

Children's Real-Time Spoken Word Recognition while Listening in Two-Talker Speech

Kelsey Klein, AuD, PhD

Center for Pediatric Hearing Health Research, House Institute Foundation

INTRODUCTION

- To comprehend speech, listeners must quickly and efficiently employ two processes (Dahan & Tanenhaus, 2004):
 - 1) Activate lexical candidates consistent with the speech signal
 - 2) Activate semantic features of these candidates
- Under *ideal listening conditions*, listeners immediately activate many lexical candidates that match early portions of the input. Candidates compete over time until only one remains (Marslen-Wilson, 1987). Lexical access speed directly predicts semantic activation speed (Klein et al., 2023).
- When listening in high-intensity noise, adults are slower to initiate lexical access and slower to suppress lexical competition, relative to listening in quiet (Ben-David et al., 2011; Brouwer & Bradlow, 2016).
- It is unknown how a competing signal, like *two-talker speech (TTS)*, affects the time course of real-time speech recognition processing in children, including lexical access and semantic activation.

RESEARCH QUESTION

What are the effects of **low-intensity** and **high-intensity** TTS on the time course of real-time **lexical access** and **semantic activation** in school-age children with typical hearing?

Hypothesis 1: As long as the competing TTS is at an intensity that is low enough so that word recognition accuracy is minimally impacted, the time course of children's lexical access and semantic activation will not be affected.

Hypothesis 2: TTS of any intensity will slow children's lexical access and semantic activation, due to reducing auditory access and/or drawing attention away from the target speech.

