Journal of Visualized Experiments

Using the Visual World Paradigm to Study Real-Time Sentence Comprehension in Mandarin-speaking Children with Autism
--Manuscript Draft--

TITLE:

Using the Visual World Paradigm to Study Sentence Comprehension in Mandarin-Speaking Children with Autism

AUTHORS & AFFILIATIONS:

Peng Zhou[1*], Weiyi Ma[2*], Likan Zhan[3*], Huimin Ma[4*]

[1]Department of Foreign Languages and Literatures, Tsinghua University, Beijing, China

[2]School of Human Environmental Sciences, University of Arkansas, Fayetteville, AR, USA

[3]Institute for Speech Pathology and the Brain Science, Beijing Language and Culture University, Beijing, China

[4]Department of Electronic Engineering, Tsinghua University, Beijing, China

*These authors contributed equally to this work

Corresponding Author:

Peng Zhou   (zhoupeng1892@mail.tsinghua.edu.cn)

Email Addresses of Co-authors:

Weiyi Ma   (weiyima@uark.edu)

Likan Zhan   (zhanlikan@hotmail.com)

Huimin Ma   (mhmpub@mail.tsinghua.edu.cn)

KEYWORDS:

Eye-tracking, visual world paradigm, linguistic cues, morphological markers, child language, real-time sentence comprehension

SUMMARY:

We present a protocol to examine the use of morphological cues during real-time sentence comprehension by children with autism.

ABSTRACT:

Sentence comprehension relies on the ability to rapidly integrate different types of linguistic and non-linguistic information. However, there is currently a paucity of research exploring how preschool children with autism understand sentences using different types of cues. The mechanisms underlying sentence comprehension remains largely unclear. The present study presents a protocol to examine the sentence comprehension abilities of preschool children with autism. More specifically, a visual world paradigm of eye-tracking is used to explore the moment-to-moment sentence comprehension in the children. The paradigm has multiple advantages. First, it is sensitive to the time course of sentence comprehension and thus can provide rich information about how sentence comprehension unfolds over time. Second, it requires minimal task and communication demands, so it is ideal for testing children with autism. To further minimize the computational burden of children, the present study measures eye movements that arise as automatic responses to linguistic input rather than measuring eye movements that accompany conscious responses to spoken instructions.

INTRODUCTION:

Sentence comprehension relies on the ability to rapidly integrate different types of linguistic and non-linguistic information[1-11]. Prior research has found that young typically developing (TD) children incrementally compute the meaning of a sentence using both linguistic and non-linguistic cues[12-19]. However, there is currently a paucity of research exploring how preschool children with autism understand a sentence using different types of cues. The mechanisms underlying their sentence comprehension remains largely unclear.

It is generally acknowledged that there is enormous variability in the language abilities of children with autism, especially in their expressive language; for example, some children with autism have relatively good structural language, some exhibit deficits in both lexical and grammatical domains, some demonstrate impaired grammar, and some never acquire functional spoken language[20-25]. In addition, prior research seems to suggest that their receptive language is relatively more impaired than their expressive language[26-29]. Most research that has assessed sentence comprehension abilities of children with autism have used offline tasks (*e.g.,* standardized tests, caregiver reports), and the findings suggest that their sentence comprehension abilities might be particularly impaired[30-37]. However, it has been pointed out that poor comprehension abilities are more likely related to these children's overall lack of social responsiveness than to language processing deficits[38, 39]. Note that these offline tasks used in previous research often require high response demands or interactions with the experimenters, which might pose particular difficulties for children with autism, because they often exhibit various challenging behaviors or symptoms. As a result, this may interact with the high task and communication demands and mask their comprehension abilities [for an overview of methods for assessing receptive language in children with autism, see Kasari et al. (2013)[27] and Plesa-Skwerer et al. (2016)[29]]. Thus, experimental paradigms that can better control these confounding factors are required to further understand the nature of sentence-processing mechanisms in autism.

In the current study, we present an eye-tracking paradigm that can directly and effectively assess sentence comprehension abilities of children with autism. Compared to offline tasks, eye-tracking is a more sensitive testing paradigm to demonstrate children's comprehension abilities. It is sensitive to the time course of the comprehension process and requires no explicit motor or language responses from the participant, making it a promising method to study younger children and minimally verbal children with autism. In addition, we record eye movements as automatic responses to linguistic input instead of measuring eye movements that accompany conscious responses to linguistic input.

PROTOCOL:

This study has been approved by the Ethics Committee of the School of Medicine at Tsinghua University. Informed consent has been obtained from all individual participants included in the study.

1. Participant Screening and Study Preparation

1.1 Recruit Mandarin-speaking preschool children with autism.

89

90　Note: Their diagnoses should be confirmed by pediatric neurologists at hospitals using DSM-IV-
91　TR[40] or DSM-5[41] and, ideally, the number of participants should be no less than 15. The present
92　study recruited 25 participants with confirmed diagnoses.

93

94　1.2 Evaluate each participant independently using gold-standard diagnostic instruments like the
95　Autism Diagnostic Observation Schedule[42].

96

97　1.3 Measure the verbal IQ of participants using the Wechsler Preschool and Primary Scale of
98　Intelligence-IV (CN), a standardized IQ test designed for Mandarin-speaking children between the
99　ages of 2-6 and 6-11[43].

100

101　Note: The verbal IQ scores of the children with autism in the present study were all above 80.
102　They were all high-functioning children with autism.

103

104　1.4 Calculate each participant's mean length of utterance (MLU) by dividing the total number of
105　words by the number of utterances in each speech sample. Record 100 utterances for each
106　participant from either their interactions with parents or with teachers. Then, calculate the MLU
107　by dividing the total number of words in each participant's utterances by 100.

108

109　Note: MLU indicates the participant's sentence complexity levels.

110

111　1.5 Recruit TD children. Ideally match the TD children to the children with autism for age (TD
112　group 1), MLU (TD group 2), and verbal IQ (TD group 3).

113

114　Note: The present study recruited 50 TD children (25 boys and 25 girls) from local kindergartens.
115　25 matched to the children with autism for age and 25 matched to the children with autism for
116　both MLU and verbal IQ.

117

118　**2. Warm-up Session**

119

120　2.1 Invite the participants for a warm-up session before the actual test. Introduce the participant
121　to the research environment and interact with him or her to establish a good rapport.

122

123　Note: This can be done on the same day as the testing session or organized on a different day. In
124　the warm-up session, two experimenters are typically involved and interact with the participant
125　using toys and props.

126

127　**3. Conditions and Experimental Design**

128

129　3.1 Construct the test stimuli. Create 12 target items, each comprised of a visual stimulus, and
130　two spoken sentences containing the morphological markers BA and BEI, respectively. Construct
131　the spoken sentences using the same structure: morphological marker + noun phrase (NP) +
132　adverb + verb phrase (VP) (see Examples 1a and 1b below).

