How do we talk to children? Leveraging speech corpora to quantify how we simplify speech to children

Ruthe Foushee[1], Yang Xu[2] & Mahesh Srinivasan[1]

University of California, Berkeley, University of Toronto

Background

- Children learn from overhearing in lab,[1] but are less likely to at home:[2] *why?*
- The amount of child-directed–but not overheard–speech in children's homes between 18 and 30 months predicts vocabulary growth[2]
- Infants preferentially attend to learnable stimuli of intermediate complexity[3,4]
- Across, cultures, adults simplify their speech to children[5]
- Early in development, might children **initially fail to learn from naturalistic overheard speech, because it is *too complex to capture their attention?***

- We test the idea that overheard speech – which will often consist of speech between adults (ADS) – is too complex for children relative to child-directed speech, leading them to disattend from it until it is of equivalent complexity
- We apply empirically grounded text-based metrics of processing and semantic complexity to child-directed and conversational adult corpora.

Data Sources

Child-directed Speech

CHILDES[6]

- exclusively single adult-child dyads
- 46,234 tokens from 140 children
- CASE STUDIES[7,8]
  - Providence:364 transcripts, 6 dyads
  - Manchester: 12 20-36 month-olds

HOMEBANK

- 53 children (M_{age} = 29.8 mos)
- 159 5-minute transcribed excerpts of daylong recordings from the VanDam corpus[9], 63,807 tokens

Adult-directed Speech

CHILDES

- adult utterances which precede other adult utterances
  - 9222 tokens from 363 adults

HOMEBANK

- 23,744 tokens

CALLHOME CORPUS[8]

NATURALISTIC

BRITISH NATIONAL CORPUS[10]

- 100 million written & spoken tokens
  - 11 million spoken words

LARGE

SANTA BARBARA CORPUS[11]

- informal conversations children might overhear
  - 19 transcripts, 87, 496 tokens

HIGH-QUALITY

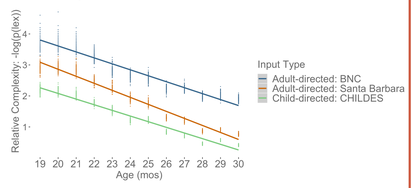
(*Preprocessing…* untranscribable fillers removed; tokens lemmatized)

Computing Complexity

Lexical Complexity

- *How frequent are familiar words?*
  - surprisal (negative log probability) of words known by most same-age children on the M-CDI
  - speech contains fewer highly child-friendly words as children age (B = -6.35 [-6.51, -6.18])
- remains more complex at 30 mos

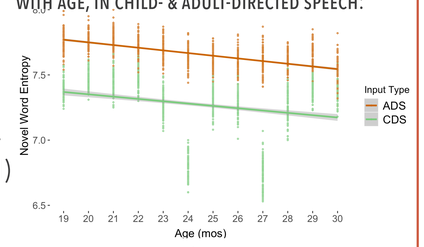
WITH AGE, ACROSS CORPORA:



Entropy

- *How unpredictable, or diverse, are the* unfamiliar *words?*
- entropy decreases with age, but remains greater at 30 mos. ($p < .001$)

WITH AGE, IN CHILD- & ADULT-DIRECTED SPEECH:



Age of Acquisition (AoA)

- AoA is associated with faster online processing; adult judgments[12] may be proxies for subjective complexity
- Of the words with ratings, AoAs were reliably higher for the BNC (M=5.78) & Santa Barbara Corpus (M=6.75), compared to CHILDES CDS (M=4.47)
- AoAs from ADS in CHILDES increased with age (B=.0036)

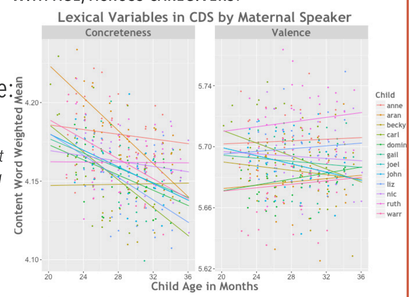
Concreteness

- *How semantically complex is CDS?*
- Concrete language may index here-&-now speech, & be easier to acquire:

A concrete word … refers to something you can have immediate experience of through your senses (smelling, tasting, touching, hearing, seeing) and the actions you do. The easiest way to explain a word is by pointing to it or by demonstrating it, you do not need other language[13]

- CDS is more concrete (difference in means: 0.412 [0.35, 0.48], p < 0.001) & higher valence
- …but CDS decreases in concreteness alone (B= -0.412 [-0.438, -0.385])

