Journal of Autism and Developmental Disorders (2021) 51:1562–1583
https://doi.org/10.1007/s10803-020-04645-7

ORIGINAL PAPER

Structural and Contextual Cues in Third-Person Pronoun Interpretation by Children with Autism Spectrum Disorder and Their Neurotypical Peers

Marisa Nagano[1] · Emily Zane[2] · Ruth B. Grossman[3]

Published online: 12 August 2020

Abstract
This study investigates the use of structural and discourse contextual cues in the interpretation of third-person pronouns by children and adolescents with autism and their neurotypical peers. Results show that referent-biasing contextual information influences pronominal interpretation and modulates looking patterns in both groups compared to a context-neutral condition. These results go against the predictions of Weak Central Coherence and the notion that pragmatics in general is impaired in ASD, since the ASD group was able to use details in discourse context to influence the pronominal interpretation process. However, although discourse context influenced looking patterns in both groups, the groups nevertheless diverged in the nature of these patterns, suggesting that behavioral differences may emerge in more complicated discourse tasks.

Keywords Autism · Pragmatics · Pronoun · Reference · Weak central coherence · Eye-tracking

The theory of weak central coherence (WCC) proposes that individuals with ASD tend to focus on local detail while ignoring global context (Frith and Happé 1994; Happé and Frith 2006). For language processing specifically, WCC predicts that individuals with autism will have difficulty using contextual information to influence linguistic interpretation during discourse. Such a prediction is in line with substantial evidence that individuals with ASD particularly struggle with pragmatics (DSM-V, American Psychological Association 2013). Previous studies on the possible effects of WCC on language behavior in ASD groups have investigated a limited range of linguistic phenomena, mostly focusing on the effects of discourse context on lexical access (e.g., Happé 1997; Brock et al. 2008) or the use of discourse information in inferencing (e.g., Saldaña and Frith 2007). These tasks focus largely on semantics/pragmatics, such as using content words read/heard earlier in a sentence to predict upcoming content words, or using semantic details in a story to draw conclusions about the larger meaning of a passage.

The current study expands upon this previous work by investigating WCC via an aspect of language that involves both structural and semantic/pragmatic cues: specifically, the interpretation and processing of third-person pronouns. As far as we are aware, no previous study has investigated WCC by examining third-person pronoun interpretation and processing in ASD. Third-person pronominal interpretation provides an ideal test for WCC, since determining pronominal reference involves both local cues (a structural preference for a local subject antecedent) and global cues (the integration of semantic information from the larger discourse context). This study investigates the relative contribution of each of these cues in both ASD and neurotypical (NT) groups by examining the interpretation and on-line processing of third-person pronouns for which local and global cues conflict with one another—i.e., structural information suggests one referent while discourse contextual information suggests

✉ Marisa Nagano
marisa.nagano@liu.edu

Emily Zane
zaneer@jmu.edu

Ruth B. Grossman
ruth_grossman@emerson.edu

[1] Department of Communication Sciences & Disorders, Long Island University—Brooklyn, 1 University Plaza, Brooklyn, NY 11201, USA

[2] Department of Communication Sciences and Disorders, James Madison University, 800 S Main St, Harrisonburg, VA 22807, USA

[3] Department of Communication Sciences & Disorders, Emerson College, 120 Boylston Street, Boston, MA 02116, USA

another. WCC predicts that the ASD group (more so than the NT group) should overlook discourse context information and instead favor the structural preference for a local subject, even when context biases a non-subject referent. On the other hand, if the ASD group *does* use the discourse contextual cues, this would provide evidence against a WCC account of language processing in ASD.

Weak Central Coherence and Language in ASD

One particularly prolific strain of research testing potential WCC effects on language behavior in ASD groups focuses on the use of intrasentential context to disambiguate a homograph in reading. This paradigm originated in Frith and Snowling (1983), with an influential adaptation by Happé (1997). In the latter study, children and adults with ASD and NT children read aloud sentences containing homographs (words with the same spelling but two pronunciations) that varied on two conditions: pronunciation of the homograph (frequent, infrequent), and position of homograph (before or after disambiguating context), as in the example below:

(1) Frequent Pronunciation, Before Context:
There was a big tear on her cheek.
(2) Frequent Pronunciation, After Context:
Molly was very happy but on Lily's cheek there was a big tear.
(3) Infrequent Pronunciation, Before Context:
There was a big tear in her dress.
(4) Infrequent Pronunciation, After Context:
They climbed over the hedge. Mary's dress was spotless but in Lucy's dress there was a big tear.

