

How do children learn to integrate social knowledge with information from talker voice? How does this bear on the perpetuation of gender stereotypes?

Background

Speech perception allows us to study children's burgeoning ability to comprehend social information in the speech stream.

The speech signal contains:

- **Phonemic information:** 'sound' information necessary for word recognition and word segmentation.
- **Talker information:** 'identity' information conveying speaker age, gender, class, etc. Conveys social identity and is also important for word recognition (Johnson, 2006).

By adulthood, listeners use talker information to make social inferences about a talker's likely behavior (Van Berkum et al., 2008), especially when they expect talker identity to be useful or find it to be a reliable cue (Creel, Aslin, & Tanenhaus, 2008).

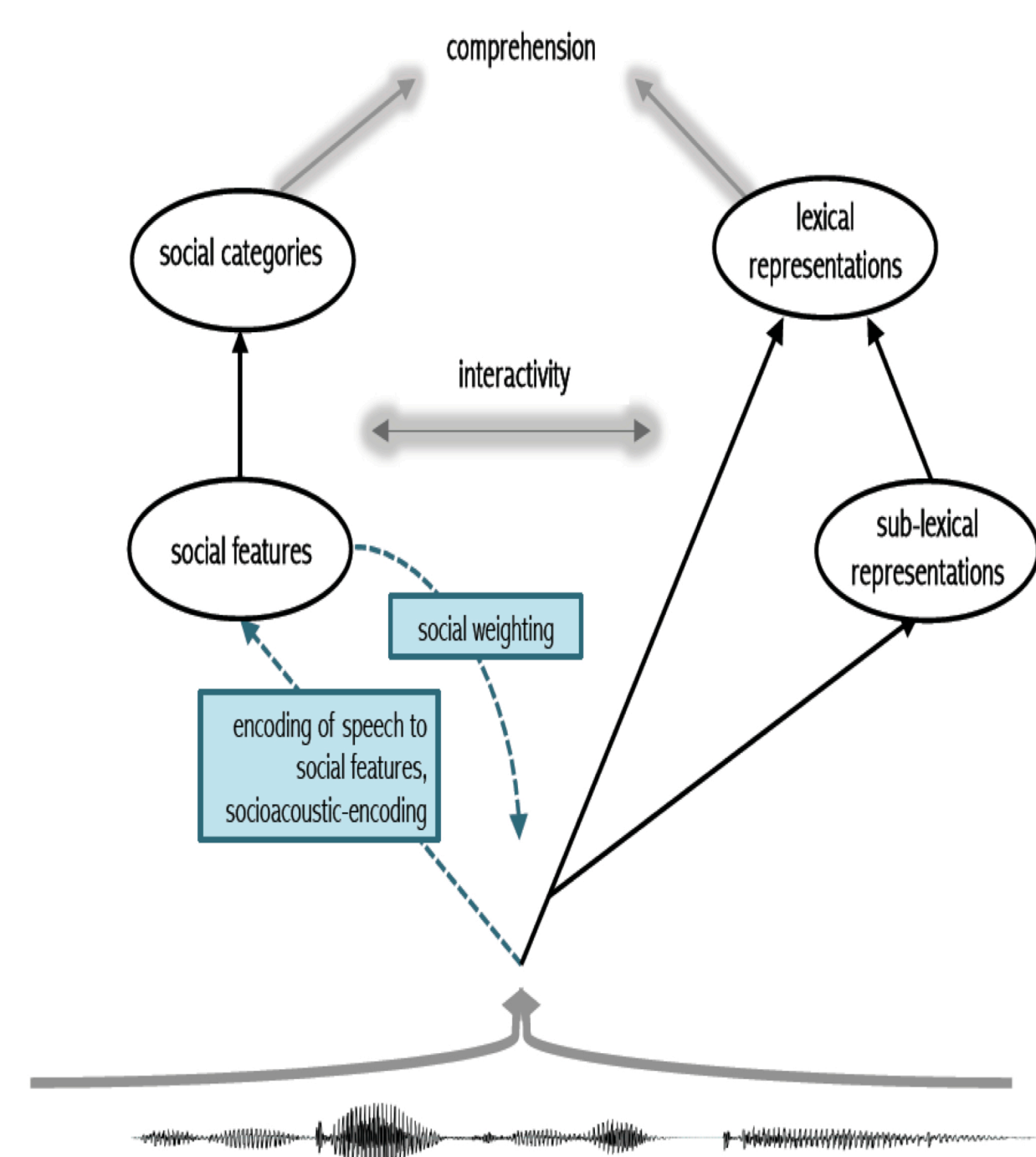
Recent work suggests that children use acoustic cues to talker identity to constrain comprehension of spoken language (Creel, 2012), **though the way in which children learn to integrate social knowledge with information from talker voice remains poorly understood.**