WITH AGE, ACROSS CAREGIVERS:

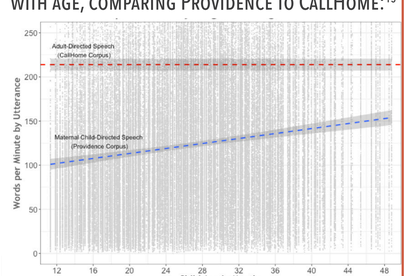


CDS from all but two mothers in the Manchester corpus decreased in concreteness

Speech Rate

- *Is speech between adults faster?*
- 368,009 CDS utterances with times
- 1.28 [1.02, 1.54] wpm increase per month and 0.03 increase in syllables per second [0.02, 0.04]
- IN LAST 6 MONTHS: maternal speech rate 136 wpm [134.14, 137.88], ADS rate 214 wpm [207.27, 220.73]

WITH AGE, COMPARING PROVIDENCE TO CALLHOME:[15]



Conclusions

- Speech to children is reliably less complex & easier to process:
  - it contains higher proportions of words children are likely to know, & less diverse words they may not
- *New* words in child-directed speech are more likely to be learnable compared to those in adult-directed speech:
  - they are more likely to be about the here-and-now and/or object of the child's attention, and easier to learn without other language, through demonstration
  - they are likely to be delivered more slowly and contingently[16]
- If complexity-based attention is relevant for language-learning, overheard speech may not maintain children's attention until at least three years of age.

Limitations

- Many aspects of complexity are missing, e.g., contextual support, syntax…
- Not all overheard speech is between adults!
- Data sparsity: little overheard speech (and labor-intensive to verify)
- CDS and ADS are from different adults, households: ideal corpus would allow analysis of comparative complexity from same speaker around vs. *to* the child.

Future Directions

- *Can children learn from overheard speech equivalent in complexity to the child-directed speech they typically receive?*
- *What do the different trajectories for different complexity measures* ***mean****?*
  - Ongoing experiments test children's attention to language stimuli of contrasting complexity, & qualitative variablity in early overhearing environments.

References

[1] Akhtar, N. (2005). The robustness of learning through overhearing. *Developmental Science*, 8(2).

[2] Shneidman, L. A., Arroyo, M. E., Levine, S. C., & Goldin-Meadow, S. (2013). What counts as effective input for word learning? *Journal of Child Language*, 40(3).

[3] Kidd, C., Piantadosi, S., & Aslin, R. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PloS one*, 7(5), e36399.

[4] Gerken, L., Balcomb, F. K., & Minton, J. L. (2011). Infants avoid 'labouring in vain'by attending more to learnable than unlearnable linguistic patterns. *Developmental Science*, 14(5), 972-979.

[5] Shneidman, L., & Goldin-Meadow, S. (2012). Mayan and US caregivers simplify speech to children. In: *Proceedings of BUCLD 36*.

[6] MacWhinney, B. (2000). The CHILDES project: Tools for analyzing talk.

[7] Demuth, K., Culbertson, J. & Alter, J. (2006). Word-minimality, epenthesis, & coda licensing in the acquisition of English. *Language & Speech, 49*

[8] Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2001). The role of performance limitations in the acquisition of verb-argument structure: An alternative account. *Journal of Child Language*, 28, 127-152.

[9] VanDam, M. (2018). VanDam Public 5-minute HomeBank Corpus.

[10] The British National Corpus (2014) .

[11] Du Bois, J. W., Chafe, W. L., Meyer, C., Thompson, S. A., Englebretson, R., & Martey, N. (2005). Santa Barbara corpus of spoken American English.

[12] Canavan, A., David G., & George Z. (1997). CALLHOME American English Speech

[13] Kuperman et al. (2012) Age-of-acquisition (AoA) norms for over 50 thousand English words. *Behavior Research Methods*.

[14] Brysbaert, M., Warriner, A.B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*.

[15] Yuan, J., Liberman, M., & Cieri, C. (2006). Towards an integrated understanding of speaking rate in conversation.

[16] Yurovsky, D., Doyle, G., & Frank, M. C. (2016). Linguistic input is tuned to children's developmental level. *In CogSci*.

We thank the James S. McDonnell Foundation, NSF GRFP # 1752814, Terry Regier, Julian Arni, Madeleine Peng, Anna Shang, Sathvik Nair, & the creators of childes-db & wordbank.edu. Code at github.com/foushee.