Results of Happé (1997) found that NT participants showed greater improvement on the infrequent pronunciations after context had been provided than the ASD group did, suggesting that NT children made better use of context than the ASD group. After Happé (1997)'s study, numerous replications have shown that ASD groups tend to be more accurate on frequent pronunciations of homographs regardless of context (Jolliffe and Baron-Cohen 1999; Lopez and Leekam 2003). Taken together, these studies support the WCC theory, since the authors claim that the ASD groups cannot use the contextual information given in sentences like (4) in order to arrive at the less-frequent pronunciation. There is some evidence, however, that performance on this task is sensitive to how familiar the ASD participants are with the less-frequent pronunciations; Snowling and Frith (1986) found no difference in ASD vs. NT groups when they checked to make sure all participants knew the less-frequent pronunciations before the experiment. Thus, group differences reported in other homograph studies could have been due to unfamiliarity with less common words in the ASD group.

Results from modified versions of these homograph studies have further cast doubt on the argument that WCC consistently accounts for differences in homograph interpretation in ASD and NT groups. A recent study by Brock et al. (2017) used the homograph paradigm to test children and adolescents with ASD who are speakers of Hebrew, a language that contains many more homographs than English (English has only 4–5 usable pairs). Results showed that performance on the homograph task was best predicted by each participant's score on a picture-naming task. The authors tentatively conclude that the results of their experiment (and perhaps previous homograph paradigm experiments) may reflect a difficulty using semantics to retrieve the appropriate phonological form of the critical word (as in a picture-naming task), not in comprehension ability per se. In another variation, Norbury (2005) had participants listen to sentences that ended in a homophone (words that are pronounced identically, regardless of spelling), such as *Bill stole/fished from the bank*. Participants were then showed a picture of either money or a river (reflecting the two meanings of *bank*) and were asked to decide if the picture matched the sentence. Using this indirect method, the author found no accuracy differences between the NT and ASD group, when the children had not been diagnosed with a comorbid language impairment (LI). Further, these two groups performed significantly better than children with SLI and children with ASD and comorbid LI. Norbury (2005) argues that since WCC predicts that both ASD groups would diverge from NT (and SLI) peers on this task, the results of the study do not support WCC.

Outside of studies on homographs/phones, recent studies have used eye-tracking in a visual world paradigm to provide evidence against WCC (Brock et al. 2008; Bavin et al. 2016). The results of these studies found that children and adolescents with ASD can use a verb's meaning to predict an upcoming target word (e.g., predicting *cake* in *The boy will eat/move the cake*). In fact, in Bavin et al. (2016), the ASD group showed faster processing of the target word with a context-biasing verb compared to a context-neutral condition; in other words, context facilitated processing for them. Instead of diagnosis, Brock et al. (2008) found that standardized language scores better predicted variability in performance, and Bavin et al. (2016) found that scores on an attention assessment covaried with looking patterns. To summarize, studies that have focused on the effect of intrasentential context on lexical access have had mixed findings: some support a WCC account for language comprehension in ASD, while others do not. These latter studies' findings implicate factors aside from WCC, including

comorbid language impairment, lexical production abilities, and attentional differences.

Another productive strain of research centered on WCC and language in ASD groups focuses on the ability of ASD and NT counterparts to make inferences that support the global comprehension of a longer passage. These inferencing studies target the ability to use semantic information in a passage to come to a conclusion, as can be seen in the example (5) from Saldaña and Frith (2007), in which participants use the semantic information in the item to infer that Maria is crying because she is happy to have won the race:

(5) Maria had never won a race before.
The tears streamed down her face.
Can people cry because they are happy?

