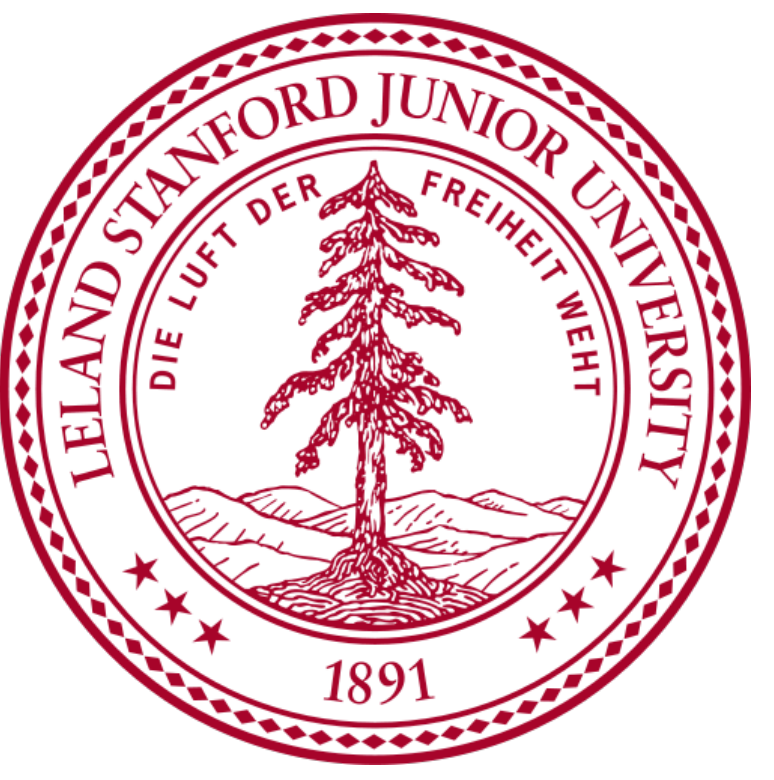


# Large-scale investigations of variability in children's first words

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

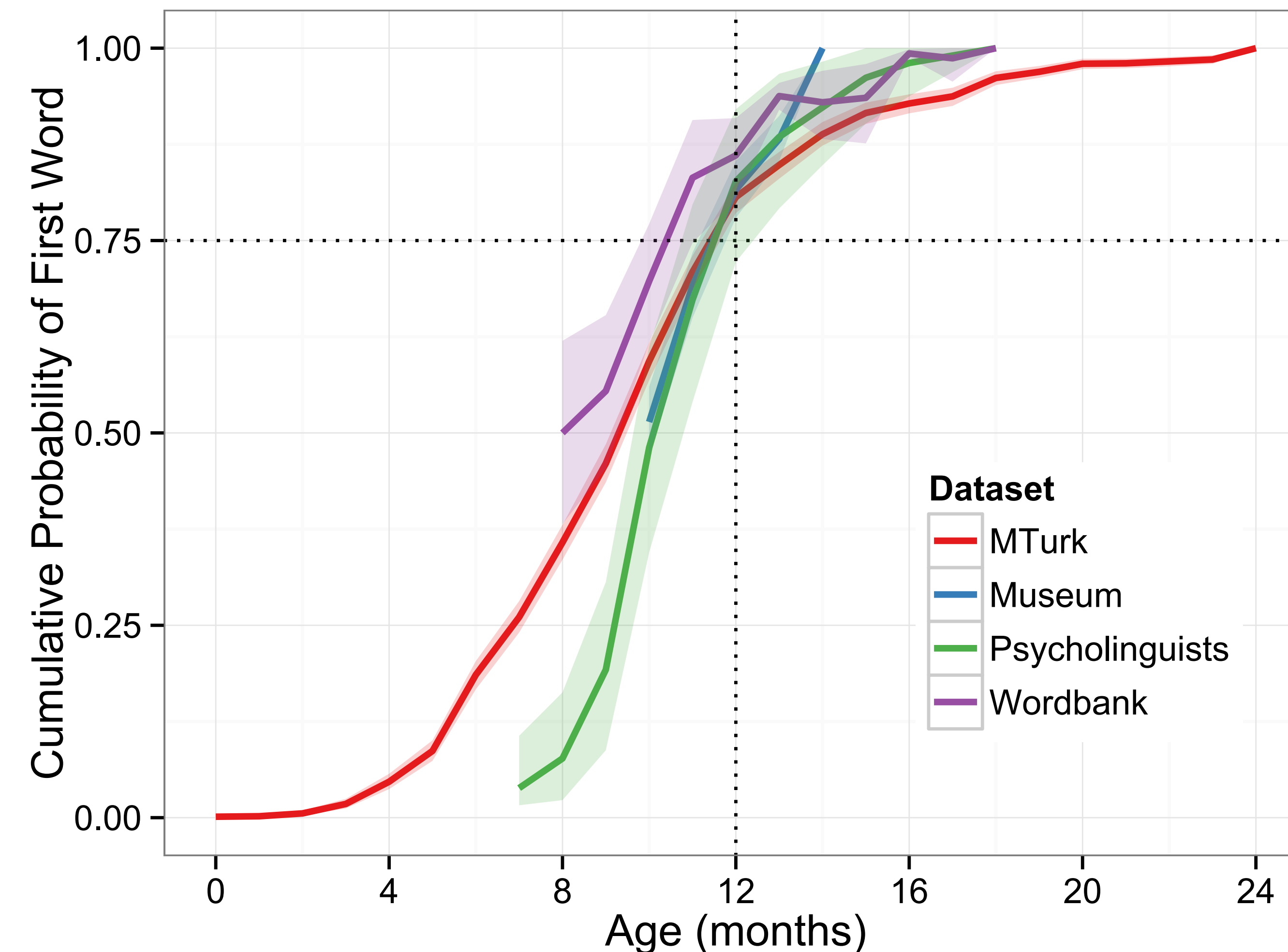
- 52 children's first words
- 26 female, 26 male
- M age = 11.16 mo., median age = 11 mo.

