



JEWEL Project

## Mini-Symposium on Paths to Word Meaning 4

# Bridging the data gap between children and AI models



## Michael C. Frank

Stanford University

Commentator



Krista Byers-Heinlein  
Concordia University

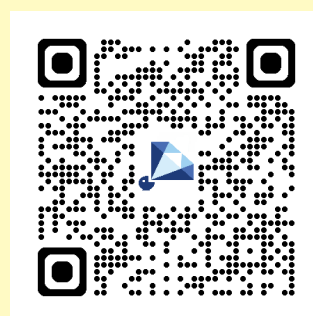
Host



Reiko Mazuka  
RIKEN CBS

**27 Jun. 2025**  
**10:00-11:30 (JST)**  
via Zoom Webinar

Registration site ↓



JSPS Grant-in-Aid for Specially Promoted Research: Grant number 20H05617

Bridging the data gap between children and AI models

Michael C. Frank

(Stanford University)

Large language and language-vision models show intriguing emergent behaviors, yet they receive at least three to four – and sometimes as much as six – orders of magnitude more language data than human children. What accounts for this vast difference in sample efficiency? I will describe steps towards a paradigm in which we can address this question. In particular, I'll discuss the use of child language and egocentric video data for model training, and the use of developmental data for model evaluation. This paradigm provides a model-based framework for exploring the nature of children's early language learning.