Adult Processing of Child-Produced Speech with and without Background Noise

Marzie Samimifar, Federica Bulgarelli, University at Buffalo

Buffalo EARLY Learning LAB

Introduction

Processing child-produced speech is challenging

- Child-produced speech is non-canonical, like accented speech
- Processing accented-speech impairs spoken word recognition[1]
- Adults exhibit difficulty processing child-produced speech[2,3]

Processing speech in noise is also challenging

- Both artificial and natural background noise hinder speech perception[4,5]
- Some types of background noise help prediction[6]

Listeners can predict upcoming speech

- Context helps listeners predict upcoming speech[7]
- Listeners can predict speech based on the speaker[8]
- Prediction is helpful for processing speech in noisy conditions[9]

Current Study

Research Questions:

1. How do young adults process child-produced speech?
2. How does the child-specificity of target items influence speech perception?
3. How do different types of background noise impact the ability to predict and process child-produced speech?

Method:

Two picture Visual World eye-tracking paradigm (see Figure 1)

Participants: n = 121 (Exp 1 = 41, Exp 2 = 41, Exp 3 = 39)

Three experiments:

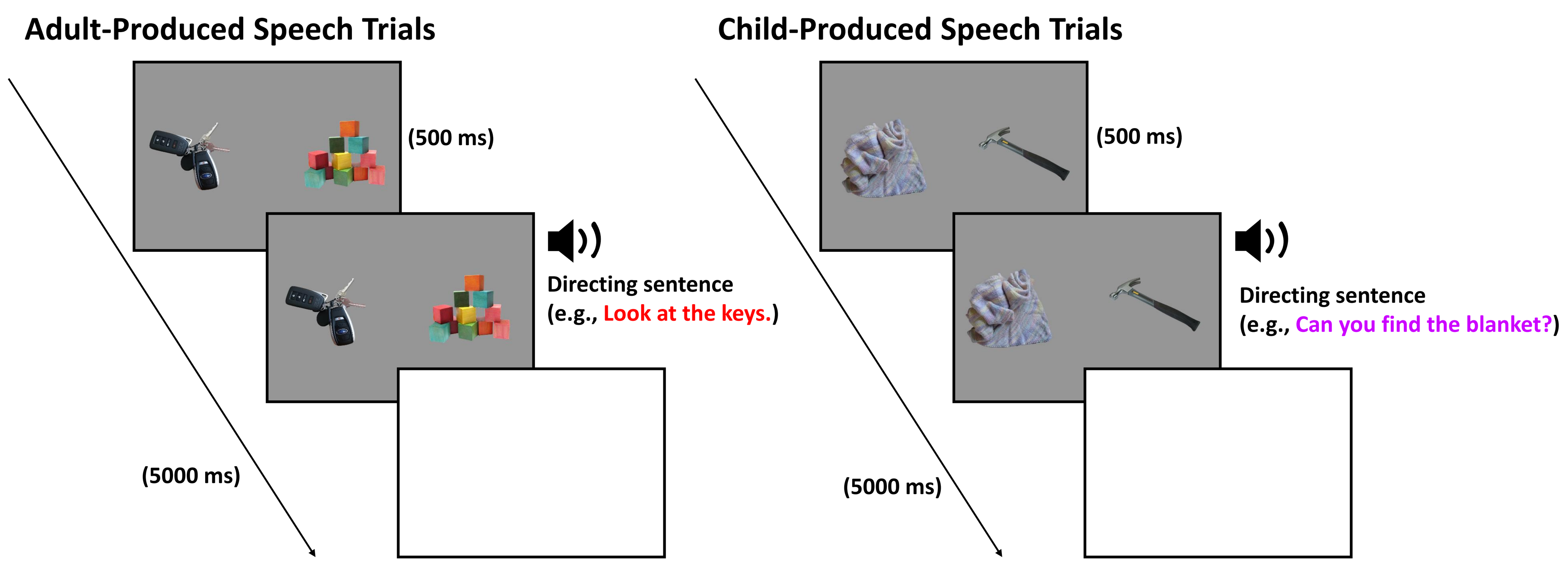
- Exp 1: No background noise
- Exp 2: Artificial background noise (pink noise)
- Exp 3: Real-world background noise (from LENA recordings: noise from children's homes)

48 trials divided to:

- 12: Child speaker, child-specific item
- 12: Child speaker, generic item
- 12: Adult speaker, child-specific item
- 12: Adult speaker, generic item

Procedure

Figure 1. Schematic of Experiment.