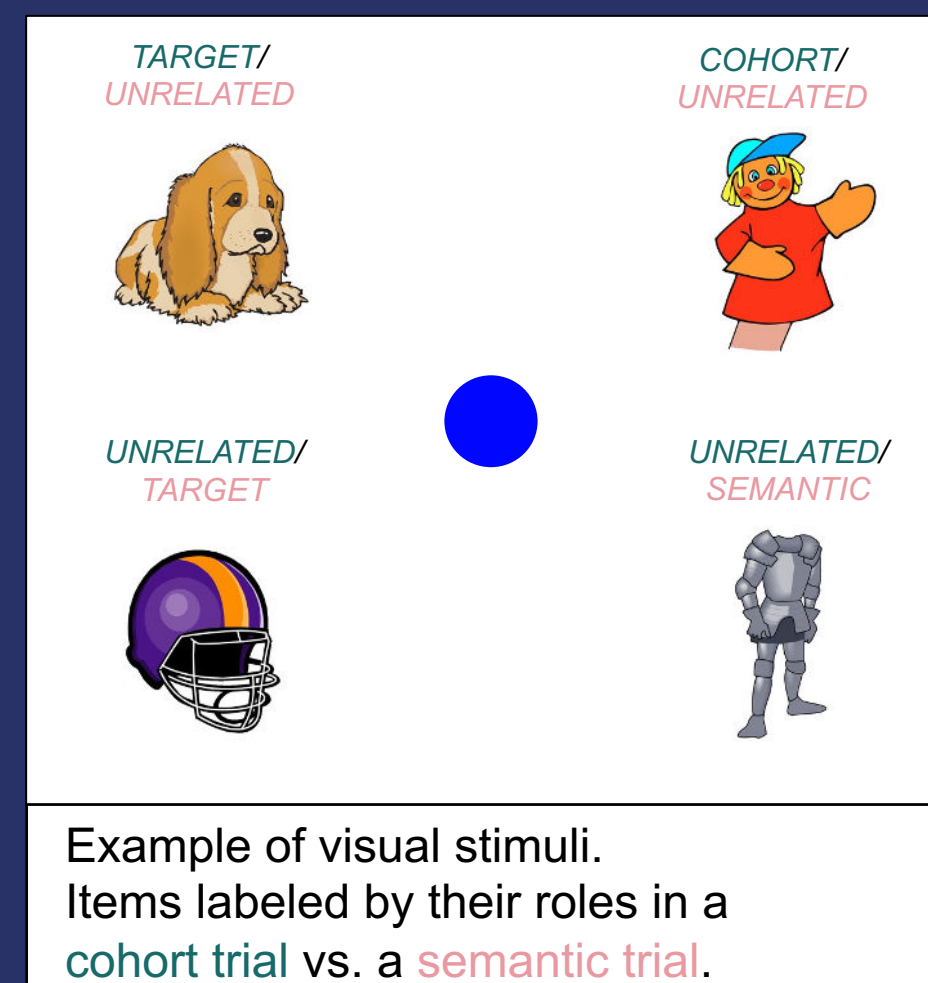
METHOD

Participants

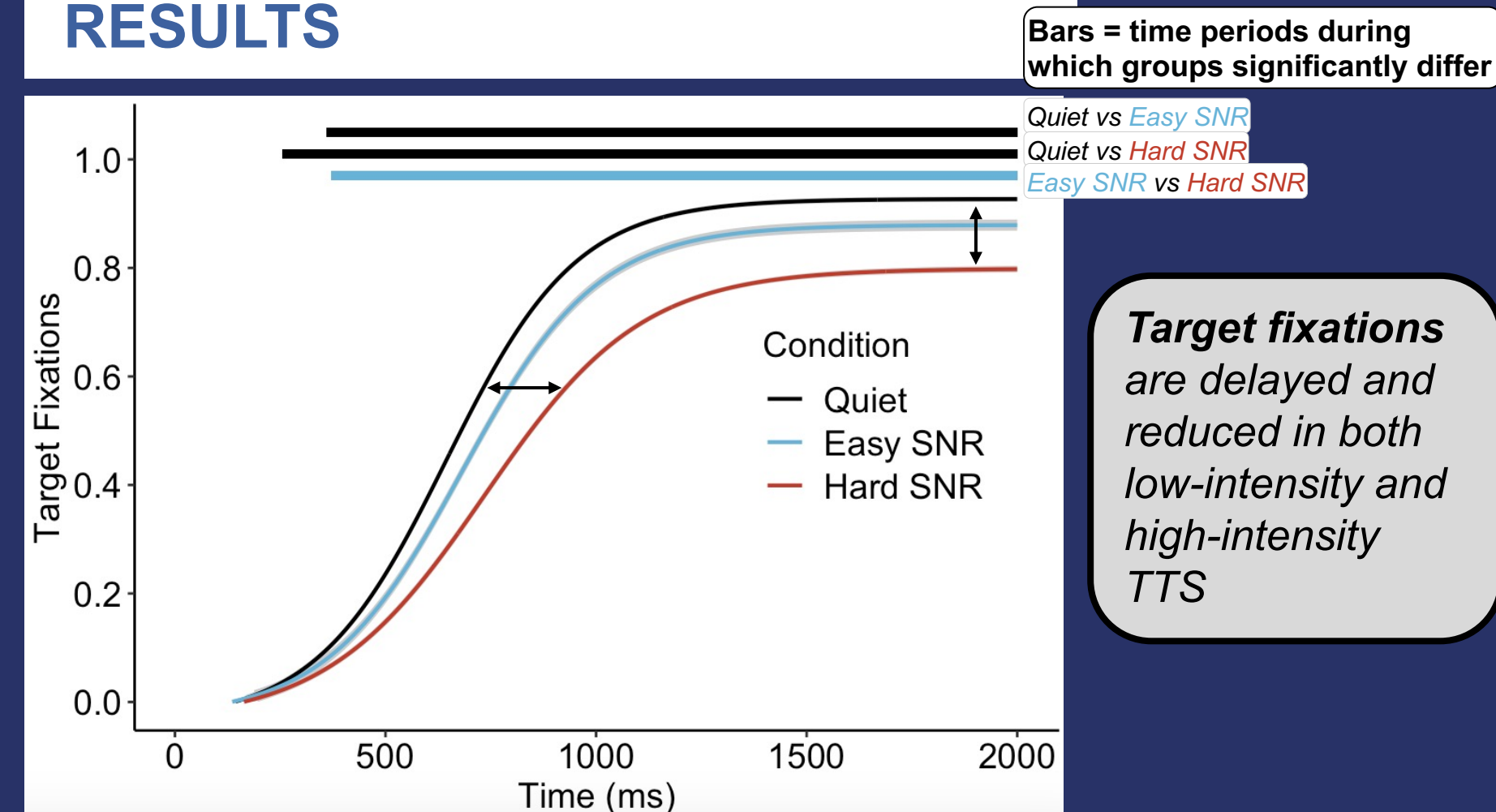
- 19 children with typical hearing, ages 10-12 years
- 13 girls, 6 boys
- Receptive vocabulary: M = 116.3, SD = 13.3

