

## **Bigger data about smaller people: Exploring language learning at scale**

Prof. Michael C. Frank, *Stanford University*

### **Abstract**

Every typically developing child learns to talk, but children vary tremendously in how and when they do so. What predicts this variability, and what is consistent across children and across learners of different languages? In this talk, I'll describe our efforts to create predictive models of early language learning as a way of formalizing hypotheses in this space. This goal has led us to create open data resources like Wordbank, childes-db, and Peekbank that capture data from tens of thousands of children learning dozens of different languages.

### **Bio**

Michael C. Frank is David and Lucile Packard Professor of Human Biology at Stanford University and Director of the Symbolic Systems Program. He received his PhD from MIT in Brain and Cognitive Sciences in 2010. He studies language use and language learning, focusing especially on early word learning. He is the founder of the ManyBabies Consortium, a collaborative replication network for infancy research, and has led open-data projects including Wordbank and MetaLab. He was a Jacobs Foundation Fellow and has received the Troland Award from the National Academy of Sciences, the FABBS Early Career Impact Award, and the Marr Prize and Glushko Dissertation Prize from the Cognitive Science Society. He served as Chair of the Governing Board of the Cognitive Science Society and has edited for journals including *Cognition* and *Child Development*.