**Audiovisual Speech Perception in Children with Autism Spectrum Disorders**

**Julia R. Irwin**

Haskins Laboratories, New Haven, CT

---

**BACKGROUND**

For typically developing perceivers, visual information assists in the recognition of speech in noise (Sueur & Pollack, 1994). Visual information is also used in the perception of unambiguous auditory speech (Irwin et al., 2006). One powerful demonstration of the influence of visual information is the McGurk effect (McGurk & MacDonald, 1976). Autism Spectrum Disorders (ASD) refer to a continuum of severe neurodevelopmental disorders characterized by marked deficits in social reciprocity, communication, and the presence of restricted or repetitive behaviors. Children with autism spectrum disorders (ASD) appear to be less influenced by visual speech information (e.g., De Gelder et al., 1991; Massaro & Boessler, 2003; Williams et al., 2006).

Given the tendency of individuals with ASD to avert gaze from the faces of others, attenuated visual influence on heard speech in ASD may reflect less visual fixation on the face of a speaker. Alternatively, children with ASD may have an underlying weakness in perception of AV speech.

In a first step to investigate these two possible underlying causes of atypical sensitivity to visual speech in ASD, we compared visual influence on heard speech in children with ASD and typically developing (TD) controls when the children were fixated on the face of a speaker.

**METHOD**

**Participants**

- 8 ASD children (3 boys, mean age 9.2 years, range 8-13 years)
- 3 TD children (2 girls, 1 boy, mean age 9.3 years, range 7.5-10.5 years)

- Normal hearing and normal or corrected vision
- Native speakers of American English
- ASD participants met criteria for autism on the ADOS, ADI-R and by clinical diagnosis.

**Stimuli**

A male native English speaker was videotaped producing the consonant-vowel syllables /ma/, /na/, and /ga/. These syllables were digitally edited to create matched and mismatched auditory-visual (AV) stimuli:

- Matched, cross-spliced syllables /ma/ + /na/
- Mismatched syllables, auditory /ma/ + visual /na/ + /ma/

**Procedure**

- Participants were asked to report what they heard.
- Eye gaze was tracked during presentation of the matched and mismatched (McGurk) speech stimuli.
- A priori “look zones” were created to assess location of gaze on the face of the speaker.
- Only those trials where the participant was fixated on the face of the speaker were included in analyses.

**Stimuli**

<table>
<thead>
<tr>
<th>Signal</th>
<th>Percept</th>
</tr>
</thead>
<tbody>
<tr>
<td>/ma/1 + /na/</td>
<td>/na/</td>
</tr>
<tr>
<td>/ma/1 + /ma/2</td>
<td>/ma/</td>
</tr>
<tr>
<td>/ma/1 + /ga/3</td>
<td>/ma/</td>
</tr>
</tbody>
</table>

**Data from this pilot project suggest:**

- Children with ASD are significantly less visually influenced than TD controls, even when fixated on the face of the speaker.
- Children with ASD may have an underlying weakness in perception of AV speech.
- Patterns of gaze to a speaking face do not differ between children with ASD and TD controls.

**Conclusions**

We are currently comparing children with ASD to both chronological age and verbal mental age matched TD controls on:

- Visual influence on heard speech with matched and mismatched (McGurk) AV speech
- Visual gain from auditory to AV speech in noise
- Sensitivity to asynchrony in AV speech
- Identification of visual only (lipread) speech stimuli

**References**