133
134 Note: The marker BA indicates that the following NP is the recipient of the holding event (see
135 2a), and BEI indicates that the following NP is the initiator of the event (see 2b). The subject NP
136 of a sentence in Mandarin can often be omitted when the referent of the NP is contextually
137 available.
138
139 **Example:**
140 (1) a. **BA** shizi qingqingdi bao-le   qilai.
141   **BA** lion   gently   hold   up
142   Meaning: Someone gently holds the lion.
143   b. **BEI** shizi qingqingdi bao-le   qilai.
144   **BEI** lion   gently   hold   up
145   Meaning: Someone is gently held by the lion.
146 (2) a. BA + [NP]$_{Recipient}$
147   b. BEI + [NP]$_{Initiator}$
148
149 3.1.1. Use Pixelmator (or another image editor) to create visual images. Open Pixelmator. Click
150 on the Pixelmator icon. Create a visual image from a template. Click **Show Details** in the template
151 chooser. Double-click the template to open it. Adjust the width, height, resolution, and color
152 depth from the pop-up menus. Enter the relevant parameters. Click **OK**.
153
154 3.1.2. Use Praat (or another audio editor) to construct spoken sentences. Set up the microphone.
155 Open Praat. Click on the Praat icon. Select **Record Mono sound** from the **New** menu. Set the
156 recording conditions by clicking the sample rate option of **22050**. Click the **Record** button.
157
158 3.1.3. Record the spoken sentences by asking a native Beijing Mandarin-speaker to produce the
159 sentences in a child-directed manner. Save the recordings by clicking **Save**.
160
161 Note: Typically, 12 to 16 target items are constructed for sentence comprehension studies with
162 children. Test stimuli can be created using other image and audio editors for a visual world study.
163
164 3.1.2 Construct visual images, each containing two pictures. The two pictures depict the same
165 event involving the same characters. Reverse the event roles (initiator or recipient) of the two
166 characters in the two pictures. Make one picture compatible with the construction containing BA
167 (BA-target event) and one with the construction containing BEI (BEI-target event). An example is
168 provided in **Figure 1**.
169
170 Note: This figure has been reprinted with permission from Zhou and Ma (2018)[19].
171
172 3.2 Counterbalance and randomization: divide the target trials into two experimental lists, with
173 a participant seeing each visual stimulus but listening to only one of the recorded sentences for
174 the stimulus. Counterbalance the spoken sentences containing BA and BEI across the two
175 experimental lists, with 6 constructions containing BA and 6 containing BEI. Add 12 filler items to
176 each experimental list and arrange the target and filler trials in a random order. Randomly assign

177 the participants to the two lists.
178
179 **4. Experimental Procedure**
180
181 4.1 Eye-tracking procedure.
182
183 4.1.1. Invite the participants to sit comfortably in front of the display monitor of the remote eye
184 tracker. Set the distance between the participants' eyes and the monitor around 60 cm. Perform
185 the standard calibration and validation procedures by asking the participants to fixate on a grid
186 of five fixation targets in random succession.
187
188 4.1.2. Present the participants with a spoken sentence while they are seeing a visual image, as
189 done in the standard visual world paradigm[10,44]. Use the monocular eye-tracking option by
190 tracking the eye that is on the same side as the illuminator of the eye tracker. Record the
191 participant's eye movements using the eye tracker.
192
193 Note: The eye tracker used in the present study allows remote eye-tracking with a sampling rate
194 of 500 Hz.
195
196 4.2 Testing and measuring.
197
198 4.2.1. Test the participants individually. Simply tell the participants to listen to the spoken
199 sentences while they are viewing the pictures. Ask one experimenter to monitor the participant
200 on the computer and one to stand behind the participant and gently rest her hands on the
201 participant's shoulders to minimize the participant's sudden movements.
202
203 4.2.2. Measure the participant's eye movements that arise as automatic responses to the
204 linguistic input using the eye tracker.
205
206 Note: The task does not ask participants to make any conscious judgments about the spoken
207 sentences to minimize their computational burden. The eye tracker automatically records the
208 eye movements.
209
210 4.3 Monitoring during the test: use the live viewer mode on the computer screen, exhibited by
211 the eye tracker during the test, to observe the participant's looking behavior. Ask the
212 experimenter who monitors data collection via the live viewer mode to signal to the
213 experimenter who stands behind the participant to reorient the participant if his or her eye gaze
214 wanders off the computer screen.
215
216 **5. Data Treatment and Analysis**
217
218 5.1 Code the participants' fixations in two interest areas. Use Data Viewer to draw the two
219 interest areas: BA-target event area and BEI-target event area (see **Figure 1**). Open Data Viewer.
220 Select one of the interest area shape icons on the tool bar. Use the mouse to drag a box around

221 the region you want to define as an interest area. Save the interest area in the Interest Area Set
222 folder. Apply the interest area to other visual images.
223
224 Note: The depicted event in the upper panel of **Figure 1** matches the BA-construction, hence the
225 BA-target event, and the event depicted in the lower panel matches **Figure 1b**, hence the BEI-
226 target event. The software used for data coding is Data Viewer, which comes with the eye tracker
227 used in the study. Other data analysis software is also available.
228
229 5.2 Analyze the eye gaze patterns using Data Viewer.
230
231 5.2.1. Open Data Viewer. Choose the sample report function from the menu to set the time
232 windows for analysis (*e.g.,* every 200 ms for the time window in the present study). Use the same
233 function to time lock the fixation proportions in the interest areas to the onset of the marker for
234 each trial. Export the raw data into an excel file using the export function from the menu.
235
236 5.2.2. Use the excel functions to average the fixation proportions following the onset of the
237 marker for each area. Use the excel functions to compute the fixation proportions in each time
238 window of 200 ms over a period of 5200 ms (the mean length of the target sentences + 200 ms)
239 from the onset of the marker for the two areas. Apply linear mixed-effects models to the eye
240 movement data, detailed in Representative Results below.
241
242 Note: The use of 200 ms as a time window is based on the standard procedure for analyzing child
243 eye gaze data in the literature[12,13,18,19,45-47], and it is generally assumed that it takes about 200 ms
244 to observe thhe effects of linguistic markers on eye movements[48].
245
246 **REPRESENTATIVE RESULTS:**
247 The present study uses minimal pairs as in Examples 1a and 1b to investigate if and how fast
248 children with autism can use event information encoded in two morphological markers during
249 real-time sentence comprehension. It was predicted that if they are able to rapidly and effectively
250 use event information in the two markers during real-time sentence comprehension, then they
251 should look more at the BA-target event when hearing BA than when hearing BEI. Also, they
252 should fixate more on the BEI-target event after listening to BEI than after listening to BA.
253
254 The comparison between 5-year-olds with autism and their age-matched TD peers is presented
255 in the representative results. **Figure 2** shows the average fixation proportions of the TD 5-year-
256 olds on the BA-target event (Panel A) and BEI-target event (Panel B) in the two conditions. **Figure**
257 **3** summarizes the average fixation proportions of the 5-year-olds with autism.
258
259 The figures show that the autism group displayed eye movement patterns similar to the age-
260 matched TD group. Both groups exhibited more fixations on the BA-target event when hearing
261 BA than when hearing BEI, occurring after onset of the object NP and before onset of the adverb.
262 To be specific, the effect occurred in the TD group during the window between 1400 and 1600
263 ms (**Figure 2**), whereas the effect occurred in the autism group during the window between 1800
264 and 2000 ms (**Figure 3**). By contrast, an opposite eye movement pattern was found in the BEI-

