Do children with autism spectrum disorder use prosodic cues to infer others’ mental states?
Peng Zhou (Tsinghua University), Weiyi Ma (University of Arkansas) & Likan Zhan (Beijing Language and Culture University)
zhoupeng1892@mail.tsinghua.edu.cn

**Background** Unravelling the kinds of cues children with autism spectrum disorder (ASD) can and cannot utilise when inferring others’ mental states is crucial to understanding the nature of their core deficits in social communication. The paper reports two eye-tracking studies that investigated whether preschool Mandarin-speaking children with ASD use prosodic cues to infer others’ mental states – communicative intentions (Study 1) and emotional states (Study 2).

**Participants** Both studies tested the same participants including 36 high-functioning children with confirmed ASD diagnoses and 25 typically developing (TD) peers. Table 1 shows the descriptive characteristics of the participants. Note that the 5-year-olds with ASD were matched with the TD 4-year-olds on both verbal-IQ and MLU. The participants were all native speakers of Mandarin, and had no reported history of hearing or visual impairment.

**Methods** Both studies used the visual world paradigm (Tanenhaus et al., 1995). The participants were simply told to listen to spoken sentences while looking at visual images. We measured their eye movements that arose as automatic responses to the test stimuli. In Study 1, the participants heard sentences as in (1) presented in two prosodic versions ((1a) and (1b)) while viewing visual scenes as in Fig1. In Study 2, they heard sentences as in (2) produced in two prosodic versions ((2a) and (2b)) while viewing visual images as in Fig2. In Study 1, manipulation of prosody distinguishes between two speech acts: rising intonation on the wh-phrase *shenme dongwu* ‘what animal’ indicates a question reading, whereas level intonation on the same wh-phrase indicates a statement reading. In Study 2, manipulation of prosody indicates two emotional states: high pitch, strong intensity and fast tempo indicate positive emotion (happiness), whereas the opposite features indicate negative emotion (sadness).

**Analyses and Results** When analysing the eye movement data, we divided each visual image into different areas of interest (AoSs). The two critical AoIs in Study 1 were the question-compatible area (e.g., the dog and the monkey in Fig1) and the statement-compatible area (e.g., Xiaoming and the horse). The two AoIs in Study 2 were the two areas containing the happy face and the sad face respectively. Fixation proportions on a particular AoI in a specific temporal bin were treated as the dependent variable. To statistically assess the eye movement patterns, we then transformed the fixation proportions using the empirical logit formula (Barr, 2008). In both studies, a series of linear mixed-effects models were then fitted to the transformed probability for each participant group. The best-fitting model results of Study 1 show that in both AoIs, significant effects of group were observed. The eye movement patterns of the TD 4-year-olds reflected their ability to effectively use prosodic cues to distinguish between the two speech acts, whereas the age-matched and the verbal-IQ-matched ASD groups failed to do so. The best-fitting model results of Study 2 found no significant effect of group. Like the TD group, the two ASD groups exhibited eye movement patterns that reflected their ability to use prosodic cues to infer others’ emotional states.

**Conclusions** The findings point towards a deficit in ASD children’s ability to use prosodic cues to understand others’ communicative intentions, but suggest a relatively preserved ability to use prosodic cues to infer others’ emotional states. We discuss the discrepancy between the two types of mental states in ASD children, as well as the implications in relation to their core deficits in the social and communicative domain.
Table 1
Verbal IQ scores and MLU of each participant group (SD in parentheses)

<table>
<thead>
<tr>
<th>Group</th>
<th>Number</th>
<th>Verbal IQ</th>
<th>MLU</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASD group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-year-olds</td>
<td>18</td>
<td>95.56 (9.46)</td>
<td>4.77 (1.25)</td>
</tr>
<tr>
<td>5-year-olds</td>
<td>18</td>
<td>101.17 (9.53)</td>
<td>5.92 (1.26)</td>
</tr>
<tr>
<td>TD group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-year-olds</td>
<td>25</td>
<td>101.28 (9.02)</td>
<td>5.89 (1.22)</td>
</tr>
</tbody>
</table>

(1) Xiaoming meiyou mo shenme dongwu
   Xiaoming not pat what animal
   (1a) Rising intonation - Question reading: What animal did Xiaoming not pat?
   (1b) Level intonation - Statement reading: Xiaoming didn’t pat any animal.

(2) Zhe shi yi-ge laji tong.
   this is one-CL trash can
   ‘This is a trash can.’
   (2a) Prosody with high pitch, strong intensity and fast tempo – positive emotion (happiness)
   (2b) Prosody with low pitch, weak intensity and low tempo – negative emotion (sadness)

Fig 1. An example test image in Study 1

Fig 2. An example test image in Study 2

References