One Hypothesis

Children are able to disambiguate between objects with gendered associations (a men's pair of gloves and a women's pair of gloves, say) based on talker voice

Children should get better at using talker information in voice as a predictor of speaker meaning as they get older

One Model



A schematic of the model of encoding of social information during speech perception developed in Sumner, Kim, King, & McGowan (2013).

Language comprehension results from integrating the linguistic and social information in voice

Experiment 1: Forced Alternative on iPad

Do children use phonetically-cued talker information to infer speaker meaning?

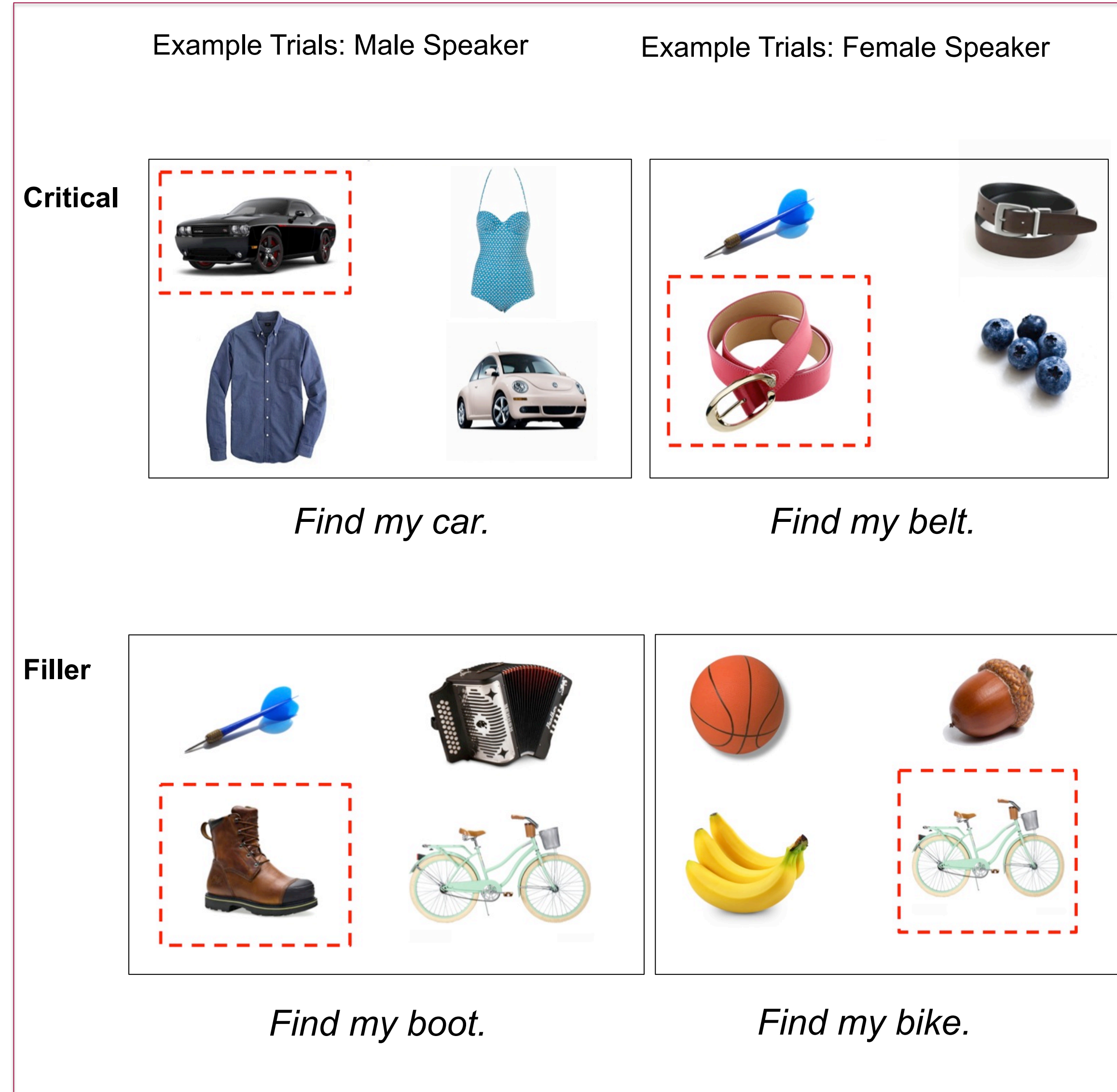
Participants: 41 children at Bing Nursery School, ages 3 to 5, counterbalanced for gender, counterbalanced for list