265 target event for both groups: more fixations on the BEI-target event were observed when hearing
266 BEI than when hearing BA, again occurring after onset of the object NP and prior to onset of the
267 adverb.

269 Fixation proportions were then transformed using the empirical logit formula[49]: *probability =*
270 *ln[(y+0.5)/(n-y+0.5)]*, where *y* is the number of fixations on the areas of interest during a
271 particular temporal bin and *n* is the total number of fixations in that temporal bin. Linear mixed-
272 effects models were then fitted to the transformed data. Statistical models were computed for
273 the two groups separately based on their fixations in the two interest areas in the critical time
274 windows, where time and marker type (BA versus BEI) were treated as fixed effects. Random
275 intercepts and slopes were included for both participants and items[50]. The fitting process was
276 conducted via functions lmer from package lme4 (v1.1-12)[51] of the R (v3.2.5) software
277 environment[52]. A Wald test was then used to compute p-values for each fixed effect.

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

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

297 Overall, the eye patterns exhibited by the 5-year-olds with autism provide evidence that they
298 were able to use the event information encoded in two morphological markers rapidly and
299 effectively during real-time sentence comprehension. The results show that the recorded eye
300 movements as automatic responses to the linguistic input are sensitive measures of sentence
301 comprehension abilities in both TD children and children with autism.

303 **FIGURE AND TABLE LEGENDS:**
304 **Figure 1: Example visual image**. (A) Indicates a BA-target event. (B) Represents a BEI-target
305 event. This figure has been reprinted with permission from Zhou and Ma (2018)[19].

307 **Figure 2: Average fixation proportions from marker onset in both conditions in TD 5-year-olds.**

308   (A) Shows the fixation proportions on the BA-target event. (B) Illustrates the fixation proportions
309   on the BEI-target event.
310
311   **Figure 3. Average fixation proportions from marker onset in both conditions in 5-year-olds with**
312   **autism.** (A) Shows the fixation proportions on the BA-target event. (B) Illustrates the fixation
313   proportions on the BEI-target event.
314
315   **DISCUSSION:**
316   In the current study, we present an eye-tracking paradigm that can directly and effectively assess
317   the sentence comprehension abilities of children with autism. We found that 5-year-old children
318   with autism, like their age-matched TD peers, exhibited eye gaze patterns that reflect effective
319   and rapid use of linguistic cues during real-time sentence comprehension.
320
321   The findings provide evidence that eye-tracking (in particular, the visual world paradigm) is a
322   sensitive measure of real-time sentence comprehension in children with autism. Compared to
323   offline methods, the paradigm has several advantages. First, it is sensitive to the time course of
324   sentence comprehension. Second, it minimizes the task and communication demands involved,
325   thus is a better-suited method that can be used with children exhibiting challenging behavioral
326   features. Third, it simply records eye movements as automatic responses to linguistic input
327   without asking participants to provide conscious judgments about the input, significantly
328   reducing the computational burden of the subjects.
329
330   The visual world paradigm is based on a linking assumption that eye movements in the visual
331   world are synchronized to the real-time processing of concurrent linguistic stimuli. Thus, an
332   effective language comprehension study using the visual world paradigm requires a close
333   mapping between eye gaze patterns in the visual world and referential processing of the spoken
334   language. To ensure a close mapping between the two, it is important to first design the visual
335   stimuli in a way so that eye movements in the visual images reflect only the processes underlying
336   comprehension of the spoken language, and that other factors that may affect participants' eye
337   movements are well-controlled. Second, it is important to time-lock participants' eye movements
338   to the onset of a critical linguistic marker in the spoken language, and to make sure that each
339   element and boundaries of the spoken language elements can be clearly identified for later
340   analyses.
341
342   The visual world paradigm has been used successfully to test TD children's language abilities. The
343   present study explored the potential of conducting visual world studies on language
344   comprehension in preschool children with autism. As discussed, these findings provide evidence
345   for the validity and sensitivity of the paradigm in testing linguistic knowledge in children with
346   autism. The findings also invite us to rethink questions surrounding the language comprehension
347   abilities of children with autism. As discussed, previous research seems to suggest that the
348   sentence comprehension abilities of children with autism might be severely impaired; however,
349   as noted by Kasari et al.[27] and Plesa-Skwerer et al.[29], it is often difficult to evaluate the
350   comprehension abilities of children with autism using traditional methods like standardized tests
351   or other off-line tasks, because these tasks require high response demands or interactions with

352 the experimenters; as a result, this might pose particular difficulties for children with autism.
353 Using the visual world paradigm, the present study shows for the first time that when minimal
354 task and communication demands are involved, young children with autism are able to use
355 linguistic cues effectively and rapidly during real-time sentence comprehension. Their sentence
356 comprehension abilities are far better than has been suggested by previous research. The
357 findings also provide evidence that poor comprehension performance of children with autism in
358 past research is perhaps due to a lack of social responsiveness and the high task and
359 communication demands involved in these traditional tasks.
360
361 The visual world paradigm can be systematically applied to establish eye gaze patterns associated
362 with language processing in autism, which will help us better understand the nature of sentence
363 processing mechanisms in autism as well as help to identify early clinical markers for autism.
364

