# Gaze Distribution Patterns for Audiovisual Speech Stimuli in Preschool Children with and without Hearing Loss

Nicholas A. Smith, Mark VanDam & Mary Pat Moeller

Boys Town National Research Hospital, Omaha, Nebraska

**4pPSC**

158th Meeting of the Acoustical Society of America, San Antonio Texas, October 2009

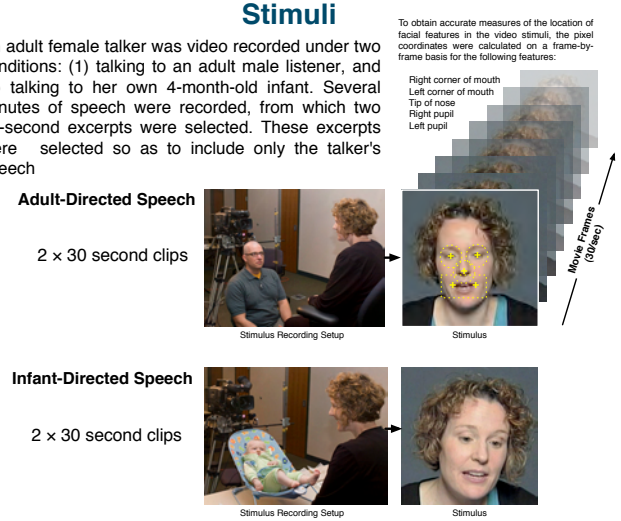
## Introduction

Speech provides both auditory and visual information to listeners. Many studies have demonstrated how visual information affects speech perception (Sumby & Pollack, 1954; McGurk & McDonald, 1976), but much less is known about how this visual information is processed by perceivers. The goal of this study was to examine the looking behavior of preschool-aged children while they watch audiovisual speech, and to examine how this behavior is affected by hearing loss. Specifically, we asked how children distribute their visual gaze fixations to different facial features of the talker when viewing audiovisual samples of speech directed to infants and adults.

Previous studies have shown that adults direct their gaze to the upper part of the face when making judgments about intonation and emotion and to the lower face when making segmental judgments or perceiving speech in noise (Buchan et al, 2007; Lansing & McConkie, 1999; Vatikiotis-Bateson et al., 1998). It was therefore predicted that children would fixate on the eyes to a greater degree in the prosody-rich infant-directed speech condition. It was also predicted that children with hearing loss may rely on visual information around the mouth to a greater degree that do children with normal hearing.

## Stimuli

An adult female talker was video recorded under two conditions: (1) talking to an adult male listener, and (2) talking to her own 4-month-old infant. Several minutes of speech were recorded, from which two 30-second excerpts were selected. These excerpts were selected so as to include only the talker's speech




## Subjects

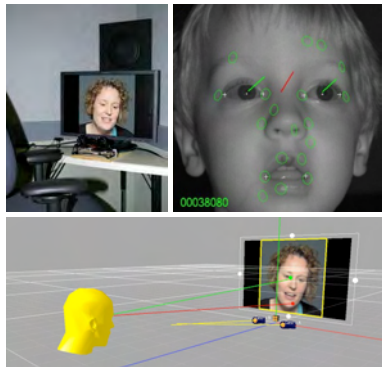
Ten children were recruited from the Boys Town Preschool Program, which comprises children who are deaf and hard of hearing, who use cochlear implants (CIs), as a well as their peers with normal hearing.

**Normal Hearing**: n = 5, mean age = 3.7 years

**CI users:** n = 5, mean age = 3.8 years
Better ear PTA = 94 dB

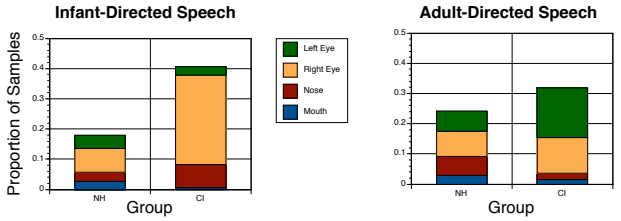
## Procedure

Video stimuli were presented on a 26" 1920 × 1200 pixel LCD display, while children's eye movements were recorded using faceLAB 4 eye-tracking system. This system logged gaze screen intersection in terms of screen pixel coordinates at rate of 60 samples/second. The audio signal was presented at 65 dB SPL through an audiometric loudspeaker located behind and above the display.




## Results

The series of coordinates for each talker facial feature were resampled using a nearest neighbor interpolation procedure, thus bringing stimuli and recordings into a common temporal framework for analysis. For each sample of eye tracking data, the distance between the subjects gaze screen intersection and each facial feature in the stimulus was calculated. Subjects were considered to be looking a facial feature when subjects' gaze fell within 50 pixels of the location of the feature on the current video frame. The figure below shows the mean proportion of samples falling within the areas surrounding each facial feature.




## Conclusions

Cochlear implant users spent more time fixating on facial features than did children with normal hearing.

Children's gaze is directed to the upper portion of the face more than half the time, particularly in the infant-directed speech condition.

All children showed an assymmetrical pattern of fixations to the talker's right versus left eye, as reported previously with adult viewers (Everdell et al. 2007).

The talker's nose, perhaps due to its centrality was a greater focus of fixation than the mouth.

Children with cochlear implants adopt different visual strategies when perceiving speech, compared to children with normal hearing. It remains to be seen whether these strategies facilitate speech perception.

## References

Buchan, J. N., Paré, M. & Munhall, K. G. (2007). Spatial statistics of gaze fixations during dynamic face processing. *Social Neuroscience, 2,* 1-13.

Everdell, I.T., Yurick, M. D., Munhall, K. G. & Paré, M. (2007). Gaze behaviour in audiovisual speech perception: Asymmetrical distribution of face-directed fixations. *Perception, 36,* 1535-1545.

Sumby, W. H. & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *JASA,* 26, 212-215.

McGurk, H. & McDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264,* 746-748.

Lansing, C. R. & McConkie, G. W. (1999). Attention to facial regions in segmental and prosodic visual speech perception tasks. *Journal of Speech, Language and Hearing Research, 42,* 526-539.

Vatikiotis-Bateson, E., Eigsti, I. Yano, S. & Munhall, K. G. (1998). Eye movement of perceivers during audiovisual speech perception. *Perception & Psychophysics, 60,* 926-940.

**Acknowledgements**

We thank Colleen Shane, for assistance in collecting these data, as well Amy Tyler Krings, Kris Mixan, Cathy Carotta, and the preschool staff for supporting this project.