Department of Foreign Languages and Literatures

Child Cognition Lab, Tsinghua University

Beijing 100084, China

[zhoupeng1892@mail.tsinghua.edu.cn](mailto:zhoupeng1892@mail.tsinghua.edu.cn)

4 July, 2018

Dear Editor,

Thank you very much for the editorial comments on our second revised manuscript. The following is a summary of how we have responded to your thoughtful comments.

**(1) In response to the 1st comment: Please remove the trademark and any commercial language.**

We have now removed the trademark.

**See lines 102-104:** “1.3 Measure the verbal IQ of the participants using the Wechsler Preschool and Primary Scale of Intelligence - IV (CN) – a standardized IQ test designed for Mandarin-speaking children between the ages of 2; 6 and 6;1143.”

**(2) In response to the 2nd comment: Step 3.1, 3.1.1, 3.1.2, 3.2: For steps that are done using software, a step-wise description of software usage must be included in the step. Please mention what button is clicked on in the software, or which menu items need to be selected to perform the step. The current step is too general, therefore, it will be hard for the readers to replicate this step.**

We have now included a step-wise description of software usage whenever possible. We also wish to note that there are different options available to construct test stimuli for a visual world study. The present study just provides one option.

**See lines 137-147:** “3.1 Construct test stimuli. Create 12 target items, each comprised of a visual stimulus, and two spoken sentences containing the morphological markers BA and BEI respectively. Use Pixelmator (or other image editors) to create visual images. Open Pixelmator. Click on the Pixelmator icon. Create a visual image from a template. Click “Show Details” in the template chooser. Double-click the template to open it. Adjust the width, height, resolution, and color depth from the pop-up menus. Enter the relevant parameters. Click OK. Use Praat (or other audio editors) to construct spoken sentences. Set the microphone. Open Praat. Click on the Praat icon. Select “Record Mono sound” from the “New” menu. Set the recording conditions by clicking the sample rate option of 22050. Click the “Record” button. Record the spoken sentences by asking a native Beijing Mandarin-speaker to produce the sentences in a child directed manner. Save the recordings by clicking “Save”. ”

**(3) In response to the 3rd comment: Step 4.1-4.3: Details are missing in these steps. All details should be included in the steps instead of the notes, otherwise, the steps cannot be filmed. Please write the steps in imperative tense.**

We have now added details about these steps and we have moved most of the information from notes to the steps using imperatives.

**See lines 193-220:** “4.1 Eye-tracking procedure. Invite the participants to sit comfortably in front of the display monitor of the remote eye tracker. Set the distance between the participants’ eyes and the monitor around 60 cm. Perform the standard calibration and validation procedures by asking the participants to fixate on a grid of five fixation targets in random succession. Present the participants with a spoken sentence while they are seeing a visual image, as in the standard visual world paradigm10, 44. Use the monocular eye tracking option by tracking the eye that is on the same side as the illuminator of the eye tracker. Record the participants’ eye movements using the eye tracker.

Note: The eye tracker used in the present study allows remote eye tracking with a sampling rate of 500 Hz.

4.2 Testing and measuring. Test the participants individually. Simply tell the participants to listen to the spoken sentences while they are viewing the pictures. Ask one experimenter to monitor the participant on the computer and one to stand behind the participant and gently rest her hands on the participant’s shoulders to minimize the participant’s sudden movements. Measure the participant’s eye movements that arise as automatic responses to the linguistic input using the eye tracker.

Note: The task does not ask the participants to make any conscious judgments about the spoken sentences, so as to minimize the computational burden of the participants. The eye tracker automatically records the participants’ eye movements

4.3 Monitoring during the test. Use the live viewer mode on the computer screen, exhibited by the eye tracker during the test, to observe the participant’ looking behavior. Ask the experimenter who monitors data collection via the live viewer mode to signal to the experimenter who stands behind the participant to reorient the participant if the participant’ eye gaze wanders off computer screen.”

**(4) In response to the 4th comment: Step 5.1-5.2: These steps cannot be filmed unless a step-wise description of software usage is provided.**

We have now added a step-wise description of software usage. We also wish to note that different eye trackers use different software for coding and analyzing data. The present study just provides one option.

**Lines 266-286:** “5.1 Code the participants’ fixations in two interest areas. Use Data Viewer to draw the two interest areas: BA-target event area and BEI-target event area (see Figure 1). Open Data Viewer. Select one of the interest area shape icons on the tool bar. Use the mouse to drag a box around the region you want to define as an interest area. Save the interest area in the “Interest Area Set folder”. Apply the interest area to other visual images.