### 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

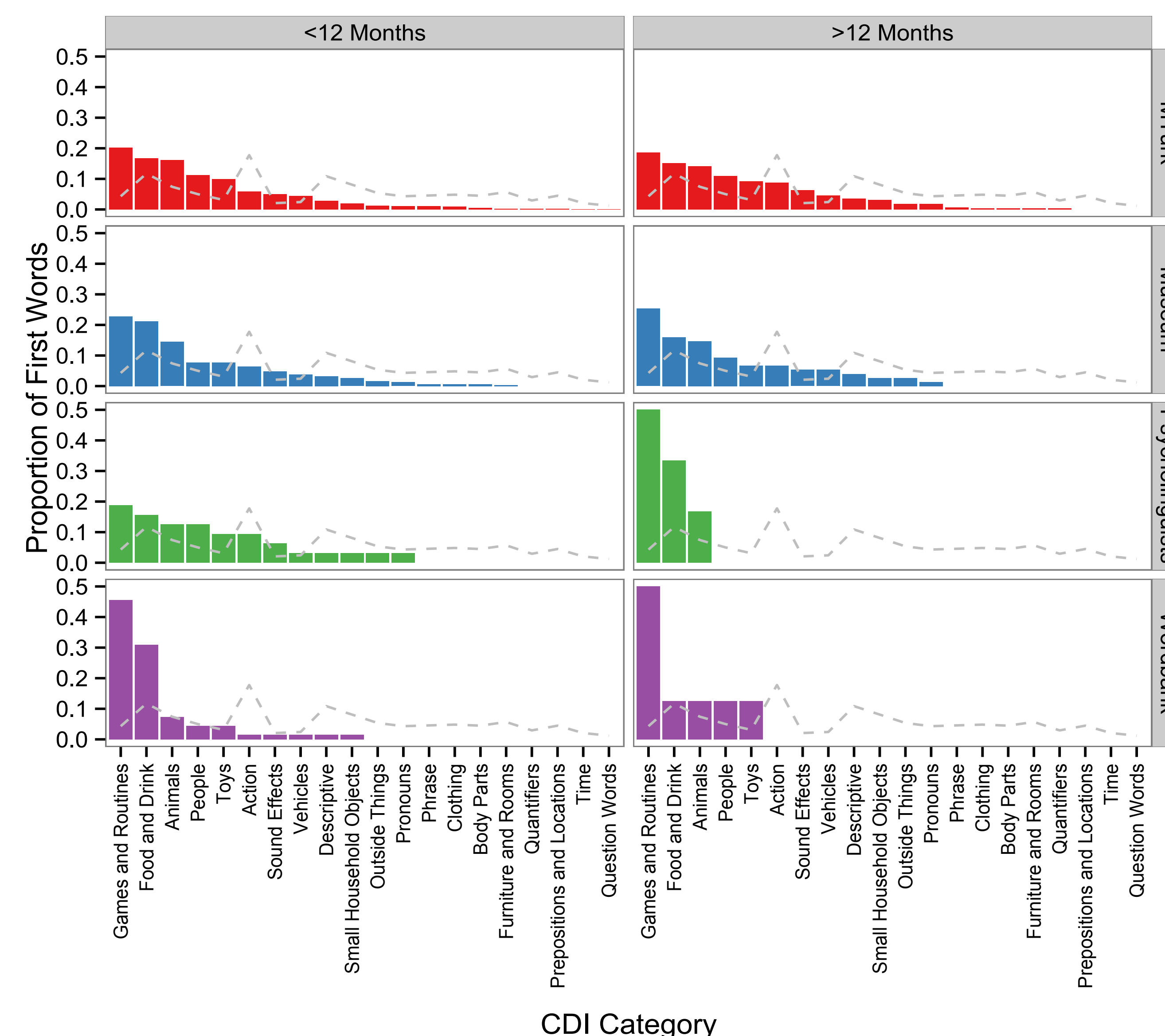
## Analysis 1: When does a first word emerge?



