LEXICAL AND GESTURAL ALIGNMENT IN INTERACTION AND THE EMERGENCE OF NOVEL SHARED SYMBOLS

MARLOU RASENBERG[*,1,2], MARK DINGEMANSE[1,2], and ASLI ÖZYÜREK[1,2,3]

[*]Corresponding Author: marlou.rasenberg@mpi.nl
[1]Centre for Language Studies, Radboud University, Nijmegen, The Netherlands
[2]Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
[3]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, the Netherlands

A challenge in the field of language evolution and emergence of new languages is to explain how shared symbols are created for concepts for which conventional labels do not exist. Converging on shared labels can be needed even in interactions between people who do share a conventional language but need to align their concepts for mutual understanding (e.g., when talking about new ideas, plans, innovations or designs). Here we investigate the role of multimodal (gestural and lexical) alignment – that is, using multimodal labels with similar semantic and form features during collaborative referring – in the creation of shared symbols for novel referents among dyads.

Previous research has shown that repetition of lexical items plays an important role in collaborative referring, be it through incremental grounding of shared labels (e.g., Brennan & Clark, 1996) or priming (e.g., Pickering & Garrod, 2004). Gestures too can be a powerful interactional resource, as they allow for iconic mappings and can convey meaning beyond what is expressed in speech (McNeill, 1992). Like words, gestures can be repeated (or "mimicked"), and such repetition has also been shown to aid collaborative referring (Holler & Wilkin, 2011). An open question is the relation between alignment in these two modalities in the context of creating shared labels for novel referents, which remains understudied (cf. Oben & Brône, 2016).

To investigate patterns of multimodal alignment in interaction and the creation of novel labels, we used a director-matcher task in which dyads communicate about unfamiliar 3D objects (i.e., 'Fribbles', Barry, Griffith, De

Rossi, & Hermans, 2014) in 6 consecutive rounds. For this *interaction task*, speech has been transcribed and co-speech gestures annotated. Before and after this interaction, participants individually labelled each object (in 1-3 words) for their partner in a *naming task*. For the purposes of this paper we analyzed data from 8 dyads, 8 target objects and the first two rounds of the interaction task.

To see if dyads converged in their labels for objects before and after the interaction in the *naming task*, we used pretrained word2vec word embeddings (based on the NLPL Dutch CoNLL17 corpus), and quantified the semantic similarity of these descriptions, yielding a simple measure of relative convergence in naming practices. There was a significant increase in the semantic similarity of descriptions, when comparing pre-interaction naming (M = 0.48, SD = 0.16) with post-interaction naming (M = 0.69, SD = 0.24); t(63) = -5.80, p < .001 (one-tailed). In contrast, dyads of participants who did not interact with each other did not show similar post-interaction convergence.

As a next step, the *interactional task* was analyzed to see if lexical and gestural alignment occurred and if they correlated, possibly giving rise to post-interaction convergence. Among the extensive use of iconic gestures (N=561 for 8 dyads), more than half (66.5%) was found to be 'referentially aligned'; that is, these gestures depicted the same subparts of the target objects. Within this set of referentially aligned gesture pairs, the degree of form alignment was calculated by scoring overlap in handedness, position, handshape, orientation and movement. Full alignment of all five form features was uncommon in these dyads (only 4% of all cases), while partial alignment of one up to four features occurred frequently (90%). Subsequently, for those target objects where referential alignment of gestures occurred for a given dyad, a lexical alignment score for that target was computed by comparing the referential speech from both participants (yielding a cosine similarity score). For these cases, a moderate correlation between the number of referentially aligned gestures and lexical alignment was found (r_s = .39; p = .010).

These results show that when creating shared labels for novel objects, interactants align in their iconic gestures as well as their speech, and these two types of alignment are correlated with each other. However, while the presence of gestural alignment is evident at the level of semantics, full form alignment does not seem to be a necessary feature. Existence of partial alignment in gesture is in line with previous reports of grounding processes for speech used in interaction (Clark & Wilkes-Gibbs, 1986). These findings underscore the role of interaction and strategic multimodal alignment as important resources for the emergence of novel shared symbols.

Acknowledgements

References

Barry, T. J., Griffith, J. W., De Rossi, S., & Hermans, D. (2014). Meet the Fribbles: Novel stimuli for use within behavioural research. *Frontiers in Psychology*, *5,* 103.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(6), 1482–1493.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*, 1–39.

Holler, J., & Wilkin, K. (2011). Co-Speech Gesture Mimicry in the Process of Collaborative Referring During Face-to-Face Dialogue. *Journal of Nonverbal Behavior*, *35*(2), 133–153.

McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. Chicago, IL: University of Chicago Press.

Oben, B., & Brône, G. (2016). Explaining interactive alignment: A multimodal and multifactorial account. *Journal of Pragmatics*, *104*, 32–51.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*(2), 169–190.