WCC predicts that participants with ASD would have difficulty connecting separate pieces of semantic information in order to make such inferences, which are necessary for accurate reading/listening comprehension. Results for off-line comprehension measures, usually in the form of comprehension questions after reading/hearing a passage, have been mixed, with some studies finding divergence between ASD and NT groups, thereby supporting WCC (Norbury and Bishop 2003) and others finding no group differences (Saldaña and Frith 2007; Sansosti et al. 2013; Micai et al. 2017). However, on-line measures, in particular eye-tracking during reading, have revealed processing differences between groups, even when no group differences were found in off-line comprehension tasks in the same study. For instance, Saldaña and Frith (2007) found no significant group differences in an off-line comprehension task, but Sansosti et al. (2013), using the same test items in an eye-tracking experiment, found that the ASD group made more fixations and regressions than NT peers while reading the passages. Interestingly, Sansosti et al. (2013), like Saldaña and Frith (2007), did not find any significant differences in the off-line comprehension task. Sansosti et al. (2013) conclude, in agreement with Saldaña and Frith (2007), that while WCC may not prevent ASD groups from making inferences, ASD groups may have difficulty integrating these inferences into the overall interpretation of the text, as evidenced by the increased fixations and regressions in the eye-tracking data. Another study by Micai et al. (2017), found similar results using a longer reading passage, with the ASD group again showing longer gaze durations and more regressions to a target word necessary for making an inference. Taken together, these studies suggest that while WCC may not affect the performance of ASD groups on behavioral tasks, there exist subtle processing differences between ASD and NT groups on the global processing of a text that are not incompatible with WCC. These studies also reinforce the importance of using both off-line and on-line measures, since WCC effects may show up more readily in processing rather than off-line data.

In sum, previous studies focusing on WCC and language behavior in ASD groups show conflicting results. Studies that investigate the ability to use intrasentential context to predict an upcoming word have sometimes shown group differences (Happé 1997; Jolliffe and Baron-Cohen 1999; Lopez and Leekam 2003), although other studies, including those that have used indirect/implicit measures, have found no group differences, instead implicating non-WCC factors in order to explain variability in the data (Norbury 2005; Brock et al. 2008, 2017; Bavin et al. 2016). Studies on general inferencing skills in ASD groups have tentatively supported WCC effects in on-line data, although off-line measures have often failed to find behavioral differences between groups (Saldaña and Frith 2007; Sansosti et al. 2013; Micai et al. 2017).

Pronominal Interpretation: Integrating Structural and Contextual Cues

Previous studies investigating WCC effects on language behavior have largely targeted the effect of semantic/pragmatic information (intrasentential or discourse context) on semantic/pragmatic processes (lexical access, inferencing), while structural information has largely been ignored. Especially given the claim ASD groups diverge from NT counterparts on pragmatic but not syntactic ability (Naigles and Tek 2017), examination of an aspect of language where both structural and contextual pragmatic factors influence interpretation and processing, such as third-person pronoun interpretation, should be useful in investigating WCC effects on language.

Interpretation and Processing of Pronouns in English

Psycholinguistic studies of pronominal interpretation often ground themselves theoretically in some kind of an anaphora hierarchy, a discourse-pragmatic mechanism in which more minimal forms (e.g., pronouns like *he* and *she*) tend to retrieve referents that are highly salient in the discourse, whereas heavier forms (e.g., full noun phrases like *the tall lady*) tend to retrieve less-salient referents (Givon 1983; Ariel 1990; Gundel et al. 1993). While the nature of "highly salient" has been debated, there is psycholinguistic evidence that native English speakers tend to interpret a pronoun as referring back to a local subject (Fukumura and van Gompel 2014), as in (6), where *she* refers to *Julia*.

(6) While Julia watches some birds, she drinks water.

There is also evidence that English speakers initially process a pronoun as referring to a subject even if they later revise their interpretation due to semantic contextual information within the discourse (Karimi and Ferreira 2016). In English, therefore, there is a strong structural component to pronoun interpretation, with grammatical position playing a key role in determining the "salience" of a referent. Nevertheless, this structural preference for a subject can be overridden by contextual pragmatic information, as in (7), where interpreting *she* as *Cassie* is possible due to the contextual information in the first sentence.

(7) Cassie is very thirsty. While Julia watches some birds, she drinks water.

