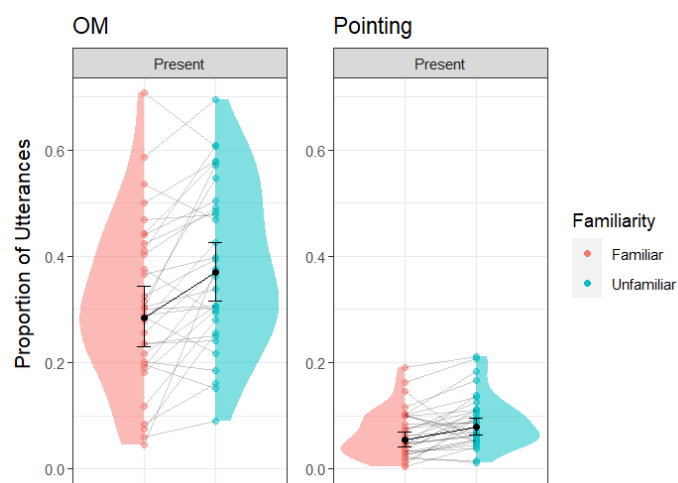


Figure 2: *As Fig. 1, for object manipulations (OM) and pointing (only for utterances produced while objects present)*

References

- Bavelas, J. B., Chovil, N., Lawrie, D. A., & Wade, A. (1992). Interactive Gestures. *Discourse Processes*, 15(4), 469–489. <https://doi.org/10.1080/01638539209544823>
- Berman, R. A., & Slobin, D. . (1994). *Relating events in narrative: A crosslinguistic developmental study*. Lawrence Erlbaum Associates.
- Grzyb, B., Frank, S., & Vigliocco, G. (2022). *Communicative efficiency in multimodal language*. <https://doi.org/10.31234/osf.io/a9wt3>
- Kelly, S. D., Creigh, P., & Bartolotti, J. (2010). Integrating Speech and Iconic Gestures in a Stroop-like Task: Evidence for Automatic Processing. *Journal of Cognitive Neuroscience*, 22(4), 683–694. <https://doi.org/10.1162/jocn.2009.21254>
- Kita, S., Alibali, M. W., & Chu, M. (2017). How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis. *Psychological Review*, 124(3), 245–266. <https://doi.org/10.1037/rev0000059>
- Murgiano, M., Motamedi, Y., & Vigliocco, G. (2021). Situating Language in the Real-World: The Role of Multimodal Iconicity and Indexicality. *Journal of Cognition*, 4(1), 1–18. <https://doi.org/10.5334/joc.113>
- Özyürek, A. (2014). Hearing and seeing meaning in speech and gesture: Insights from brain and behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1651). <https://doi.org/10.1098/rstb.2013.0296>
- Sloetjes, H., & Wittenburg, P. (2008). Annotation by category – ELAN and ISO DCR. *Proceedings of the 6th International Conference on Language Resources and Evaluation*.
- Trujillo, J. P., & Holler, J. (2023). Interactionally Embedded Gestalt Principles of Multimodal Human Communication. *Perspectives on Psychological Science*, 174569162211414. <https://doi.org/10.1177/17456916221141422>
- Valenzano, L., Alibali, M. W., & Klatzky, R. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology*, 28(2), 187–204. [https://doi.org/10.1016/S0361-476X\(02\)00007-3](https://doi.org/10.1016/S0361-476X(02)00007-3)
- Zhang, Y., Frassinelli, D., Tuomainen, J., Skipper, J. I., & Vigliocco, G. (2021). More than words: word predictability, prosody, gesture and mouth movements in natural language comprehension. *Proceedings of the Royal Society B: Biological Sciences*, 288(1955), 20210500. <https://doi.org/10.1098/rspb.2021.0500>