

# **Lexical diversity and language development**

**Yawen Yu**

University of California, Los Angeles, LOS ANGELES, California, United States

**Dan Yurovsky**

University of Chicago, Chicago, Illinois, United States

## **Abstract**

Previous research has demonstrated a relationship between quantity of language input and childrens rate of language development: Children who hear more words learn faster. This work takes on two mutually-constraining questions: (1) How should we define quality, and (2) what is the relationship between input quality and language development? We analyzed a longitudinal corpus of interactions between 50 children and their parents using four measures of lexical diversity: Type Token Ratio (TTR), Moving Average TTR, and two more recent measures vocd-D and MTLT. We found that only MTLT gave a prima-facie correct characterization of childrens development, and parents MTLT was correlated with childrens over development. Results of simulations showed that MTLT was distinct from the other measures in its sensitivity to both lexical diversity and word order, suggesting that quality should be defined not just by diversity of words, but also by the variability of sentence structures in which they occur.