Therefore, both structural (grammatical position of the antecedent) and contextual pragmatic cues can play a role in the resolution of third-person pronoun reference. This provides us with an interesting test case for children with ASD. If central coherence is indeed weaker in ASD than NT groups, ASD participants should prefer the structural bias for a pronoun to refer to a local subject antecedent over pragmatic contextual information biasing some other referent; i.e., they will interpret the pronoun *she* in (7) the same as in (6) more often than NT peers.

Pronouns and ASD

The earliest report on Autism Spectrum Disorder (ASD) describes atypical pronoun use (Kanner 1943). Much subsequent research has focused on the use of first- and second-person pronouns in ASD, showing that individuals with ASD tend to "reverse" pronouns, e.g. they use a second-person pronoun to refer to themselves (Fay 1979; Jordan 1989; Lee and Hobson 1994). But there is also some literature suggesting that individuals with ASD struggle with third-person pronominal reference as well: they produce ambiguous third-person pronouns and/or they avoid third-person pronouns altogether (Novogrodsky 2013; Novogrodsky and Edelson 2016). This research (particularly research that focuses on third-person pronominal reference) has overwhelmingly examined production rather than comprehension (Malkin et al. 2018).

Previous studies examining the production of third-person pronouns during narrative tasks have reported somewhat mixed results: some studies find that ASD groups use more ambiguous pronouns than controls; some find that ASD groups overspecify referents by using full noun phrases (NPs) when pronouns are appropriate; some find both patterns; and some find no group differences at all. Multiple studies have analyzed narrative skills of children with ASD (covering among them ages between 6 and 15 years), and have reported an overproduction of ambiguous third-person pronouns, e.g., using *he* when it is unclear which character *he* refers to (Norbury and Bishop 2003; Novogrodsky 2013; Suh et al. 2014; Banney et al. 2015; Novogrodsky and Edelson 2016). However, some studies have found the opposite, where ASD groups diverge from NT groups not in overusing pronouns, so that their reference is ambiguous, but in overspecifying referents. That is, they tend to use a full NP when NT individuals would use a pronoun instead (e.g., repeating *the boy* several times rather than switching to *he* after the initial mention of the character), with some evidence that group differences may disappear as participants get older (Arnold et al. 2009; Rumpf et al. 2012). Still other studies have found inconsistent referencing patterns in ASD groups. Colle et al. (2008) showed that adults with ASD were more likely than NT adults to produce ambiguous subject pronouns during a narrative generation task in reference to secondary characters, while at the same time referring to the protagonist of the story with a full NP more often than NT peers. Norbury et al. (2014) also found erratic referencing patterns in children and adolescents with ASD, who sometimes used pronouns when the referent was unclear, but at other times generated full NPs or a proper name when a pronoun was expected.

Other narrative production studies report no differences at all in third-person pronoun use between ASD and NT child groups (1995 in English; Makinen et al. 2014 in Finnish et al. 2015 in Dutch). The authors of two of these studies suggest that the lack of divergence may result from immature pronominal development in the NT groups, who were around 7 years old in each case (1995; Makinen et al. 2014). Nevertheless, Kuijper et al. (2015), tested slightly older children (~9 years), and also found no difference between ASD and NT groups in frequency of pronoun vs. full NP use when introducing a character, maintaining reference, and reintroducing a character. Kuijper et al. (2015) did find that working memory played a role in referential choice during reintroduction (a result not differentiated by diagnosis group). Based on this, they argue that it is the cognitive load of keeping track of multiple referents in extended discourse that leads to group differences on third-person referential choice, and that this might explain why previous studies show ASD groups diverging from NT peers on more complex discourse tasks.

While there are numerous studies examining the way that individuals with ASD use third-person pronouns, there are few studies examining how they comprehend/interpret them. One exception is Perovic et al. (2013), whose primary interest was in syntactic, not pragmatic, constraints on the interpretation of pronouns (e.g., *him* in *Bart's dad is pointing to him)* compared to reflexives (e.g., *himself* in *Bart's dad is pointing to himself*), specifically testing Binding Principles A and B. Their results showed no differences among groups on interpretation of non-reflexive pronouns,

Table 1 Summary of demographic information by group. Values represent means with s.d. in parentheses