Setup of trials: half of the trials (n=24) are produced by an adult and the other half (n=24) are produced by a child. In half of the trials, the target image is a child-specific item (e.g. blocks, blanket) and in the other half is a generic item (e.g., keys, hammer).

Results

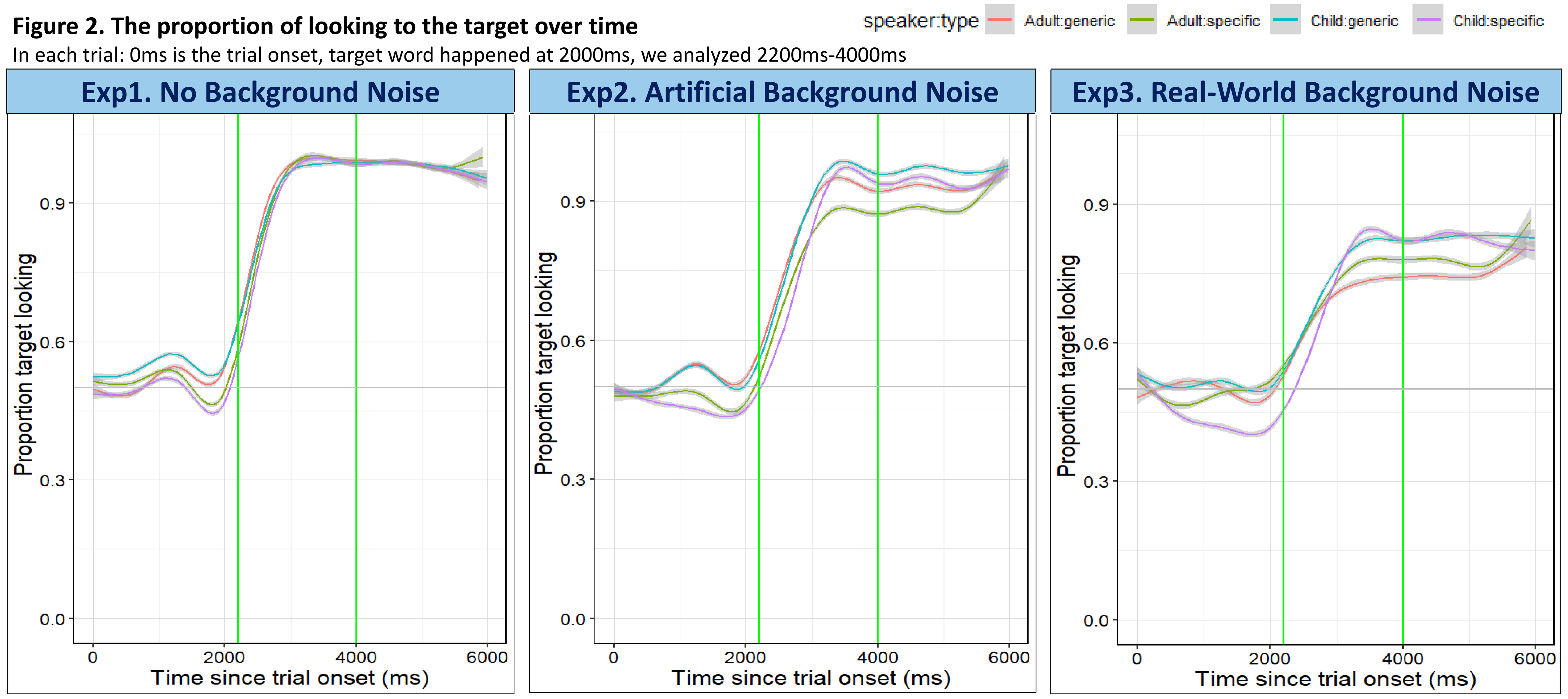
Overall Looking Time Analysis:

Exp1. No Background Noise	Exp2. Artificial Background Noise	Exp3. Real-World Background Noise
• Overall accuracy was 91% (SD = 5) • Speaker-age (p=.011): looked more when produced by an adult • Item-type (p=.005): looked more at generic items • No significant interactions	• Overall accuracy was 82% (SD = 10) • Speaker-age (not significant): adding pink background noise removed the effect of speaker • Item-type (p<.001): looked more at generic items • No significant interactions	• Overall accuracy was 71% (SD = 17) • Speaker-age (not significant) • Item-type (not significant) • Adding real-world background noise removed all main effects • No significant interactions

Growth Curve Model Analysis:

Figure 2. The proportion of looking to the target over time

In each trial: 0ms is the trial onset, target word happened at 2000ms, we analyzed 2200ms-4000ms



Exp1. No Background Noise
- Fastest looking in adult-produced speech and generic item condition

Exp2. Artificial Background Noise
- Faster to look at generic items in both child-produced speech and adult-produced speech conditions
- Lowest peak in adult-produced speech and child-specific condition

Exp3. Real-World Background Noise
- Looked most in child-produced speech conditions
- Reached a higher peak for child-specific items

Conclusions

- **RQ1:** Child-produced speech is more challenging to process than adult-produced speech
- **RQ2:** Adults are slower to look at the target for child-specific items
- **RQ3:** The type of background noise can influence processing:
  - Artificial noise seems to make processing more challenging
  - Real-world noise seems to help processing of child-produced speech by allowing listeners to make predictions
- **Listeners leverage background noise and speaker identity when making predictions about upcoming speech**

Future directions:

- Is it harder to process child-produced speech due to unfamiliarity or higher cognitive demands?
- How do toddlers process child-produced speech in silence and background noise?
- Does hearing child-produced speech in a second language make it more difficult to process?

Acknowledgements

We thank all participants and research assistants who helped with recruitment and data collection.

Citations

1. Porretta, V., Tucker, B. V., & Järvikivi, J. (2016). The influence of gradient foreign accentedness and listener experience on word recognition. Journal of Phonetics, 58, 1-21.
2. Creel, S. C., & Jimenez, S. R. (2012). Differences in talker recognition by preschoolers and adults. Journal of Experimental Child Psychology, 113(4), 487-509.
3. Cooper, A., Fecher, N., & Johnson, E. K. (2020). Identifying children's voices. The Journal of the Acoustical Society of America, 148(1), 324-333.
4. Strauß, A., Wu, T., McQueen, J. M., Scharenborg, O., & Hintz, F. (2022). The differential roles of lexical and sublexical processing during spoken-word recognition in clear and in noise. Cortex, 151, 70–88.
5. Lee, J. Y., Lee, J. T., Heo, H. J., Choi, C.-H., Choi, S. H., & Lee, K. (2015). Speech recognition in real-life background noise by young and middle-aged adults with normal hearing. Journal of Audiology & Otology, 19(1), 39.
6. Meylan, S. C., Foushee, R., Wong, N. H., Bergelson, E., & Levy, R. P. (2023). How adults understand what young children say. Nature Human Behaviour, 7(12), 2111-2125.
7. Brouwer, S., Mitterer, H., & Huettig, F. (2013). Discourse context and the recognition of reduced and canonical spoken words. Applied Psycholinguistics, 34(3), 519-539.
8. Van Berkum, J. J., Van den Brink, D., Tesink, C. M., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. Journal of cognitive neuroscience, 20(4), 580-591.
9. Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension?. Trends in cognitive sciences, 11(3), 105-110.