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

**Early first words: 75% produced prior to 12 months**

## Analysis 2: How are age and first word related?



**Figure 2:** Proportion of children's first words by CDI category, split by earlier (<12 mo.) vs. later (>12 mo.) speakers; grey line is baseline CDI distribution

**Independence between first word age and conceptual category**

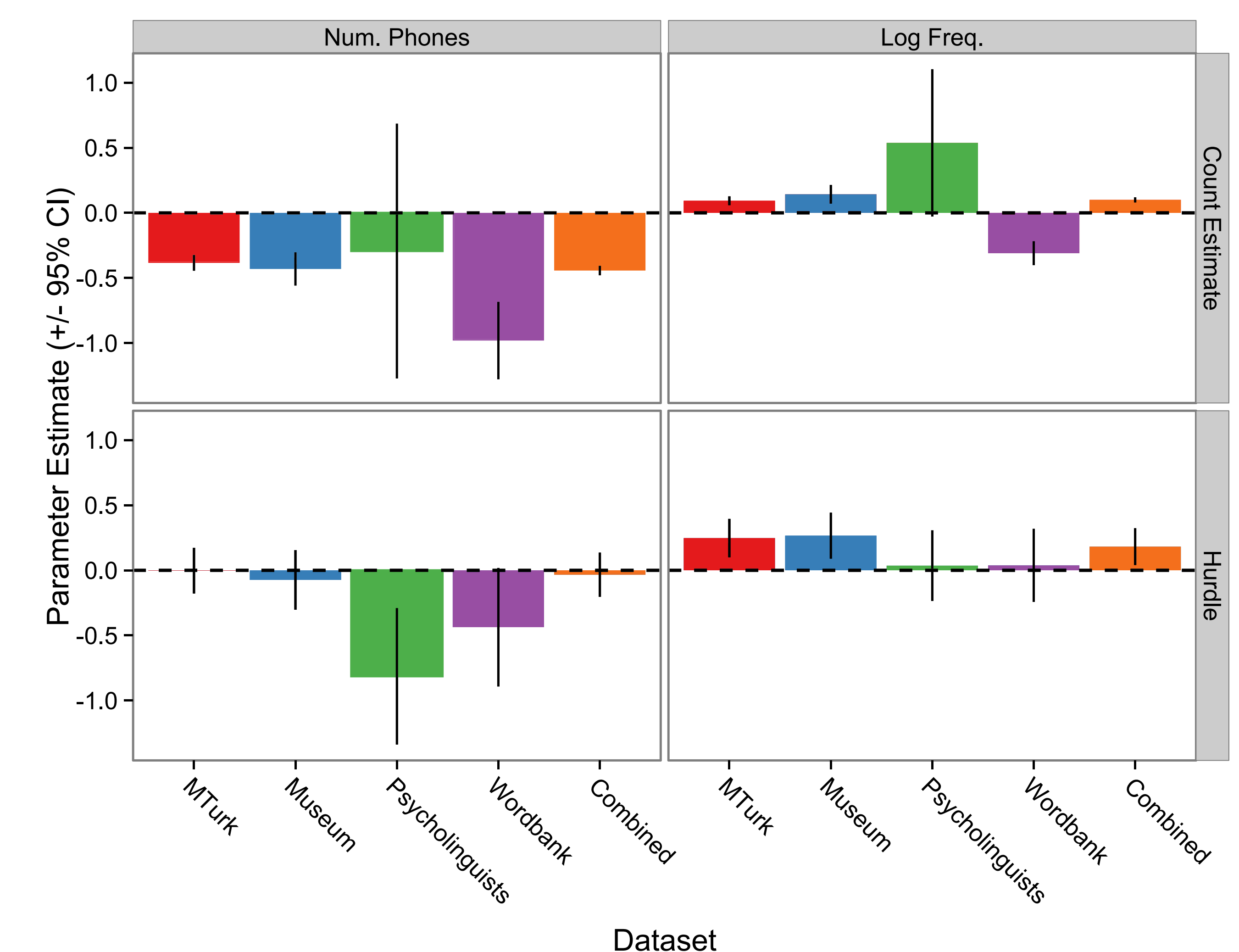
## What are the most common first words?

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

**Table 1:** Top 5 first words from all 4 datasets (collapsed across age)

**First words highly consistent across datasets**

## 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**

## 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