

## **The role of multimodal interaction in children's language acquisition**

Rowland, Caroline F<sup>1,2</sup>, Sander, Jennifer<sup>1</sup> and Zhang, Yayun<sup>1</sup>

<sup>1</sup>Max Planck Institute for Psycholinguistics, <sup>2</sup>Radboud University

Caroline.Rowland@mpi.nl, Jennifer.Sander@mpi.nl, Yayun.Zhang@mpi.nl

Children and machines learn language in very different ways. Large language models (LLM) passively learn from, and abstract across, regularities in a stream of speech/text. Children learn language in dynamic social interactions, in which they and their interlocutors actively and adaptively construct the communicative environment together. In addition, the interaction is not often a simple dyadic exchange between a person producing language and a person comprehending language, as it is when two adults converse. Instead, interactions between young children and their caregivers are often triadic, incorporating objects or actions in the environment that are perceived and referenced multimodally. These facts have important implications for how we conceptualize the language acquisition process itself.

In this talk, I will present evidence for the importance of considering multimodal interaction in our theories of language acquisition, from studies of children and caregivers conversing in both spoken and sign language in naturalistic play situations. These studies show how objects (such as toys, books), actions, and the non-verbal cues given out by interlocutors all provide key cues that guide the child's attention and affect how they interpret the interaction. They demonstrate that caregivers are highly aware of the attentional and interactional needs of their children and adjust their social interaction to accommodate learning. Correspondingly, they show that children adapt their own attention to maximise learning opportunities, following, in particular, different attentional strategies that depend on the language modality in use. Together the studies paint a picture of language acquisition as an active, dynamic process in which children and adults together direct their focus of attention within multimodal social interaction to maximise the affordances of the learning environment.