Visual World Paradigm Eye-Tracking Task

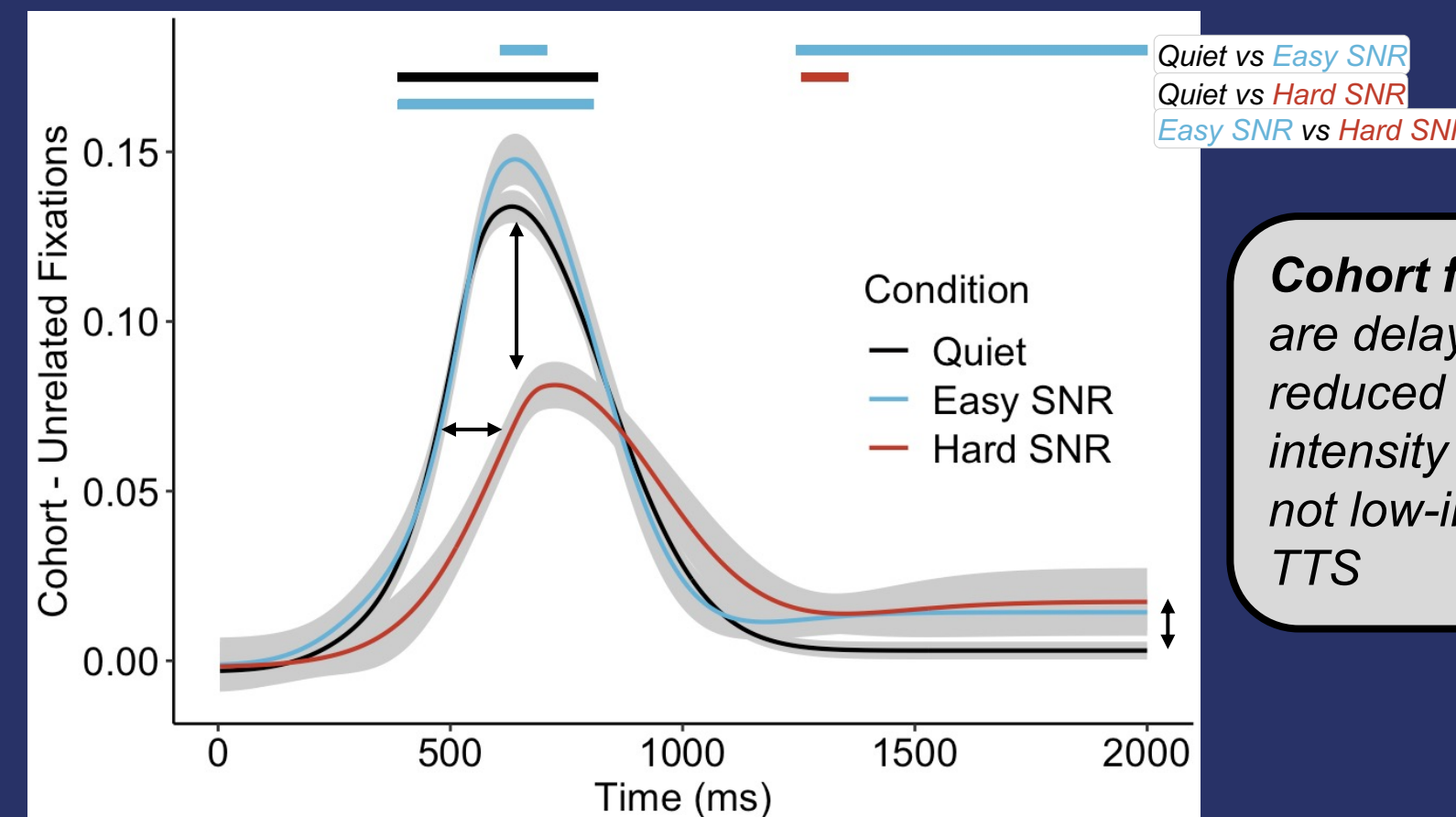
- Isolated target words presented via loudspeaker
- Screen contains four pictures: one target, two unrelated items, and either a phonologically similar cohort (e.g., *puppy*, *puppet*) or semantically related item (e.g., *helmet*, *armor*)
- Child clicks on picture corresponding to target word
- Each child completes 116 each of cohort and semantic trials, interleaved
- Eye movements recorded and analyzed every 4 ms
- Eye fixations index the extent to which child is considering each item
- Analyses include only correct trials
- Each participant completed 3 conditions, counterbalanced:
 - 1) **Quiet**
 - 2) **Hard SNR**, with TTS presented at individualized SNR-70 determined with adaptive task (i.e., high-intensity TTS)
 - 3) **Easy SNR**, with TTS presented 10 dB lower than in Hard SNR condition (i.e., low-intensity TTS)