	Age	Sex	KBIT	CELF	SCQ
ASD	13.60 (1.77)	4F; 14 M	109.22 (17.30)	104.11 (16.48)	19.77 (6.69)
NT	14.13 (1.93)	4F; 14 M	108.89 (16.32)	105.94 (16.83)	3.29 (2.44)
t	−0.848	n/a	0.058	−0.322	7.409
p	0.40	n/a	0.95	0.75	< .001

with all groups showing a similar difficulty rejecting reflexive interpretations. The authors suggest that the difficulty with non-reflexive pronouns in all groups is due to a delay in acquiring adult-like pragmatic pronoun constraints, an established finding in the typically-developing literature known as the Delay of Principle B Effect (c.f., Thornton and Wexler 1999). Since all groups in this study showed this delay in adult-like pronoun interpretation, it remains an open question as to whether this delay would persist for older children with ASD (with or without language impairment), in line with claims that pragmatics is impaired in ASD. Furthermore, Perovic et al. (2013), with their interest in syntactic constraints, did not directly test the influence of discourse-pragmatic factors in pronoun interpretation, presenting single sentences in isolation, rather than in a larger narrative context. Thus, participants did not need to integrate extrasentential information in order to interpret pronouns, which WCC predicts would be more difficult for the ASD groups to do, nor did the authors test or discuss WCC in their study.

Current Study

The current study examines whether older children and adolescents with ASD will use both structural and discourse contextual cues to determine reference for third-person pronouns in the same way as NT peers, in order to investigate the WCC account of language processing in ASD. As far as we are aware, no previous study explicitly testing WCC via language in ASD has investigated pronoun interpretation, instead investigating lexical access and/or global comprehension of a passage. The results of even these studies have been mixed, with some results supporting WCC, and others suggesting that divergence between groups is better explained by other linguistic or cognitive factors. Although previous studies have identified divergences between ASD and NT groups on third-person pronoun *production* in narrative discourse tasks, little research has examined third-person pronoun *interpretation* in ASD populations. The limited previous research on pronoun interpretation in ASD has focused on syntactic constraints on pronouns interpreted in isolated sentences with no narrative context. Further, no work has compared how individuals with ASD integrate discourse context *versus* relying on syntactic constraints to resolve pronominal reference, in order to test whether pronoun comprehension differences provide support for WCC accounts of language processing in ASD.

We hypothesize that, if the WCC account holds, children with ASD in our study will have trouble integrating discourse contextual information into the pronoun interpretation process, and therefore—to a greater extent than the NT group—will tend to choose a local subject antecedent for the pronoun in narratives where context biases a non-subject referent. Because previous studies on WCC and language in ASD have found differing results in online compared to offline measures, we also record eye-tracking data while participants listen to short narratives containing pronouns. Assuming that NT children will use discourse context to interpret pronominal reference, we predict that their looking patterns will differ among conditions: showing one pattern in conditions where context biases pronominal interpretation and another when context is neutral. In contrast, if WCC holds, children with ASD will show similar looking patterns across conditions, suggesting that they do not incorporate contextual bias during their on-line interpretation of pronouns.

Methodology

Participants

Participants included older children and adolescents with ASD (N = 18; 4 female), as well as neurotypical (NT) controls (N = 18; 4 female). Mean age for the ASD group was 13.6 years, with a range of 10.0–17.2 years; for NT the mean was 14.13 years with a range of 10.7–17.8 years. Previous studies suggest that adult-like pronoun use by English-speaking children emerges by approximately age 7 (Rooryck and Wyngaerd 2015; Thornton and Wexler 1999, 1995); the youngest children in our study are well above this age, to ensure homogeneity of the sample in this area. As seen in Table 1, participants were well-matched for age, sex, IQ (KBIT-2; Kaufman and Kaufman 2004), and language ability (CELF-4; Semel et al. 2003); as expected, groups differed significantly on the Social Communication Questionnaire (SCQ-Lifetime; Rutter et al. 2003).

Participants in the ASD group had all received a previous diagnosis of ASD; to confirm this diagnosis, the ADOS-2

Table 2 Sample test item in three conditions

	Introductory sentence	Context sentence	Test sentence	Closing sentence	Question
Neutral	Cassie and Julia are hiking in the woods	The trail is very steep	While Julia watches some birds, *she* drinks water	It's a hot day	"Who drank water?"
Biasing Subject		Julia is very thirsty			
Biasing Other		Cassie is very thirsty			

(Lord et al. 2012) was conducted by a research-reliable administrator. All participants in the ASD cohort earned ADOS-2 scores in the Autism Spectrum range.

