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

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

Background

- First words reveal much about early cognitive and linguistic development, marking what a child both wants to and is capable of talking about at that point in development
- First words often occur in dyadic parent-child moments, and are difficult for an external observer to record or measure
- Parent self-report is a potential method for studying early language development
- We leverage large-scale data from parent reports to address questions on the time-course of first words, and the relation between conceptual and linguistic development

Datasets

First words **(N= 2,279)** drawn from 4 parental reports:

Amazon Mechanical Turk Survey (N = 1650)
- 803 female, 847 male
- M age = 10 mo., median age = 10 mo.

Museum Member Web Survey (N = 501)
- 214 female, 285 male, 2 sex unreported
- M age = 11 mo., median age = 10 mo.

Psycholinguist Web Survey (N= 52)
- 26 female, 26 male
- M age = 11.16 mo., median age = 11 mo.

Wordbank Database (N = 76)
- 31 female, 45 male
- M age = 10.63 mo., median age = 11 mo.

- Datasets selected to address issues of bias
- Stressed first word to be consistent use of a form to indicate a particular meaning
- "Mama" and "Dada" excluded from analyses
- Responses standardized across datasets when possible (e.g., "Doggy", "Doggie" = "Dog)
- First words consistent both within and across datasets
- Classification of CDI categories derived from Macarthur-Bates Communicative Development Index

Analysis 1: When does a first word emerge?

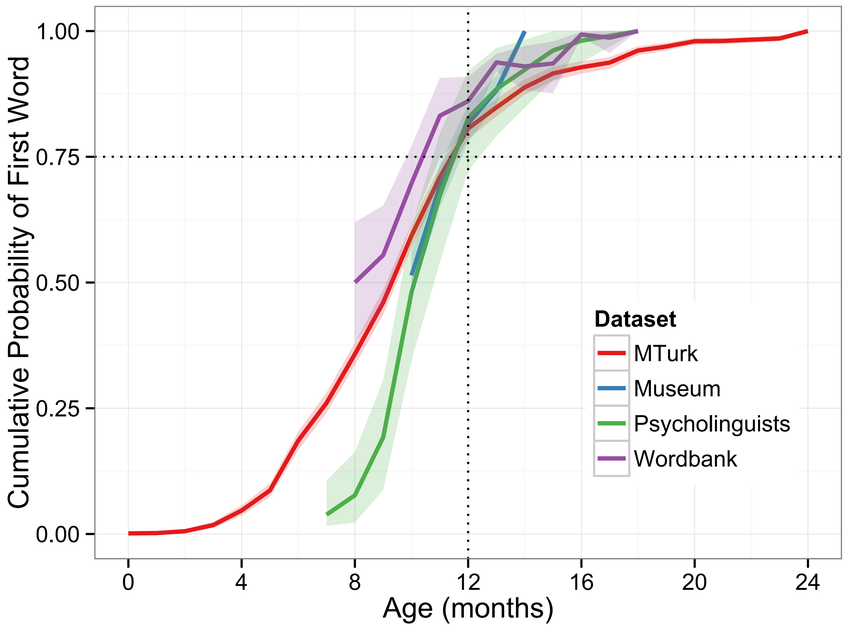

Figure 1: Cumulative probability of a child having a produced a first word across development. About 75% of children produced a first word before 12 months; more than 50% produced a word by 10 months.

Analysis 3: What linguistic factors predict a first word?

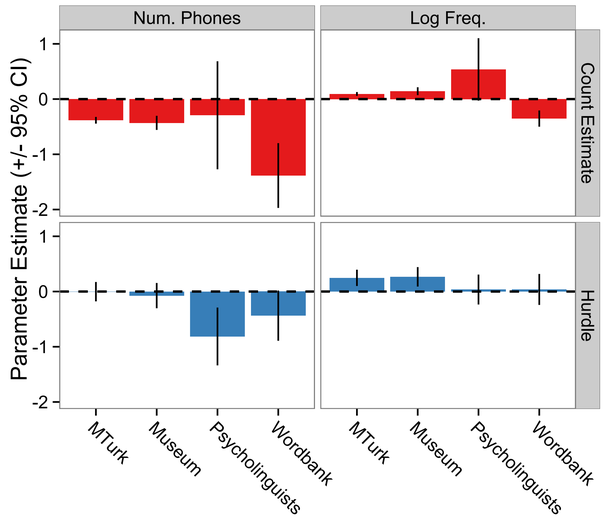

Figure 3: Parameter estimates for hurdle models predicting children's first words: first words across ages tended to have fewer phonemes and were higher frequency in parental input within CDI categories.

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

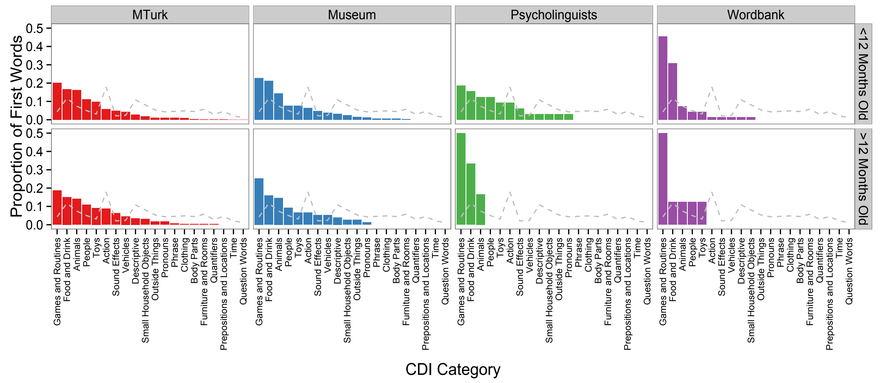

Figure 2: Proportion of children's first words by CDI category, split by earlier (<12 mo.) vs. later (>12 mo.) speakers. Overall, earlier and later producers tend to speak about the same kinds of things, and even the same things.

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 from all datasets collapsed across age. Words repeated across more than two datasets are bolded.

Discussion

- Some evidence for early first words, with 75% of words produced before 12 months
- Degree of independence between age and first word, suggesting that linguistic factors contribute to first word emergence
- Specifically, frequency in parental input and phonetic complexity predict first words
- No evidence of retrospective biases with parents of older children reporting earlier first words
- While disadvantages to parent self-report, the consistency we observed both within and across datasets indicates that first words are a highly memorable event and many of their details can be captured in surveys

Future Work

- To ensure accurate representation of both earlier and later speakers, longitudinal study begun at 6 months that will track children's very early language production
- Deeper understanding of changes in form and meaning throughout development