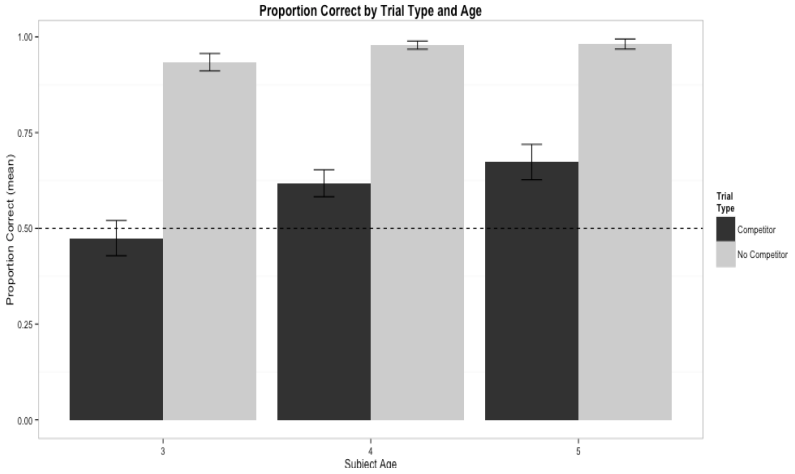
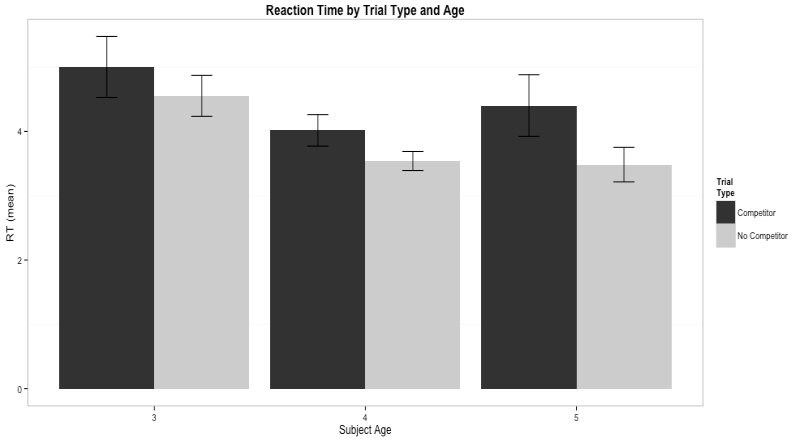
Children use phonetically-cued talker information to infer speaker meaning

Speech perception provides an avenue with which to study children's burgeoning ability to comprehend social information in the speech stream. The speech signal contains not only phonemic information necessary for word recognition but abundant "indexical information" about the talker, including gender (Perry, Ohde, & Ashmead, 2001; Goldinger, 1998; Johnson et al., 1999; Johnson, 2006). A growing body of work has begun to investigate the effects of listeners' associations between talkers' social identity and the phonetic cues in talker's voices, demonstrating that 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.





In this paper, we test the hypothesis that 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. We explore children's use of talker indexical information to infer speaker meaning through an experiment in which children interact with a web page on an iPad. Over 24 trials, children ages 3-5 were shown series of four images and asked by talkers to find one of the objects by clicking on it. During half of the trials, children heard a male talker's voice, and during the other half they heard a female talker's voice. In addition, half of the trials were non-competitor trials in which there was only one image of the talker's referent (hearing a man's voice and seeing a man's glove, say). The other half were competitor trials in which the target image competed with a variant which would stereotypically belong to a speaker of the opposite gender (hearing a man's voice but choosing between both a man's glove and a woman's glove). Children's reaction time between the utterance of the target word and their click on an image was logged, as was their choice of image.

Preliminary data show that by the age of 5, children regularly integrate phonetically-cued socially indexical talker information with their social knowledge of speaker characteristics to guide their interpretations of speaker meaning in a real world paradigm. From ages 3 to 5, the ability to disambiguate increases significantly. All ages are extremely good at the task in non-competitor trials. Interestingly, in competitor trials, as age increases, children are faster and more accurate in disambiguating the voice cues. These results suggest not only that children make use of socially-nuanced talker-specific acoustic information by a young age, but reveal a robust understanding of gender stereotypes that guide their daily interactions with interlocutors and may bear on their linguistic and social development.

Word Count = 480

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.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological review*, *105*, 251.

Johnson, K. A., Strand, E. A., & D'imperio, M. (1999). Auditory-visual integration of talker gender in vowel perception. *Journal of Phonetics*, *24*, 359-384.

Johnson, K. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of phonetics*, *34*, 485-499.

Perry, T. L., Ohde, R. N., & Ashmead, D. H. (2001). The acoustic bases for gender identification from children's voices. *The Journal of the Acoustical Society of America*, *109*, 2988-2998.

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.