Between Subjects Design: children randomly assigned to one of 8 Male or Female speaker stimulus lists of 24 (+1 training filler) four-choice forced alternative slides.

Critical Trials: 12 slides with target word, two distractors, gender competitor

Filler Trials: 12 slides with target word, three distractors, no gender competitor

Visual Stimuli: Images normed by participants on Amazon Mechanical Turk, presented on an iPad. Trials presented in random order, with an extra filler trial always presented first.



Competitor Condition

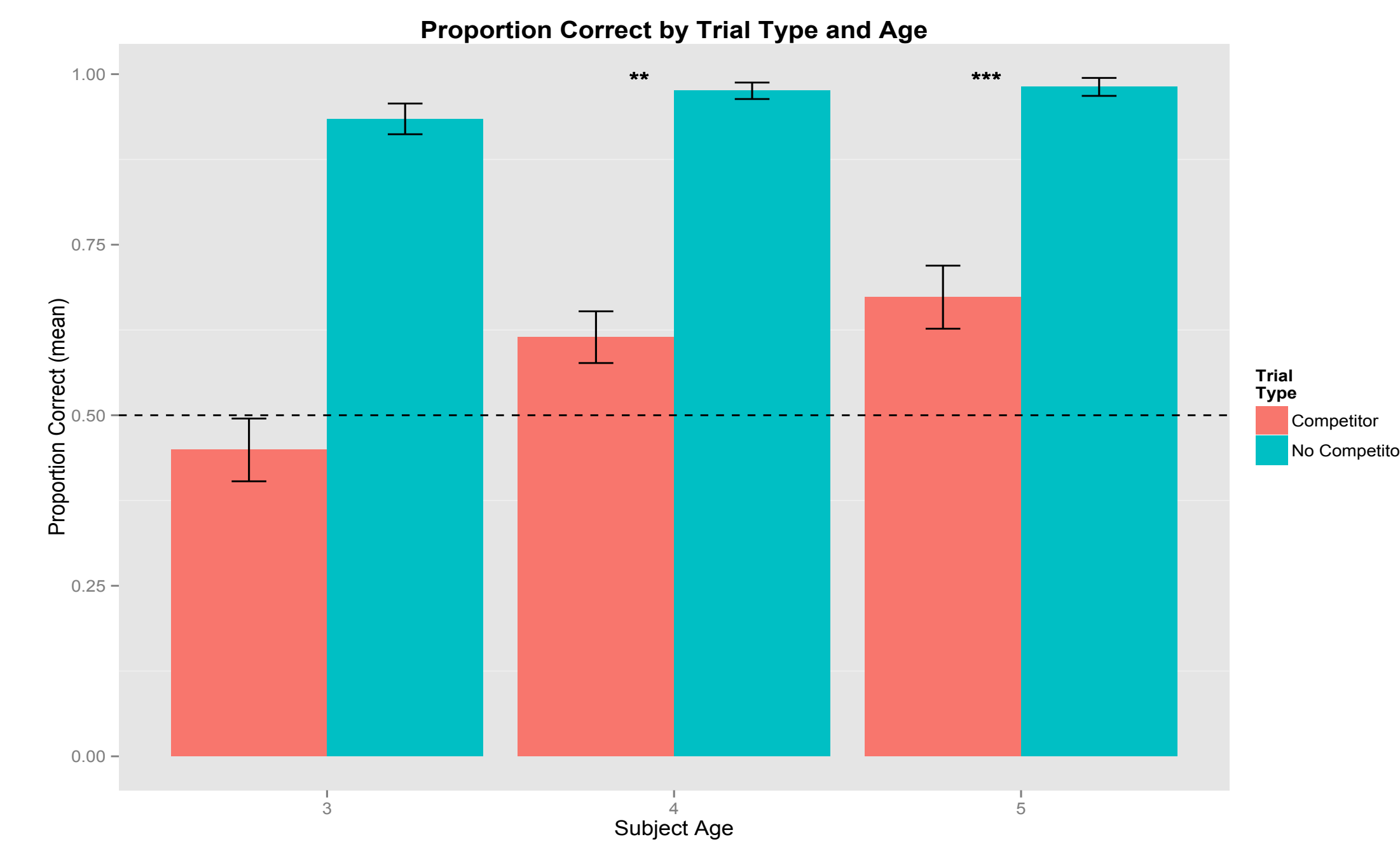
Trial featured both the target image and a competing variant that would stereotypically belong to a speaker of the opposite gender