372 **DISCLOSURES:**
373 The authors have nothing to disclose.
374

375 **REFERENCES:**
376 1. Altmann, G. T., Kamide, Y. Incremental interpretation at verbs: Restricting the domain of
377 subsequent reference. *Cognition*. **73** (3), 247-264 (1999).
378 2. Altmann, G. T., Kamide, Y. The real-time mediation of visual attention by language and world
379 knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of
380 Memory and Language*. **57** (4), 502-518 (2007).
381 3. DeLong, K. A., Urbach, T. P., Kutas, M. Probabilistic word pre-activation during language
382 comprehension inferred from electrical brain activity. *Nature neuroscience*. **8** (8), 1117-1121
383 (2005).
384 4. Kamide, Y., Altmann, G. T., Haywood, S. L. The time-course of prediction in incremental
385 sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and
386 Language*. **49** (1), 133-156 (2003).
387 5. Knoeferle, P., Crocker, M. W., Scheepers, C., Pickering, M. J. The influence of the immediate
388 visual context on incremental thematic role-assignment: Evidence from eye-movements in
389 depicted events. *Cognition*. **95** (1), 95-127 (2005).
390 6. Knoeferle, P., Kreysa, H. Can speaker gaze modulate syntactic structuring and thematic role
391 assignment during spoken sentence comprehension? *Frontiers in Psychology*. **3**, 538 (2012).
392 7. Knoeferle, P., Urbach, T. P., Kutas, M. Comprehending how visual context influences
393 incremental sentence processing: Insights from ERPs and picture-sentence verification.
394 *Psychophysiology*. **48** (4), 495-506 (2011).
395 8. Pickering, M. J., Traxler, M. J., Crocker, M. W. Ambiguity resolution in sentence processing:

396 Evidence against frequency-based accounts. *Journal of Memory and Language*. **43** (3), 447-475
397 (2000).
398 9. Staub, A., Clifton, C. Syntactic prediction in language comprehension: Evidence from *either…*
399 *or*. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. **32** (2), 425-436 (2006).
400 10. Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K., Sedivy, J. Integration of visual and
401 linguistic information in spoken language comprehension. *Science.* **268** (5217), 1632-1634 (1995).
402 11. Van Berkum, J. J., Brown, C. M., Zwitserlood, P., Kooijman, V., Hagoort, P. Anticipating
403 upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental*
404 *Psychology: Learning, Memory, and Cognition*. **31** (3), 443-467 (2005).
405 12. Choi, Y., Trueswell, J. C. Children's (in) ability to recover from garden paths in a verb-final
406 language: Evidence for developing control in sentence processing. *Journal of Experimental Child*
407 *Psychology*. **106** (1), 41-61 (2010).
408 13. Huang, Y., Zheng, X., Meng, X., Snedeker, J. Children's assignment of grammatical roles in the
409 online processing of Mandarin passive sentences. *Journal of Memory and Language*. **69** (4), 589-
410 606 (2013).
411 14. Lew-Williams, C., Fernald, A. Young children learning Spanish make rapid use of grammatical
412 gender in spoken word recognition. *Psychological Science*. **18** (3), 193-198 (2007).
413 15. Sekerina, I. A., Trueswell, J. C. Interactive processing of contrastive expressions by Russian
414 children. *First Language*. **32** (1-2), 63-87 (2012).
415 16. Trueswell, J. C., Sekerina, I., Hill, N. M., Logrip, M. L. The kindergarten-path effect: Studying
416 on-line sentence processing in young children. *Cognition*. **73** (2), 89-134 (1999).
417 17. Van Heugten, M., Shi, R. French-learning toddlers use gender information on determiners
418 during word recognition. *Developmental Science*. **12** (3), 419-425 (2009).
419 18. Zhou, P., Crain, S., Zhan, L. Grammatical aspect and event recognition in children's online
420 sentence comprehension. *Cognition*. **133** (1), 262-276 (2014).
421 19. Zhou, P., Ma, W. Children's use of morphological cues in real-time event representation.
422 *Journal of Psycholinguistic Research.* **47** (1), 241–260 (2018).
423 20. Eigsti, I. M., Bennetto, L., Dadlani, M. B. Beyond pragmatics: Morphosyntactic development
424 in autism. *Journal of Autism and Developmental Disorders*. **37** (6), 1007–1023 (2007).
425 21. Kjelgaard, M. M., Tager-Flusberg, H. An investigation of language impairment in autism:
426 Implications for genetic subgroups. *Language and Cognitive Processes*. **16** (2-3), 287-308 (2001).
427 22. Tager-Flusberg, H. Risk factors associated with language in autism spectrum disorder: clues
428 to underlying mechanisms. *Journal of Speech, Language, and Hearing Research*. **59** (1), 143-154
429 (2016).
430 23. Tager-Flusberg, H., Kasari, C. Minimally verbal school-aged children with autism spectrum
431 disorder: the neglected end of the spectrum. *Autism Research*. **6** (6), 468-478 (2013).
432 24. Tek, S., Mesite, L., Fein, D., Naigles, L. Longitudinal analyses of expressive language
433 development reveal two distinct language profiles among young children with autism spectrum
434 disorders. *Journal of Autism and Developmental Disorders*. **44** (1), 75-89 (2014).
435 25. Wittke, K., Mastergeorge, A. M., Ozonoff, S., Rogers, S. J., Naigles, L. R. Grammatical language
436 impairment in autism spectrum disorder: Exploring language phenotypes beyond standardized
437 testing. *Frontiers in Psychology*. **8**, 532 (2017).
438 26. Hudry, K., Leadbitter, K., Temple, K., Slonims, V., McConachie, H., Aldred, C., et al.
439 Preschoolers with autism show greater impairment in receptive compared with expressive

language abilities. *International Journal of Language & Communication Disorders*. **45** (6), 681-690 (2010).

27. Kasari, C., Brady, N., Lord, C., Tager-Flusberg, H. Assessing the minimally verbal school-aged child with autism spectrum disorder. *Autism Research*. **6** (6), 479-493 (2013).

28. Luyster, R. J., Kadlec, M. B., Carter, A., Tager-Flusberg, H. Language assessment and development in toddlers with autism spectrum disorders. *Journal of Autism and Developmental Disorders*. **38** (8), 1426-1438 (2008).

29. Plesa-Skwerer, D., Jordan, S.E., Brukilacchio, B.H. Tager-Flusberg, H. Comparing methods for assessing receptive language skills in minimally verbal children and adolescents with Autism Spectrum Disorders. *Autism*. **20** (5), 591-604 (2016).

30. Boucher, J. Research review: Structural language in autism spectrum disorder–characteristics and causes. *Journal of Child Psychology and Psychiatry*. **53** (3), 219–233 (2012).

31. Eigsti, I. M., de Marchena, A. B., Schuh, J. M., Kelley, E. Language acquisition in autism spectrum disorders: A developmental review. *Research in Autism Spectrum Disorders.* **5** (2), 681–691 (2011).

32. Howlin, P. Outcome in high-functioning adults with autism with and without early language delays: Implications for the differentiation between autism and Asperger syndrome. *Journal of Autism and Developmental Disorders.* **33** (1), 3–13 (2003).

33. Koning, C., Magill-Evans, J. Social and language skills in adolescent boys with Asperger syndrome. *Autism: The International Journal of Research and Practice*. **5** (1), 23–36 (2001).

