Do Toddlers Experience the McGurk Illusion? An Eye-Tracking Study

Catherine Kraper, Zsuzsa Kaldy, Ph.D., Emily Higgins, Erik Blaser, Ph.D., & Alice S. Carter, Ph.D.

University of Massachusetts Boston

Introduction

- While understanding spoken language does not require visual input, our daily experience of speech integrates auditory and visual information.
- This automatic process is apparent in illusions such as the McGurk effect, in which an auditory syllable ("ba") is mismatched with a visual syllable ("ga") mouthed by a speaker, creating a perceived syllable "fusion" - a sound ("ga" or "tha") that combines the auditory and visual input (McGurk & MacDonald, 1976). The McGurk effect has been shown to occur in infants as young as 4.5-5 months (Burnham & Dodd, 2004; Kushnerenko et al., 2008).
- The current study expanded on Burnham and Dodd's looking-time paradigm in two important ways:
  1. We studied toddlers with emerging language skills; this is the first study to our knowledge to do so;
  2. We used eye-tracking to precisely determine fixation length and location.

Hypothesis:

We hypothesize that typically developing toddlers experience a robust, consistent McGurk illusion that is reflected in their looking behavior.

Methods

Participants: 15 healthy toddlers (age: 14-31 months (mean = 28±4.9 months). Language abilities were assessed using the Mullen Scales of Early Learning (mean EL T-score = 54.3; mean RL T-score = 58.5; measure has a mean T-score of 50±10)

Procedure: We presented the children with a looped video that paired an unfamiliar face mouthing a single syllable ("ga") dubbed with a different syllable sound ("ba"). Participants were shown 4 such "fusion" familiarization trials, followed by 2 test trials. All trials lasted for 20 sec; between trials, a dynamic stimulus was shown for 5 sec to draw toddlers' attention to the center of the screen. In test trials, the same face was seen, but importantly, now static, and was paired with either an auditory "ba" or "tha" sound. Test trial order was counterbalanced across subjects. We used a **Tobii T60** eye-tracker to measure fixation duration and time-to-first-fixation to the mouth region of the face during test trials.

Analysis: Areas of Interests (AOI's) for the mouth and face were defined (Fig. 1). *T*-tests were run to determine if there were significant differences between the two test trials ("ba" vs. "tha") in either time to first fixation or looking times (fixation length) to the mouth AOI (Fig. 2 and 3).

Figure 1. Areas of Interest: mouth and face

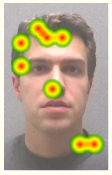

Figure 2. Fixation length heatmap, 'tha' test trial

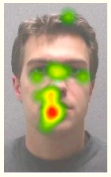

Figure 3. Fixation length heatmap, 'ba' test trial

Results

Familiarization Trials: In the four familiarization trials, participants displayed a trend of decreasing fixation length to the face region (Fig. 4).

Test Trials: In independent-samples *t*-tests, participants were significantly faster at looking to the mouth region in "ba" trials than in "tha" trials ($t(13) = 2.373$, $p < 0.034$) (Fig. 5). Two outliers (time to first fixation > 3 SD away from mean, suggesting the participants were not attending to the task) were dropped in the analysis.

Participants did not look significantly longer to the mouth region at the "ba" versus the "tha" trials ($t(27) = 1.472$, $p < 0.153$); however, there was a trend toward longer fixation length on the mouth in the "ba" trials.

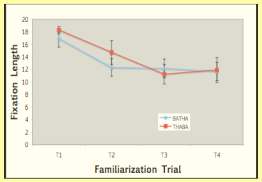

Figure 4. Fixation length on face during familiarization trials

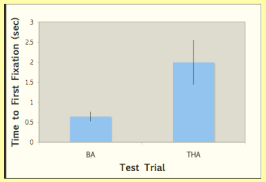

Figure 5. Time-to-first fixation to the mouth area in test trials

Discussion

We familiarized participants to the combined audio-visual stimulus that induces the classic McGurk illusion in adults and then tested their reactions to the "fused" ("tha") vs. the non-fused ("ba") audio stimulus. Toddlers looked significantly faster to the mouth region in "ba" test trials than in "tha" test trials, suggesting that they perceived the "ba" sound as novel. Our results therefore suggest that typically developing toddlers do experience the McGurk illusion, indicating that toddlers utilize the auditory and visual components of a display to experience language.

We compared looking at the mouth region to looking at other salient features of the face. Participants fixated longer on the mouth than the eyes in both types of test trials - even though the entire face was static at that point ('ba': $t(10) = 4.149$, $p < 0.002$, 'tha': $t(12) = 1.972$, $p < 0.072$, two-tailed *t*-tests).

Studies of early language development suggest that infants at 6 months are able to discriminate phonemes, when the phonemes were originally paired with a correct visual stimulus, but were less consistent when the phonemes had been paired in half the familiarization trials with an incorrect visual stimulus (Teinonen et al., 2008). This suggests that visual information significantly informs the process of speech development from an early age.

This study design closely mirrored that of Burnham and Dodd (2004). An important difference between the studies was the use of eye-tracking, which allowed for more precise information regarding the location and length of participants' fixations. **Our design does not require that children understand verbal instructions, and is therefore particularly well-suited for toddlers or children with developmental delays.**

References

Burnham, D., & Dodd, B. (2004). Auditory-visual speech integration by prelinguistic infants: Perception of an emergent consonant in the McGurk effect. *Developmental Psychobiology, 45*, 204-220.

Kushnerenko, E., Teinonen, T., Volein, A., & Csibra, G. (2008). Electrophysiological evidence of illusory audio-visual speech perception in human infants. *Proceedings of the National Academy of Sciences, 105*, 11442-5.

McGurk H., & MacDonald, J.W. (1976). Hearing lips and seeing voices. *Nature, 264*, 746-8.

Teinonen, T., Aslin, R.N., Alku, P., & Csibra, G. (2008). Visual speech contributes to phonetic learning in 6-month-old infants. *Cognition, 108*, 850-5.