Materials

Test items consisted of aurally presented mini-stories containing four sentences (Introductory Sentence, Context Sentence, Test Sentence, Closing Sentence), followed by a question (see Table 2 for an example).

By manipulating the content of the Context Sentence, we created three conditions targeting the use of structural vs. contextual cues in pronominal interpretation: Neutral, Biasing Other, and Biasing Subject. These three conditions are designed to mimic contextual cues that could occur in natural conversation within the framework of controlled experimental manipulation. In the Neutral condition, nothing in the Context Sentence suggests one of the two possible choices over the other as the referent of the pronoun in the Test Sentence. In this case, the default structural constraint preferring a local subject as referent for English pronouns should prevail without interference from context, leading the participant to choose the subject of the subordinate clause of the Test Sentence (*Julia* in the example below). In the Biasing Other condition, the content of the Context Sentence biases the other, non-subject referent for coreference with the pronoun; i.e., specifically stating that "Cassie is very thirsty" in the Context Sentence supports a reading of the Test Sentence in which *she* refers to Cassie, not Julia. In the Biasing Subject condition, both the context and the structural cues bias the subject of the subordinate clause as referent for the pronoun. In addition to the three conditions described above, participants also received ten Filler items (which were identical to the Neutral, but asked about something other than the pronoun referent, such as "Where were they hiking?"), as well as some additional items targeting other aspects of reference not included in the current analysis. As such, a total of 50 items were presented, in pseudorandomized order, with each participant hearing each item one time.[1] While listening to each item, participants viewed a screen with four pictures: Cassie, Julia, an image of the setting, and a non-human object (see Fig. 1 for an example).

This non-human object was always the last word of the subordinate clause of the Test Sentence; thus, all participants were likely to be looking at the same picture leading up to the pronominal subject of the matrix clause. Eye-tracking data was recorded at a rate of 60 Hz using an SMI eye-tracker.

Procedure

All participants were tested at Emerson College and were approved by the Emerson Institutional Review Board. All participants' caregivers provided informed consent. Participants who were 12 years old or older signed an additional informed assent form. Participants were paid $15 an hour (in Amazon.com gift cards) for their time in the lab, which included their participation in the current study and in several other unrelated tasks. Before the main experiment began, participants were trained on the names of the two human figures and completed two practice items. Each trial began with a fixation cross for one second, followed by presentation of the screen with the four images accompanied by the audio of the four sentences of the mini-story. After the mini-story, the screen with the four images disappeared and was replaced by a screen with a written question with three possible answers given in a menu below. For the test conditions, the possible answers were *Julia*, *Cassie*, and *Can't tell*; for Filler items, *Julia* and *Cassie* were replaced with non-animate answers relevant to the filler question (e.g., Question: *Where were Julia and Cassie hiking?* Answers: *the woods*, *the mountains*, *Can't tell*). Participants responded to the question by clicking on their response with a mouse.

Results

Behavioral Results

For data analysis, responses to the question were coded as to whether the participant chose the subject of the subordinate clause (1) or not (0) as the referent of the matrix clause pronoun in the Test Sentence. Figure 2 displays summary statistics for the percent of trials in which participants chose

[1] In the end, only 45 items were included for analysis (9 item types × 5 conditions) due to methodological oversight: For one of the item types, disambiguating information occurred at the final noun in the sentence (the direct object of the verb), rather than at the verb itself as it did for the other nine item types.