34. Kover, S. T., Haebig, E., Oakes, A., McDuffie, A., Hagerman, R. J., Abbeduto, L. Sentence comprehension in boys with autism spectrum disorder. *American Journal of Speech-Language Pathology*. **23** (3), 385–394 (2004).

35. Perovic, A., Modyanova, N., Wexler, K. Comprehension of reflexive and personal pronouns in children with autism: A syntactic or pragmatic deficit? *Applied Psycholinguistics.* **34** (4), 813–835 (2013).

36. Rapin, I., Dunn, M. Update on the language disorders of individuals on the autistic spectrum. *Brain Development*. **25** (3), 166–172 (2003).

37. Tager-Flusberg, H. Sentence comprehension in autistic children. *Applied Psycholinguistics*. **2** (1), 5–24 (1981).

38. Rutter, M., Maywood, L., Howlin, P. Language delay and social development. In: *Specific speech and language disorders in children: Correlates, characteristics, and outcomes.* Fletcher, P., & Hall, D., eds., London: Whurr. (1992).

39. Tager-Flusberg, H. The challenge of studying language development in autism. In: *Methods for studying language production*. Menn L., & Ratner N. B., eds., Mahwah, NJ: Erlbaum. (2000).

40. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders, 4th edition, text revision (DSM-IV-TR).* American Psychiatric Association. Washington, DC. (2000).

41. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders, 5th edition (DSM-5)*. American Psychiatric Association. Washington, DC. (2013).

42. Lord, C., Rutter, M., DiLavore, P. C., Risi, S. *Autism diagnostic observation schedule.* Western Psychological Services. Los Angeles, CA. (1999).

43. Li, Y., Zhu, J. *Wechsler Preschool and Primary Scale of Intelligence™-IV (CN) [WPPSI-IV (CN)].* Zhuhai King-may Psychological Measurement Technology Development Co., Ltd. Zhuhai. (2014).

44. Cooper, R. M. The control of eye fixation by the meaning of spoken language: A new

484 methodology for the real-time investigation of speech perception, memory, and language
485 processing. *Cognitive Psychology.* **6** (1), 84 – 107 (1974).
486 45. Huang, Y., Snedeker, J. Semantic meaning and pragmatic interpretation in five-year olds:
487 Evidence from real time spoken language comprehension. *Developmental Psychology*. **45** (6),
488 1723-1739 (2009).
489 46. Snedeker, J., Yuan, S. Effects of prosodic and lexical constraints on parsing in young children
490 (and adults). *Journal of Memory and Language*. **58** (2), 574-608 (2008).
491 47. Zhou, P., Crain, S., Zhan, L. Sometimes children are as good as adults: The pragmatic use of
492 prosody in children's on-line sentence processing. *Journal of Memory and Language*. **67** (1), 149-
493 164 (2012).
494 48. Matin, E., Shao, K. C., Boff, K. R. Saccadic overhead: Information-processing time with and
495 without saccades. *Perception & Psychophysics*. **53** (4), 372-380 (1993).
496 49. Barr, D. J. Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal
497 of Memory and Language*. **59** (4), 457-474 (2008).
498 50. Baayen, R. H., Davidson, D. J., Bates, D. M. Mixed-effects modeling with crossed random
499 effects for subjects and items. *Journal of Memory and Language.* **59** (4), 390-412 (2008).
500 51. Bates, D. M., Maechler, M., Bolker, B. *lme4: Linear mixed-effects models using S4 classes*.
501 Retrieved from http://cran.r-project.org/web/packages/lme4/index.html. (2013).
502 52. R Development Core Team. *R: A Language and Environment for Statistical Computing*. R
503 Foundation for Statistical Computing. Vienna. Retrieved from http://www.r-project.org. (2017).
504
505
506
507

Figure 1. Example visual image used in the study.
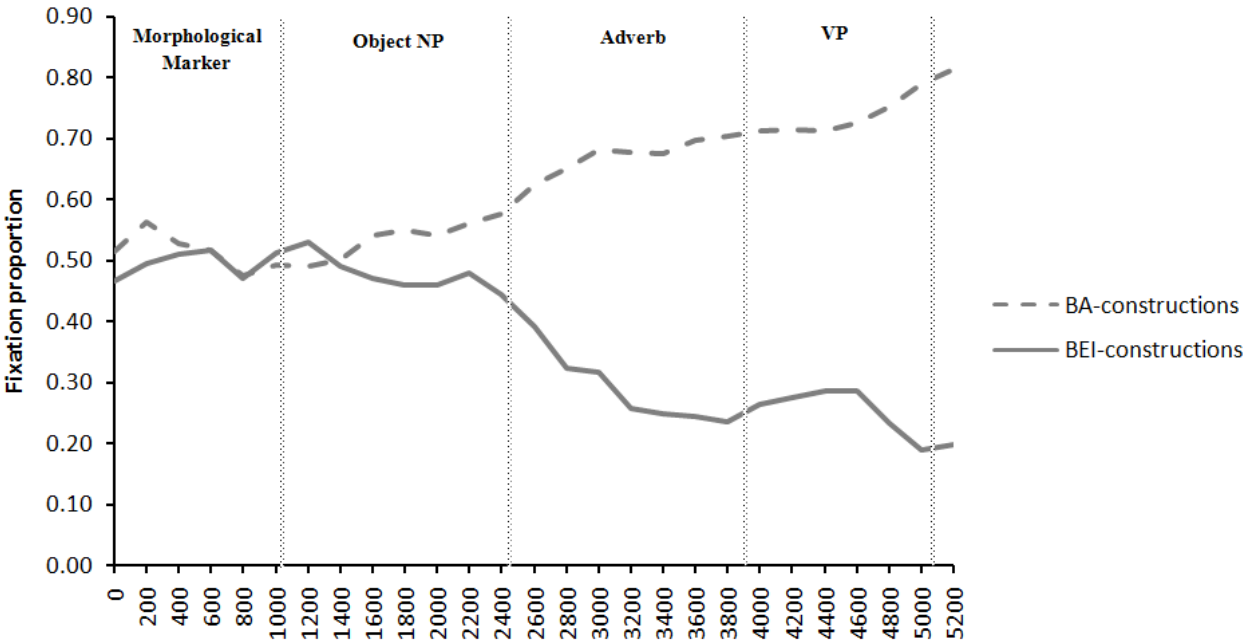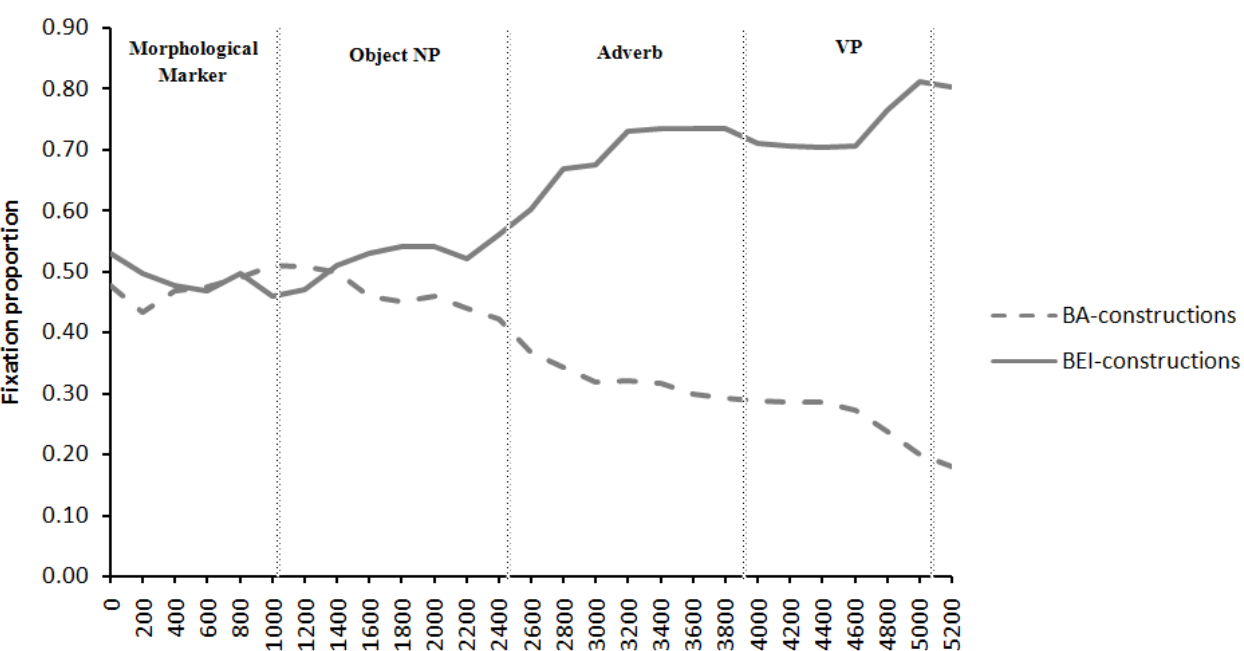
BA-Target Event

