Large-scale investigations of variability in children's first words

Rose M. Schneider, Daniel Yurovsky, and Michael C. Frank
Stanford University

Background

First words can reveal much about early cognitive and linguistic development, marking what a child is both able and willing to talk about.

However, because first words are often difficult for an external observer to record or measure, parental reports are a potential method for studying early language development

We use 4 parent-report datasets to explore both the time-course of first words and the relationship between conceptual and linguistic development in early language production

Datasets

Children's first words **(N= 2,279)** drawn from 4 parent reports:

Amazon Mechanical Turk Survey
- 1650 children's first words
- 803 female, 847 male
- M age = 10 mo., median age = 10 mo.

Museum Member Web Survey
- 501 children's first words
- 214 female, 285 male, 2 sex unreported
- M age = 11 mo., median age = 10 mo.

Psycholinguist Web Survey
- 52 children's first words
- 26 female, 26 male
- M age = 11.16 mo., median age = 11 mo.

Wordbank Database
- 76 children producing exactly one word
- 31 female, 45 male
- M age = 10.63 mo., median age = 11 mo.

- "Mama" and "Dada" excluded from analyses
- Responses standardized across datasets when possible (e.g., "Doggy", "Doggie" = "Dog")
- 21 responses excluded

Analysis 1: When does a first word emerge?

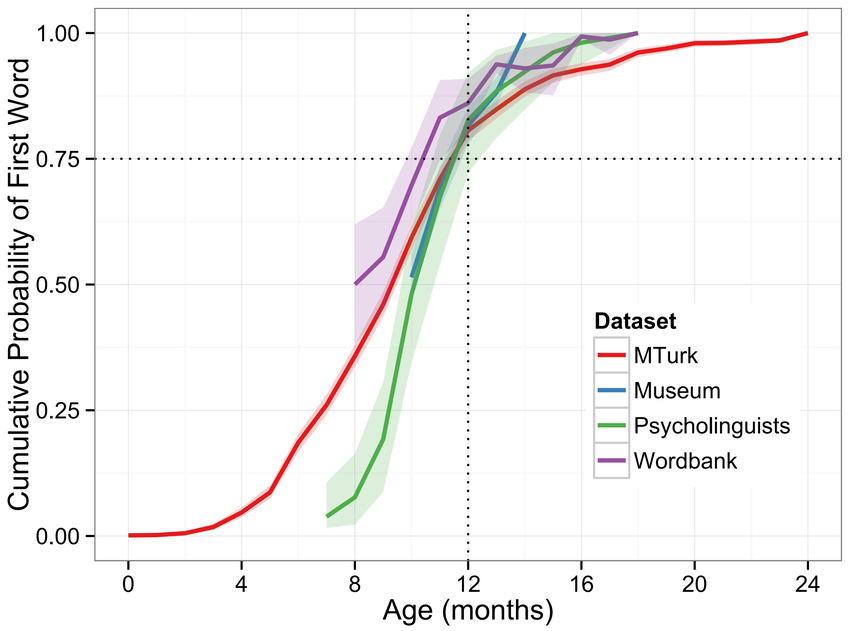

Figure 1: Cumulative probability of a child having a produced a first word across development

Early first words: 75% produced prior to 12 months

Analysis 3: What linguistic factors predict a first word?

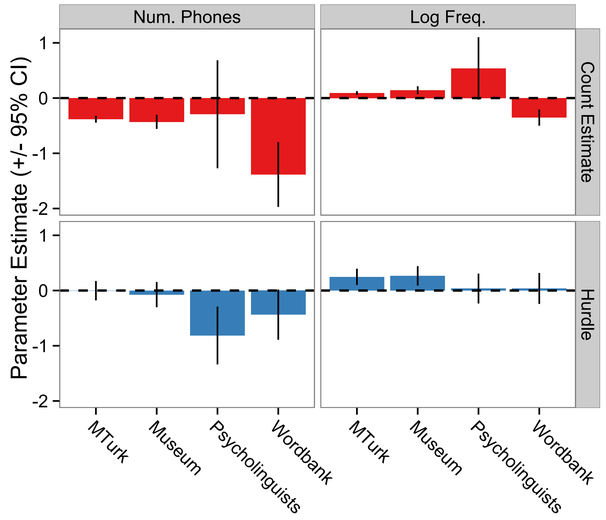

Figure 3: Parameter estimates for hurdle models predicting children's first words

First words phonologically simpler, more frequent in input

Analysis 2: Is there a relationship between age and first word?

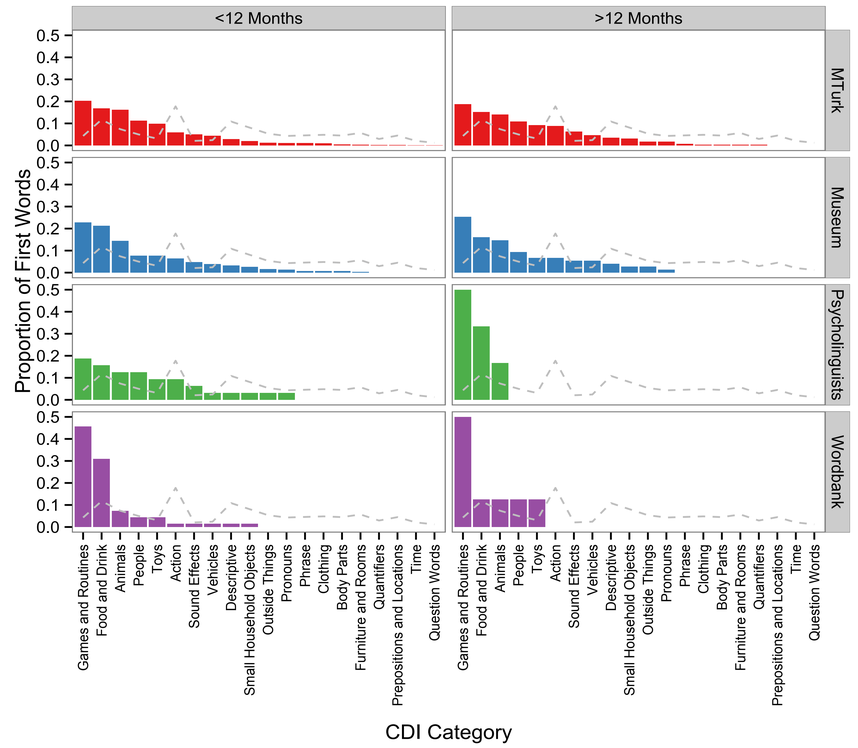

Figure 2: Proportion of children's first words by CDI category, split by earlier (<12 mo.) vs. later (>12 mo.) speakers

MTurk	Museum	Psycholinguists	Wordbank
Dog	**Ball**	Up	Baa Baa
No	**Hi**	More	Uh-Oh
Ball	**Dog**	**Hi**	Yum Yum
Bottle	Uh-Oh	**Cat**	**Woof Woof**
Hi	Duck	**Bye**	**Hi**

Table 1: Top 5 first words collapsed across age

Independence between first word age and conceptual category

Discussion

- Observed first words that were both early and displayed a degree of independence between age and conceptual category, indicating linguistic factors, specifically phonological complexity and frequency in parental speech, likely drive first word production.
- While disadvantages, parent report yielded data consistent both within and across datasets
- Future work includes longitudinal tracking of children's very early language development