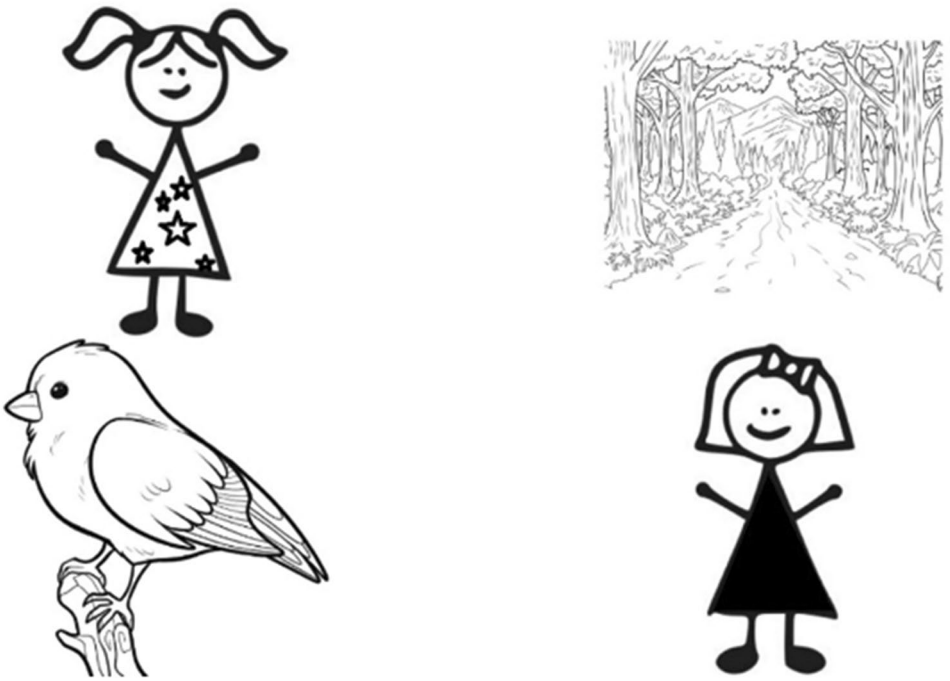

Fig. 1 Sample screen

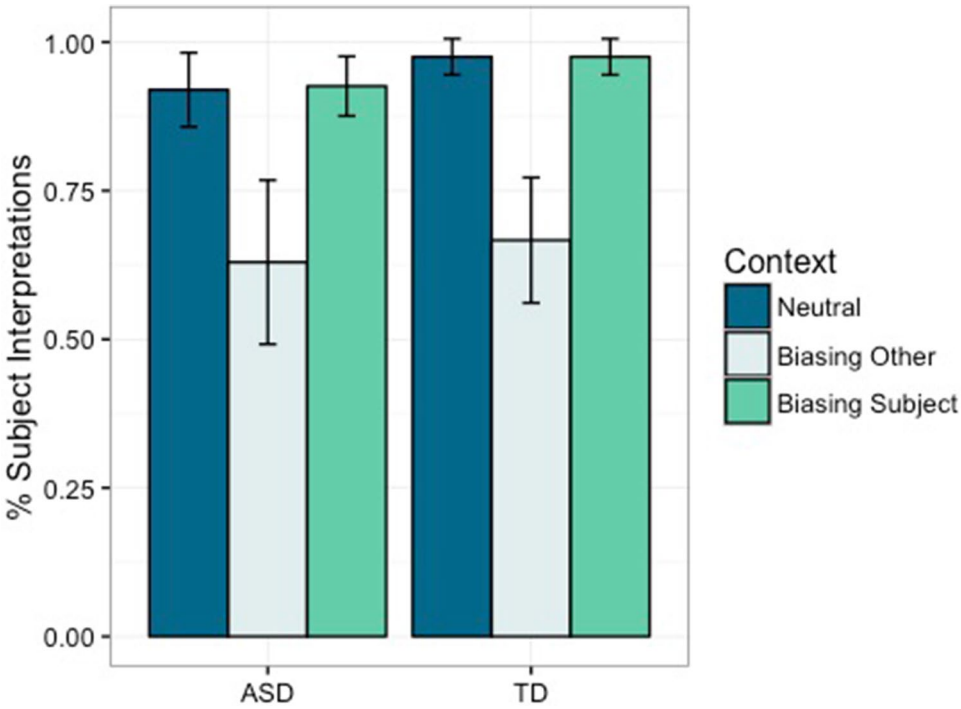

Fig. 2 Percent of trials in which participants chose the subject of the subordinate clause as referent of the matrix subject pronoun, by group and context type. Error bars represent 95% CIs

the subject of the subordinate clause as the referent for the matrix subject pronoun; in other words, the percentage of trials in which participants chose the referent based on the structural cue for a local subject. (Exact numbers are found in Table 6 in the Appendix).[2]

[2] For the vast majority of trials in which the participant did not choose the subject, they chose the other human referent. Although "Can't tell" was also an option, this accounted for only 10 out of 324 Biasing Other trials, 3/324 Biasing Subject trials, and 1/324 Neutral trials. Participants therefore seemed confident in their interpretation choices.