BEI-Target Event

Figure 2. Average fixation proportions following the onset of the morphological marker on the two critical categories across the two
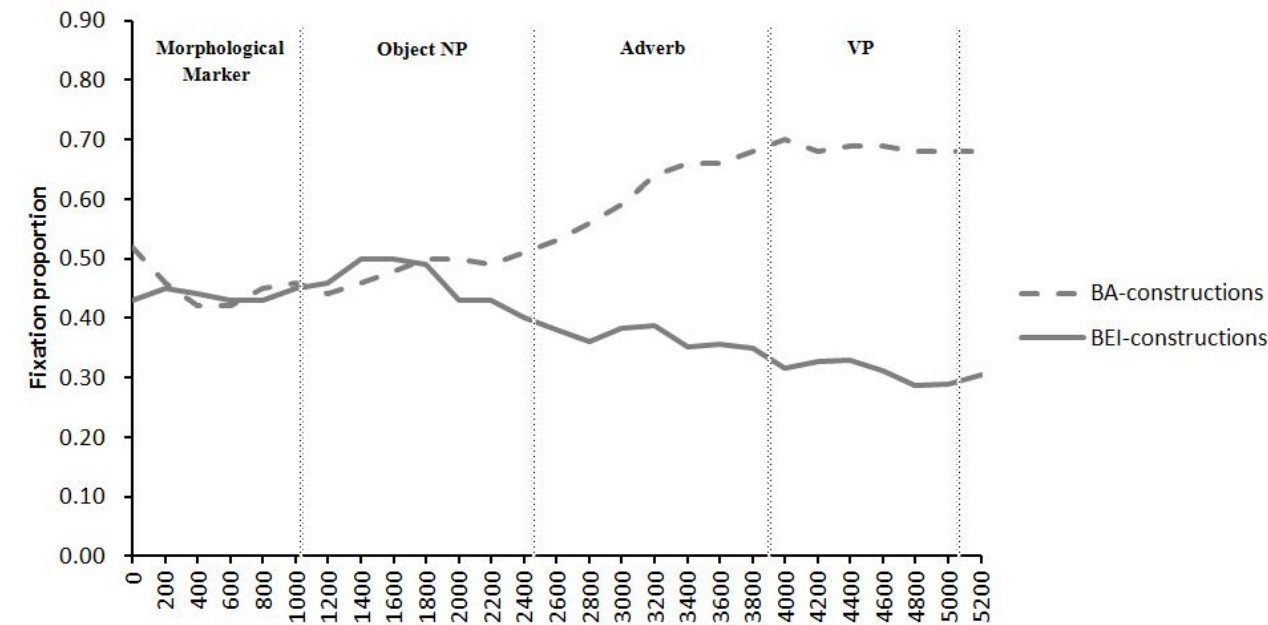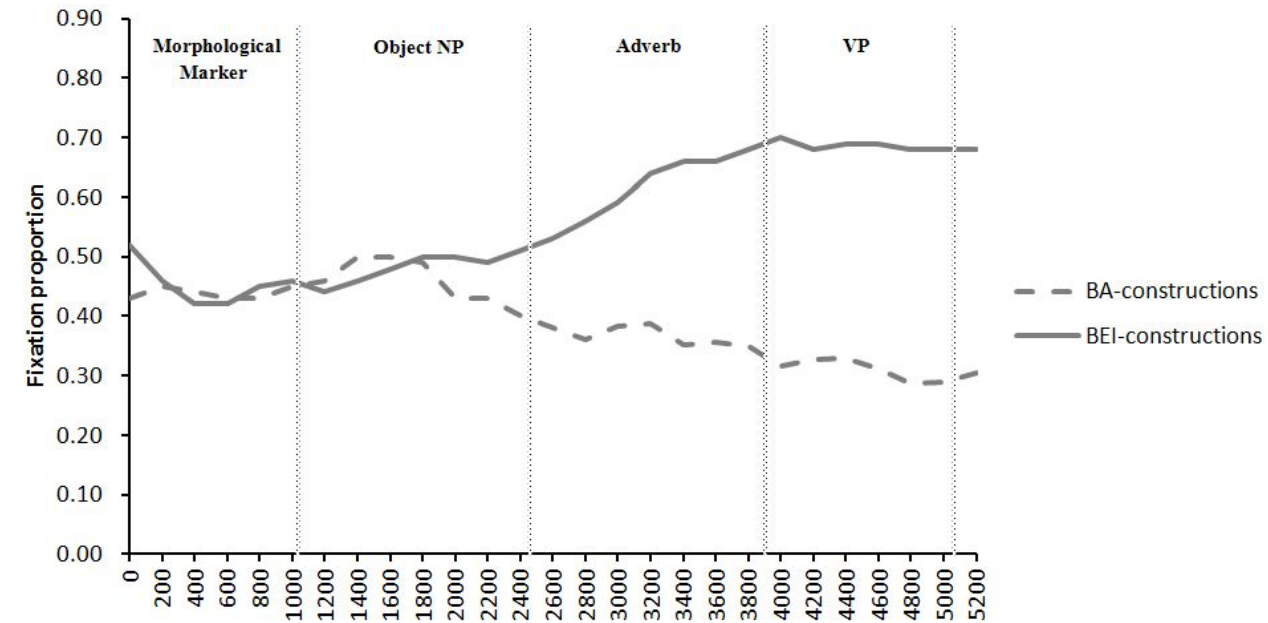
ms from onset of the morphological marker