Note: The depicted event in the upper panel of Figure 1 matches the BA-construction, and hence the BA-target event, and the event depicted in the lower panel matches (1b), and hence the BEI-target event. The software used for data coding is Data Viewer that comes with the eye tracker used in the study. Other data analysis software is also available.

5.2 Analyze the eye gaze patterns using Data Viewer. Open Data Viewer. Choose the sample report function from the menu to set the time windows for analysis (e.g., every 200 ms as a time window in the present study). Use the same function to time lock the fixation proportions in the interest areas to the onset of the marker for each trial. Export the raw data into an excel file using the export function from the menu. Use the excel functions to average the fixation proportions following the onset of the marker for each area. Use the excel functions to compute the fixation proportions in each time window of 200 ms over a period of 5200 ms (the mean length of the target sentences plus 200 ms) from the onset of the marker for the two areas. Apply linear mixed-effects models to the eye movement data, detailed in the section “Representative Results”.”

**(5) In response to the 5th comment: Line 247-267, 279-288: These text shows overlap with previous publication. Though there may be a limited number of ways to describe a technique, please use original language throughout the manuscript.**

As you noted, it is fairly hard to use completely original language to describe a technique, in particular, in a way that is different from the authors’ own publications. But we have tried to rewrite the sections to avoid significant overlap.

**See lines 306-368:** “The figures show that the autism group displayed eye movement patterns similar to the age-matched TD group. Both groups exhibited more fixations on the BA-target event when hearing BA than when hearing BEI, occurring after the onset of the object NP and before the onset of the adverb. To be specific, the effect occurred in the TD group in the time window between 1400-1600 ms (see Figure 2). The effect occurred in the autism group in the time window between 1800-2000 ms (see Figure 3). By contrast, an opposite eye movement pattern was found in the BEI-target event for both groups: more fixations on the BEI-target event were observed when hearing BEI than when hearing BA, again occurring after the onset of the object NP and prior to the onset of the adverb.

Fixation proportions were then transformed using the empirical logit formula49: *probability=ln((y+0.5)/(n-y+0.5))*, where *y* is the number of fixations on the areas of interest during a particular temporal bin; *n* is the total number of fixations in that temporal bin. Linear mixed-effects models were then fitted to the transformed data. Statistical models were then computed for the two groups separately on their fixations in the two interest areas in the critical time windows, where time and marker type (BA versus BEI) were treated as fixed effects; random intercepts and slopes were included for both participants and items50. The fitting process was conducted via functions lmer from package lme4 (v1.1-12)51 of the R (v3.2.5) software environment52. A Wald test was then used to compute *p-*values for each fixed effect.

**The model results for the TD 5-year-olds in the two interest areas**: In the BA-target event area, hearing BA led the TD children to look significantly more at this event than hearing BEI (*β*= 0.54, *p* < .001). In addition, there was a significant interaction between marker type and time (*β*= 0.33, *p* < .001), indicating that the probability of fixating on the BA-target event increased over time after the onset of BA. However, the TD children exhibited an opposite eye movement pattern in the BEI-target event area. Hearing BEI triggered more fixations on the BEI-target event than hearing BA (*β*= -0.60, *p* < .001). Again, there was a significant interaction between marker type and time (*β*= -0.21, *p* < .001), suggesting that their tendency to look at the BEI-target event declines over time after the onset of BA.

**The model results for the autism 5-year-olds in the two interest areas**: The autism group showed similar eye movement patterns. Hearing BA triggered more fixations on the BA-target event than hearing BEI (*β*= 0.50, *p* < .001). Hearing BEI triggered more looks at the BEI-target event than hearing BA (*β*= -0.54, *p* < .001). Like the TD group, the autism group exhibited significant interactions in both interest areas. In both the BA-target event area and the BEI-target event area, they displayed a significant interaction between marker type and time (*β*= 0.15, *p* < .01 in the BA-target event area; *β*= -0.16, *p* < .01 in the BEI-target event area).”

**(6) We have also added a funding source.**

**See lines 569-573:** “This work was funded by the National Social Science Foundation of China [16BYY076] to Peng Zhou, and the Science Foundation of Beijing Language and Cultural University under the “the Fundamental Research Funds for the Central Universities” [15YJ050003]. The authors are grateful to the children, the parents and the teachers at the Enqi Autism Platform and at the Taolifangyuan Kindergarten, Beijing, China, for their support in running the study. ”

Again, we wish to thank you for your careful reading and thoughtful comments.

Best,

Peng Zhou