Lexical predictors of single-word intelligibility in young children's speech

Tristan Mahr, Katherine Hustad
University of Wisconsin–Madison

Background

- Intelligibility measures how well a listener is able to recover a speaker's intended message from a speech signal.
- Research commonly examines predictors of intelligibility by looking at speaker characteristics or acoustic features.
- But speech perception is probabilistic. Frequency or structure of a word can predict how it is perceived (Vitevitch & Luce, 2016).
- Instead of asking how intelligible *speakers* are, it is also important to ask how intelligible individual *words* are.

Current study

We measured the intelligibility of young children and modeled how age and word-level characteristics (frequency, phonotactic probability, motor complexity) predicted intelligibility.

Method

- **Participants.** Speakers were 165 typically developing children aged 2;6–3;11 (months;years; 72 boys, 93 girls).
- **Task.** Speech samples were collected in a structured repetition task based on the TOCS+ (Hodge & Daniels, 2007). Prompts included 38 single words.
- **Listeners.** For each child, two listeners transcribed the child's single word productions (2 x 165 = 330 unique listeners)
- **Model.** Bayesian mixed effects model (1PL item-response):

`correct ~`	`age + [predictors]`	`+ (1 | child)`	`+ (age | word)`
We estimate the probability of a correct response.	We compute the average probability for an average child and on an average item. We use age and the other lexical features to estimate this baseline average.	We can think of each child as having an "ability" value. Some children will be less or more intelligible than the baseline average. Here, we use "by-child varying intercepts" to update the baseline probability with each child's ability.	We can think of each word as having an "easiness" value, and we update the average probability by using these easiness values ("by-item varying intercepts"). The age term here means we are allowing the easiness to change with age for each item ("by-item varying age slopes")

Lexical Features

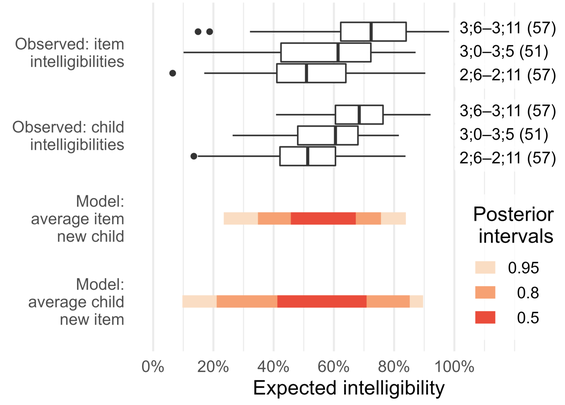
- **Frequency** counts from the SUBTLEX-US database (Brysbaert & New, 2009) (log10 frequency per million words).
- **Phonotactic probability** from the IPhOD database (Vaden, Halpin, & Hickok, 2009), using each word's *average biphone probability* (log10 transformed).
- Total **motor complexity** score (Kuruvilla-Dugdale, Custer, Heidrick, Barohn, & Govindarajan, 2018; based on Kent, 1992). This scheme assigns a score to each syllable part (onset, nucleus, coda) based on its articulatory motor demands, ranging from 1 (/ə, ɑ/) to 8 (cluster of 3 consonants).

Sources of variability

Item variability (in easiness) was greater than child variability (in ability).

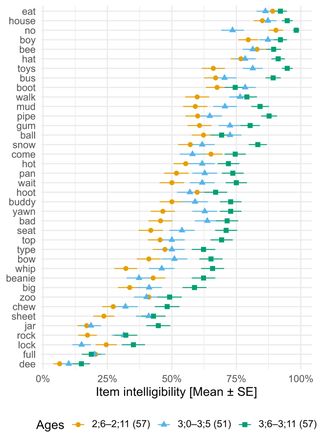
The widths of the boxplot parts for the items are wider for the item means than child means.

The random intercept variance for item difficulties was large than the variance for child abilities. We can demonstrate this effect by simulating new children and new items.

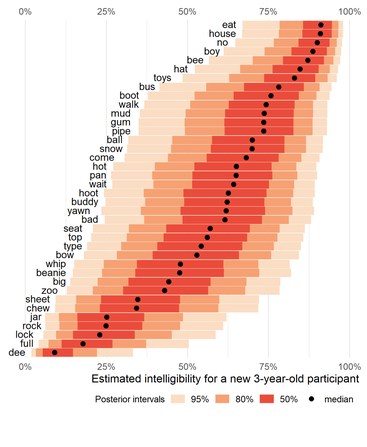


Words differed in their difficulty. What do the most difficult and least difficult words have in common?

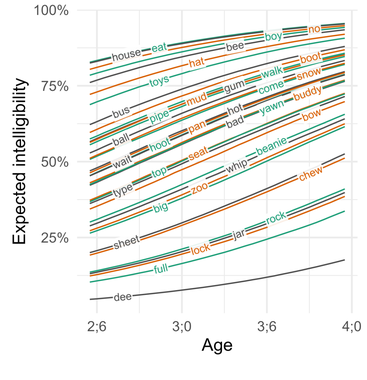
Observed item means by age.



Marginal item means.

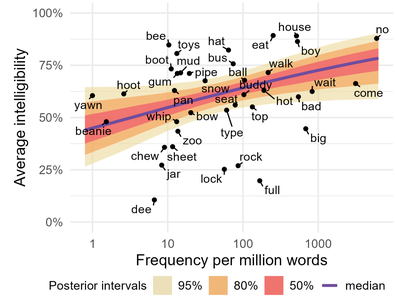


The difficulties of items decreased with age. None of these trajectories (posterior medians) cross each other, suggesting that item difficulties were consistently ranked between ages.



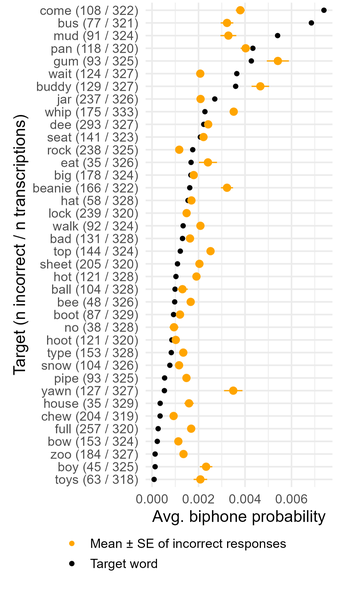
Only a frequency effect was observed

- For **an average item on an average 3-year-old child** with average frequency, phonotactic probability, and complexity, the expected intelligibility was 56%, 95% interval [47, 65].
- A 10x increase in **frequency** predicted an increase in average intelligibility of 10 percentage points, 95% interval [0, 18].
- For **phonotactic probability**, both positive and negative effects are equally plausible, odds ratio 95% interval [0.39, 1.61].
- For a 1-point increase in **motor complexity**, the median odds ratio was 0.93, 95% interval [0.79, 1.08]. The sign is likely negative (complexity penalty), but it is plausible the effect is too small to be meaningful.
- Based on model comparison (LOOIC), there were no statistically credible interactions with age.



Conclusions and future directions

- Only frequency had a statistically clear effect on intelligibility where higher frequency words were more likely to be transcribed accurately.
- However, exploration of the *incorrect* responses suggests a direction for future work. For many words, incorrect responses had higher biphone probabilities than the target on average.
- Listeners might draw on lexical statistics more when given a challenging signal.



Funding provided by NIDCD (R01DC015653)
Conference on Motor Speech
Charleston, SC: February, 2022
Contact: Tristan Mahr (tristan.mahr@wisc.edu)