Figure 3. Average fixation proportions following the onset of the
morphological marker on the two critical categories across the two

A

B

ms from onset of the morphological marker

Materials Table

Name of Material/ Equipment	Company	Catalog Number	Comments/Description
EyeLink 1000 plus eye tracker	SR Research Ltd.		The EyeLink 1000 plus allows remote eye tracking, without a head s

upport. The eye tracker provides information about the participant's point of gaze at a sampling rate of 500 Hz, and it has accuracy o

f 0.5 degrees of visual angle.

jove
VISUALIZED EXPERIMENTS

1 Alewife Center #200
Cambridge, MA 02140
tel. 617.945.9051
www.jove.com

ARTICLE AND VIDEO LICENSE AGREEMENT

Title of Article: *Using the Visual world Paradigm to study Real-Time sentence Comprehension in Mandarin-speaking children with Autism*

Author(s): *Peng Zhou, Weiyi Ma, Likan Zhan, Huimin Ma*

Item 1: The Author elects to have the Materials be made available (as described at http://www.jove.com/publish) via:

☒ Standard Access   ☐ Open Access

Item 2: Please select one of the following items:

☒ The Author is **NOT** a United States government employee.

☐ The Author is a United States government employee and the Materials were prepared in the course of his or her duties as a United States government employee.

☐ The Author is a United States government employee but the Materials were NOT prepared in the course of his or her duties as a United States government employee.

ARTICLE AND VIDEO LICENSE AGREEMENT

1.   **Defined Terms.** As used in this Article and Video License Agreement, the following terms shall have the following meanings: "**Agreement**" means this Article and Video License Agreement; "**Article**" means the article specified on the last page of this Agreement, including any associated materials such as texts, figures, tables, artwork, abstracts, or summaries contained therein; "**Author**" means the author who is a signatory to this Agreement; "**Collective Work**" means a work, such as a periodical issue, anthology or encyclopedia, in which the Materials in their entirety in unmodified form, along with a number of other contributions, constituting separate and independent works in themselves, are assembled into a collective whole; "**CRC License**" means the Creative Commons Attribution-Non Commercial-No Derivs 3.0 Unported Agreement, the terms and conditions of which can be found at: http://creativecommons.org/licenses/by-nc-nd/3.0/legalcode; "**Derivative Work**" means a work based upon the Materials or upon the Materials and other pre-existing works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which the Materials may be recast, transformed, or adapted; "**Institution**" means the institution, listed on the last page of this Agreement, by which the Author was employed at the time of the creation of the Materials; "**JoVE**" means MyJove Corporation, a Massachusetts corporation and the publisher of The Journal of Visualized Experiments; "**Materials**" means the Article and / or the Video; "**Parties**" means the Author and JoVE; "**Video**" means any video(s) made by the Author, alone or in conjunction with any other parties, or by JoVE or its affiliates or agents, individually or in collaboration with the Author or any other parties, incorporating all or any portion of the Article, and in which the Author may or may not appear.

2.   **Background.** The Author, who is the author of the Article, in order to ensure the dissemination and protection of the Article, desires to have the JoVE publish the Article and create and transmit videos based on the Article. In furtherance of such goals, the Parties desire to memorialize in this Agreement the respective rights of each Party in and to the Article and the Video.

3.   **Grant of Rights in Article.** In consideration of JoVE agreeing to publish the Article, the Author hereby grants to JoVE, subject to **Sections 4** and **7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Article in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Article into other languages, create adaptations, summaries or extracts of the Article or other Derivative Works (including, without limitation, the Video) or Collective Works based on all or any portion of the Article and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and(c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. If the "Open Access" box has been checked in **Item 1** above, JoVE and the Author hereby grant to the public all such rights in the Article as provided in, but subject to all limitations and requirements set forth in, the CRC License.

jove

1 Alewife Center #200
Cambridge, MA 02140
tel 617.945.9051
www.jove.com

JOURNAL OF VISUALIZED EXPERIMENTS

ARTICLE AND VIDEO LICENSE AGREEMENT

4.   Retention of Rights in Article. Notwithstanding the exclusive license granted to JoVE in **Section 3** above, the Author shall, with respect to the Article, retain the non-exclusive right to use all or part of the Article for the non-commercial purpose of giving lectures, presentations or teaching classes, and to post a copy of the Article on the Institution's website or the Author's personal website, in each case provided that a link to the Article on the JoVE website is provided and notice of JoVE's copyright in the Article is included. All non-copyright intellectual property rights in and to the Article, such as patent rights, shall remain with the Author.

5.   Grant of Rights in Video – Standard Access. This **Section 5** applies if the "Standard Access" box has been checked in **Item 1** above or if no box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby acknowledges and agrees that, Subject to **Section 7** below, JoVE is and shall be the sole and exclusive owner of all rights of any nature, including, without limitation, all copyrights, in and to the Video. To the extent that, by law, the Author is deemed, now or at any time in the future, to have any rights of any nature in or to the Video, the Author hereby disclaims all such rights and transfers all such rights to JoVE.

6.   Grant of Rights in Video – Open Access. This **Section 6** applies only if the "Open Access" box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby grants to JoVE, subject to **Section 7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Video in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Video into other languages, create adaptations, summaries or extracts of the Video or other Derivative Works or Collective Works based on all or any portion of the Video and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. For any Video to which this **Section 6** is applicable, JoVE and the Author hereby grant to the public all such rights in the Video as provided in, but subject to all limitations and requirements set forth in, the CRC License.