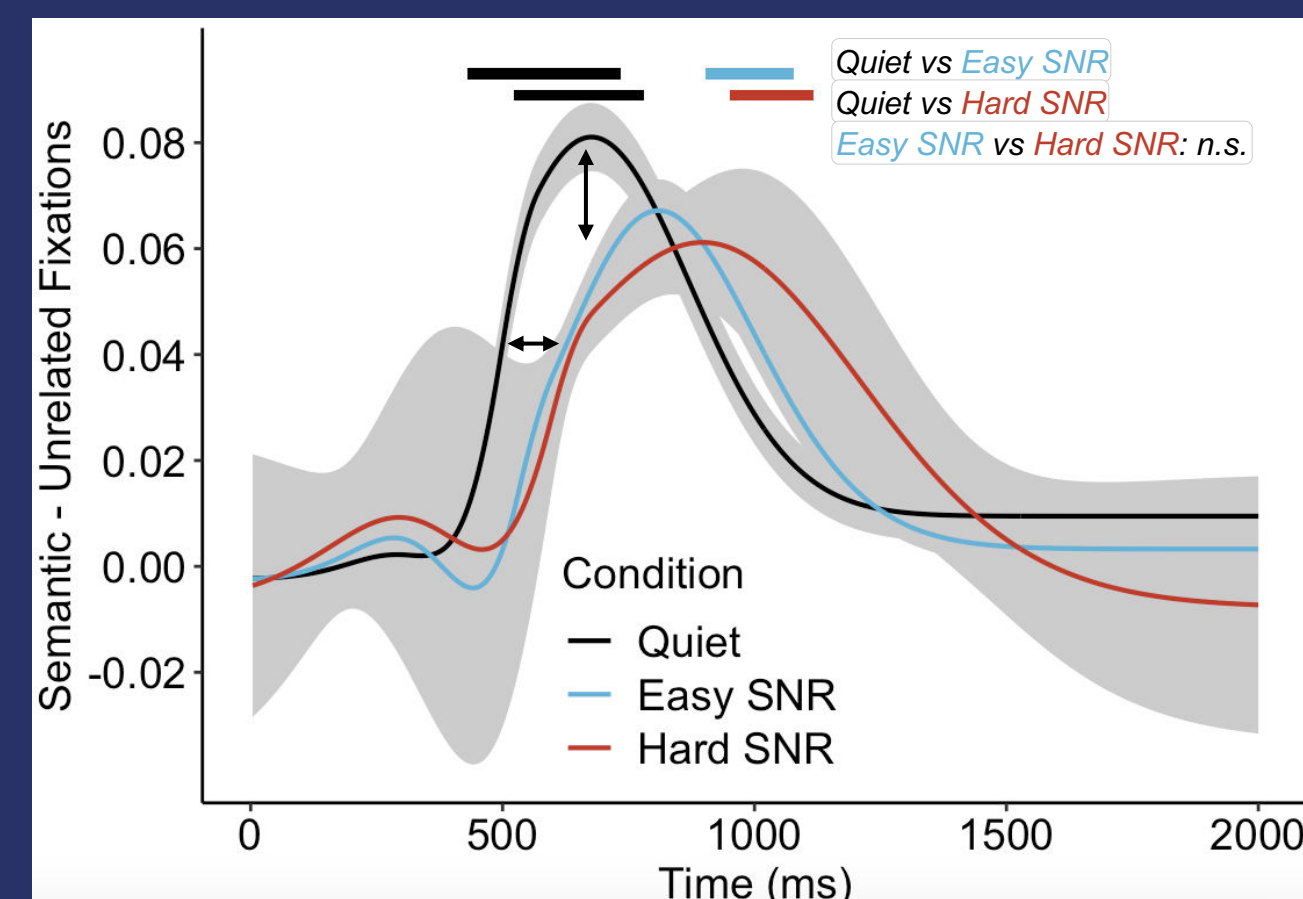
RESULTS



Target fixations are delayed and reduced in both low-intensity and high-intensity TTS



Cohort fixations are delayed and reduced in high-intensity TTS, but not low-intensity TTS



Semantic fixations are delayed and reduced in both low-intensity and high-intensity TTS

CONCLUSIONS

- Results support *Hypothesis 2*: both low-intensity and high-intensity TTS slow children's real-time lexical access (target fixations) and semantic activation (semantic fixations).
- Even when they ultimately choose the correct word, TTS reduces children's resolution of lexical competition at the end of the trial.
- Lexical competition (cohort fixations) was delayed only by the high-intensity TTS, though the low-intensity TTS caused incomplete competitor suppression.
- Low- and high-intensity TTS resulted in semantic activation that was delayed and reduced to the same extent.
- Even when TTS is at a low enough intensity so that word accuracy is minimally affected, children's semantic processing is diminished to the same extent as if high-intensity TTS is present.
- Relatively low levels of background noise in a classroom may impair children's ability to process the meaning of what they hear, even when word recognition is excellent.**

REFERENCES

Ben-David et al. (2011). *JSLHR*, 54, 243-262.
Brouwer & Bradlow (2016). *J Psycholinguist Res*, 45, 1151-1160.
Dahan & Tanenhaus (2004). *J Exp Psychol Learn Mem Cogn*, 30, 498-513.
Klein et al. (2023). *Ear Hear*, 44(2), 338-357.
Marslen-Wilson (1987). *Cognition*, 25, 71-102.

ACKNOWLEDGMENTS

Thanks to Tina Grieco-Calub for support with study design, and Layla Cameron, Grace Rhein, Lona Rhodes, Sara Smith, and Grace Sullivan for support with data collection. Funding provided by a New Investigators Research Grant from the ASHFoundation.

CONTACT

Kelsey Klein: kklein@hifla.org

		Quiet	Easy SNR	Hard SNR
SNR	Mean	-	4.5 dB	-5.5 dB
	SD	-	3.7 dB	3.7 dB
	Range	-	0 to 14 dB	-10 to +4 dB
Accuracy	Mean	99.6%	95.2%	83.1%
	SD	0.4	2.6	6.1
	Range	98.3 to 100%	90.1 to 98.7%	72.4 to 93.1%

SNR and accuracy in each eye-tracking condition.