(Ex: hearing a man's voice and seeing both a man's and woman's glove)

Non-competitor Condition

Trial featured only one image of the talker's referent. (Ex: hearing a man's voice and seeing a man's glove)

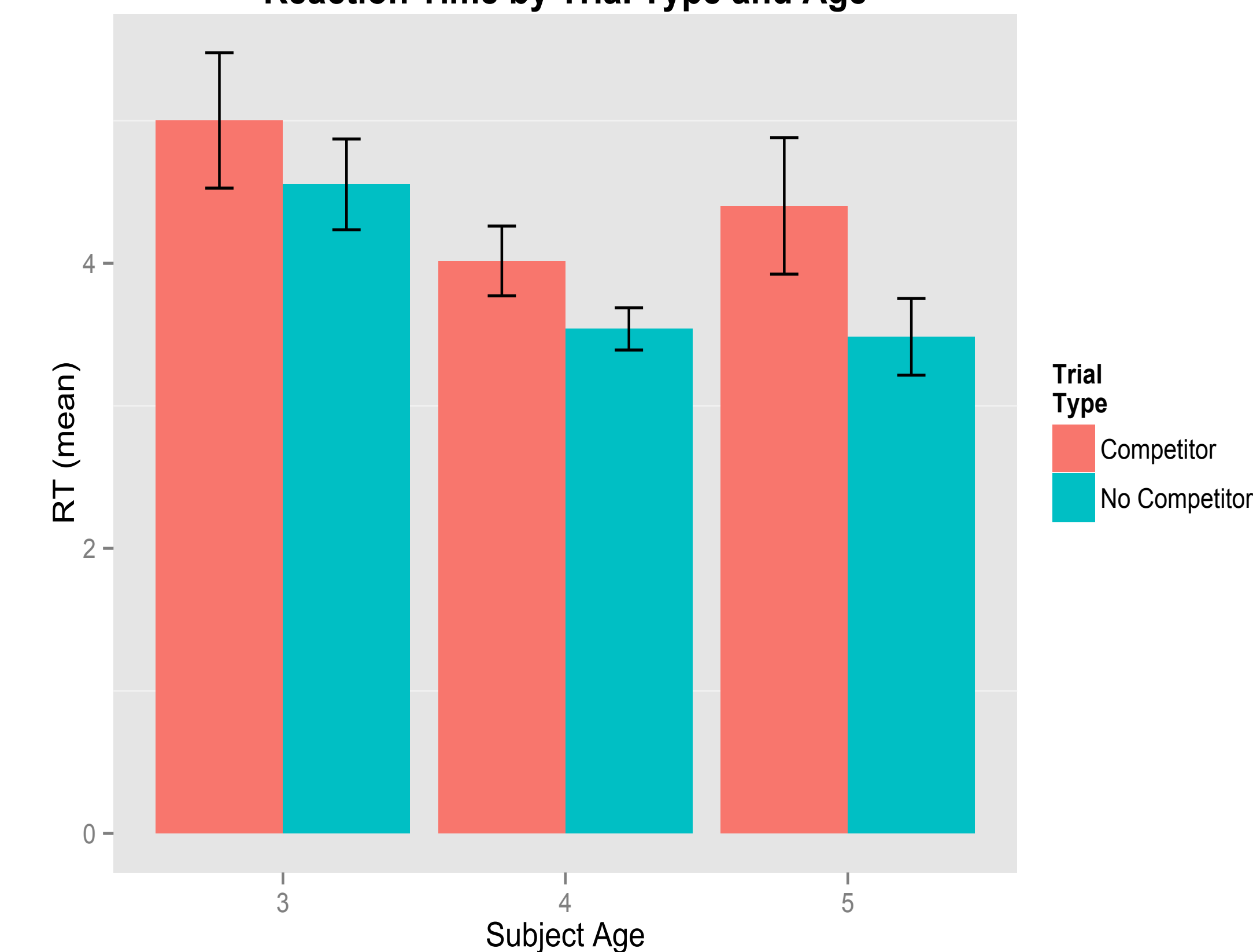
Results: Experiment 1



From ages 3 to 5, the ability to disambiguate increases significantly. By the age of 5, children regularly integrate phonetically-cued talker information with their social knowledge of speaker characteristics (4 year olds: $\beta=0.79$, $SE=0.18$, $z=2.794$, $p < 0.01$; 5 year olds: $\beta=1.10$, $SE=0.32$, $z=3.406$, $p < 0.001$). ($N=50$)

Female participants are significantly better at the task than male participants ($\beta=0.76$, $SE=0.24$, $z=3.101$, $p < 0.01$). The data suggest children were always faster to select during non-competitor trials than during competitor trials ($\beta=-0.48$, $SE=0.25$, $t(747)=-1.923$, $p=0.05$). ($N=50$)

Reaction Time by Trial Type and Age



Children's proportion correct click to target is at ceiling even from age 3. Additionally, there is no significant difference in children's performance as a function of speaker voice, whether male or female (figure not pictured).

Discussion

Strong Use of Talker Information to Facilitate Comprehension

- Children are able to integrate phonetically-cued socially indexical talker information in voice and their social knowledge of inferred speaker characteristics to guide their interpretation of speaker meaning in a real-world paradigm.
- Specific speaker gender played no role in determining whether children used the talker information in voice to infer the speaker's referent.
- Interestingly, in competitor trials, as age increases, children are both faster and more accurate in disambiguating the voice cues.
- These preliminary results suggest that children make use of socially-nuanced talker-specific acoustic information by a young age

Future Directions

We investigated at what age children are able to use talker indexical information in voice to infer speaker meaning, thus linking such talker information with children's growing knowledge of social categories.

Future work will attempt to tease apart what may be responsible for the significant coefficient for gender in the logistic mixed effect model of children's accuracy on the task.

In addition, future work can examine other stereotypes and try to determine what drives this change in performance from age 3 to age 5, and what form the data take at later ages.

References

- Creel, S. C., Aslin, R. N., & Tanenhaus, M. K. (2008). Heeding the voice of experience: The role of talker variation in lexical access. *Cognition*, 108, 633-664.
- Creel, S. C. (2012). Preschoolers' Use of Talker Information in On-Line Comprehension. *Child development*, 83, 2042-2056.
- Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of memory and language*, 59(4), 434-446.
- Johnson, K. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of phonetics*, 34, 485-499.
- Sumner, M., Kim, S. K., King, E., & McGowan, K. B. (2013). The socially weighted encoding of spoken words: a dual-route approach to speech perception. *Frontiers in psychology*, 4.
- 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, 580-591.