7.   Government Employees. If the Author is a United States government employee and the Article was prepared in the course of his or her duties as a United States government employee, as indicated in **Item 2** above, and any of the licenses or grants granted by the Author hereunder exceed the scope of the 17 U.S.C. 403, then the rights granted hereunder shall be limited to the maximum rights permitted under such statute. In such case, all provisions contained herein that are not in conflict with such statute shall remain in full force and effect, and all provisions contained herein that do so conflict shall be deemed to be amended so as to provide to JoVE the maximum rights permissible within such statute.

8.   Protection of the Work. The Author(s) authorize JoVE to take steps in the Author(s) name and on their behalf if JoVE believes some third party could be infringing or might infringe the copyright of either the Author's Article and/or Video.

9.   Likeness, Privacy, Personality. The Author hereby grants JoVE the right to use the Author's name, voice, likeness, picture, photograph, image, biography and performance in any way, commercial or otherwise, in connection with the Materials and the sale, promotion and distribution thereof. The Author hereby waives any and all rights he or she may have, relating to his or her appearance in the Video or otherwise relating to the Materials, under all applicable privacy, likeness, personality or similar laws.

10.   Author Warranties. The Author represents and warrants that the Article is original, that it has not been published, that the copyright interest is owned by the Author (or, if more than one author is listed at the beginning of this Agreement, by such authors collectively) and has not been assigned, licensed, or otherwise transferred to any other party. The Author represents and warrants that the author(s) listed at the top of this Agreement are the only authors of the Materials. If more than one author is listed at the top of this Agreement and if any such author has not entered into a separate Article and Video License Agreement with JoVE relating to the Materials, the Author represents and warrants that the Author has been authorized by each of the other such authors to execute this Agreement on his or her behalf and to bind him or her with respect to the terms of this Agreement as if each of them had been a party hereto as an Author. The Author warrants that the use, reproduction, distribution, public or private performance or display, and/or modification of all or any portion of the Materials does not and will not violate, infringe and/or misappropriate the patent, trademark, intellectual property or other rights of any third party. The Author represents and warrants that it has and will continue to comply with all government, institutional and other regulations, including, without limitation all institutional, laboratory, hospital, ethical, human and animal treatment, privacy, and all other rules, regulations, laws, procedures or guidelines, applicable to the Materials, and that all research involving human and animal subjects has been approved by the Author's relevant institutional review board.

11.   JoVE Discretion. If the Author requests the assistance of JoVE in producing the Video in the Author's facility, the Author shall ensure that the presence of JoVE employees, agents or independent contractors is in accordance with the relevant regulations of the Author's institution. If more than one author is listed at the beginning of this Agreement, JoVE may, in its sole

jove

1 Alewife Center #200
Cambridge, MA 02140
tel. 617.945.9051
www.jove.com

ARTICLE AND VIDEO LICENSE AGREEMENT

discretion, elect not take any action with respect to the Article until such time as it has received complete, executed Article and Video License Agreements from each such author. JoVE reserves the right, in its absolute and sole discretion and without giving any reason therefore, to accept or decline any work submitted to JoVE. JoVE and its employees, agents and independent contractors shall have full, unfettered access to the facilities of the Author or of the Author's institution as necessary to make the Video, whether actually published or not. JoVE has sole discretion as to the method of making and publishing the Materials, including, without limitation, to all decisions regarding editing, lighting, filming, timing of publication, if any, length, quality, content and the like.

12. **Indemnification.** The Author agrees to indemnify JoVE and/or its successors and assigns from and against any and all claims, costs, and expenses, including attorney's fees, arising out of any breach of any warranty or other representations contained herein. The Author further agrees to indemnify and hold harmless JoVE from and against any and all claims, costs, and expenses, including attorney's fees, resulting from the breach by the Author of any representation or warranty contained herein or from allegations or instances of violation of intellectual property rights, damage to the Author's or the Author's institution's facilities, fraud, libel, defamation, research, equipment, experiments, property damage, personal injury, violations of institutional, laboratory, hospital, ethical, human and animal treatment, privacy or other rules, regulations, laws, procedures or guidelines, liabilities and other losses or damages related in any way to the submission of work to JoVE, making of videos by JoVE, or publication in JoVE or elsewhere by JoVE. The Author shall be responsible for, and shall hold JoVE harmless from, damages caused by lack of sterilization, lack of cleanliness or by contamination due to the making of a video by JoVE its employees, agents or independent contractors. All sterilization, cleanliness or decontamination procedures shall be solely the responsibility of the Author and shall be undertaken at the Author's expense. All indemnifications provided herein shall include JoVE's attorney's fees and costs related to said losses or damages. Such indemnification and holding harmless shall include such losses or damages incurred by, or in connection with, acts or omissions of JoVE, its employees, agents or independent contractors.

13. **Fees.** To cover the cost incurred for publication, JoVE must receive payment before production and publication of the Materials. Payment is due in 21 days of invoice. Should the Materials not be published due to an editorial or production decision, these funds will be returned to the Author. Withdrawal by the Author of any submitted Materials after final peer review approval will result in a US$1,200 fee to cover pre-production expenses incurred by JoVE. If payment is not received by the completion of filming, production and publication of the Materials will be suspended until payment is received.

14. **Transfer, Governing Law.** This Agreement may be assigned by JoVE and shall inure to the benefits of any of JoVE's successors and assignees. This Agreement shall be governed and construed by the internal laws of the Commonwealth of Massachusetts without giving effect to any conflict of law provision thereunder. This Agreement may be executed in counterparts, each of which shall be deemed an original, but all of which together shall be deemed to me one and the same agreement. A signed copy of this Agreement delivered by facsimile, e-mail or other means of electronic transmission shall be deemed to have the same legal effect as delivery of an original signed copy of this Agreement.

A signed copy of this document must be sent with all new submissions. Only one Agreement is required per submission.

CORRESPONDING AUTHOR

Name: PENG ZHOU

Department: Department of Foreign Languages and Literatures

Institution: Tsinghua University

Title: Using the Visual World Paradigm to study Real-Time Sentence Comprehension in Mandarin-speaking children with Autism

Signature: [signature]   Date: June 30, 2018

Please submit a **signed** and **dated** copy of this license by one of the following three methods:
1. Upload an electronic version on the JoVE submission site
2. Fax the document to +1.866.381.2236
3. Mail the document to JoVE / Attn: JoVE Editorial / 1 Alewife Center #200 / Cambridge, MA 02140

612542.6   For questions, please contact us at submissions@jove.com or +1.617.945.9051.

Department of Foreign Languages and Literatures
Child Cognition Lab, Tsinghua University
Beijing 100084, China
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;11[43]."

(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 paradigm[10, 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 formula[49]: *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 items[50]. The fitting process was conducted via functions lmer from package lme4 (v1.1-12)[51] of the R (v3.2.5) software environment[52]. 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.

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

Best,

Peng Zhou