Table 3 Parameters for the best-fitting model predicting participants' choosing the subject of the subordinate clause as the referent for the pronoun in the matrix clause

Model parameters				
Fixed effects	Estimate	SE	*z*	*p*
Intercept	3.245	0.332	9.79	< .001 ***
Other-bias context	− 2.523	0.288	− 8.77	< .001 ***
Subject-bias context	0.067	0.358	0.17	0.853
Random effects	Variance	SD		
Participant	0.669	0.818		
Item	0.176	0.420		

SE standard error, *SD* standard deviation

Data were analyzed using the generalized linear mixed effects model function (glme4) in the R statistics software package (R Core Team 2017; Bates et al. 2012). Random intercepts for Participants and Item were used to control for repeated measures effects. Adding the fixed effect of Context (where Neutral was used as the reference) significantly improved the model compared to that with random intercepts alone ($\chi^2(2) = 156.71$, $p < 0.001$), with the analysis of fixed effects showing that participants chose the subject significantly less often in the Biasing Other condition than in the Neutral and Biasing Subject conditions. A model with Group as a fixed effect did not significantly improve the null model ($\chi^2(1) = 1.76$, $p = 0.18$), nor did a model including both Context and Group as fixed effects significantly improve the model with Context alone ($\chi^2(1) = 1.69$, $p = 0.19$). The final, best-fitting model, which included a fixed effect of Context and random intercepts for Participant and Item, appears in Table 3.

Overall, behavioral results suggest that both ASD and NT groups were able to use both structural and contextual cues to pronoun resolution equally well, with both groups choosing the subject significantly less often when context biased a non-subject referent than in conditions with no bias or where context biased the subject. There were no significant group differences in interpretation.

Eye-Tracking Results

Eye-Tracking Data Processing and Analysis

The eye-tracking analysis focused on looks to the two possible human referents during the matrix clause of the Test Sentence, beginning with onset of the pronoun, continuing through the verb and direct object, and including the pause before the onset of the next sentence. We were particularly interested in how looks to these referents were modulated by the pronoun (which should bias the subject of the matrix clause) and the verb (which contained information that called back to the Context Sentence in the Biasing Other and Biasing Subject conditions, e.g., *is very thirsty* connecting to *drinks*). For preliminary data processing, a value of 1 was assigned to the image the participant was looking at and 0 to the three other images, with choices of SUB (subject), OTH (other, non-subject human), SET (setting), LW (last word of the subordinate clause) for each line of eye-tracking data (~ 1 line recorded every 17 ms). For example, if the participant was looking at the subject of the subordinate clause, the data for that line was coded as SUB = 1, OTH = 0, SET = 0, LW = 0. We excluded individual lines with track loss and trials where tracking ratios were below 70%. Proportion of looks to each possible image was then calculated across 50 ms time bins for the duration of the clause in order to create the time-course figures below.

For our main statistical analysis, we used growth curve analysis (GCA) to model the proportion of looks to the four different pictures (SUB, OTH, SET, LW) across time and to compare looking behavior between groups (ASD and NT) and Context (Neutral, Biasing Other, Biasing Subject) (Mirman 2014). GCA is a multilevel regression technique designed for analysis of data across a time course. In GCA, time is transformed into independent, polynomial vectors. The approach provides a model of the impact of differences between conditions and groups on features of condition curves of movement over time (Mirman et al. 2008). In our analysis, we modeled time as linear and quadratic. For looks to Subject and looks to Other, we were interested in looking behavior during the second, main clause, so we used GCA to analyze proportion of looks from the onset of the pronoun *she* to 2 s later, including the pronoun, verb, direct object, and a brief pause before the following sentence. Fixed effects were Group (ASD and NT) and Context (Neutral, Biasing Other, Biasing Subject). For all four comparisons, the

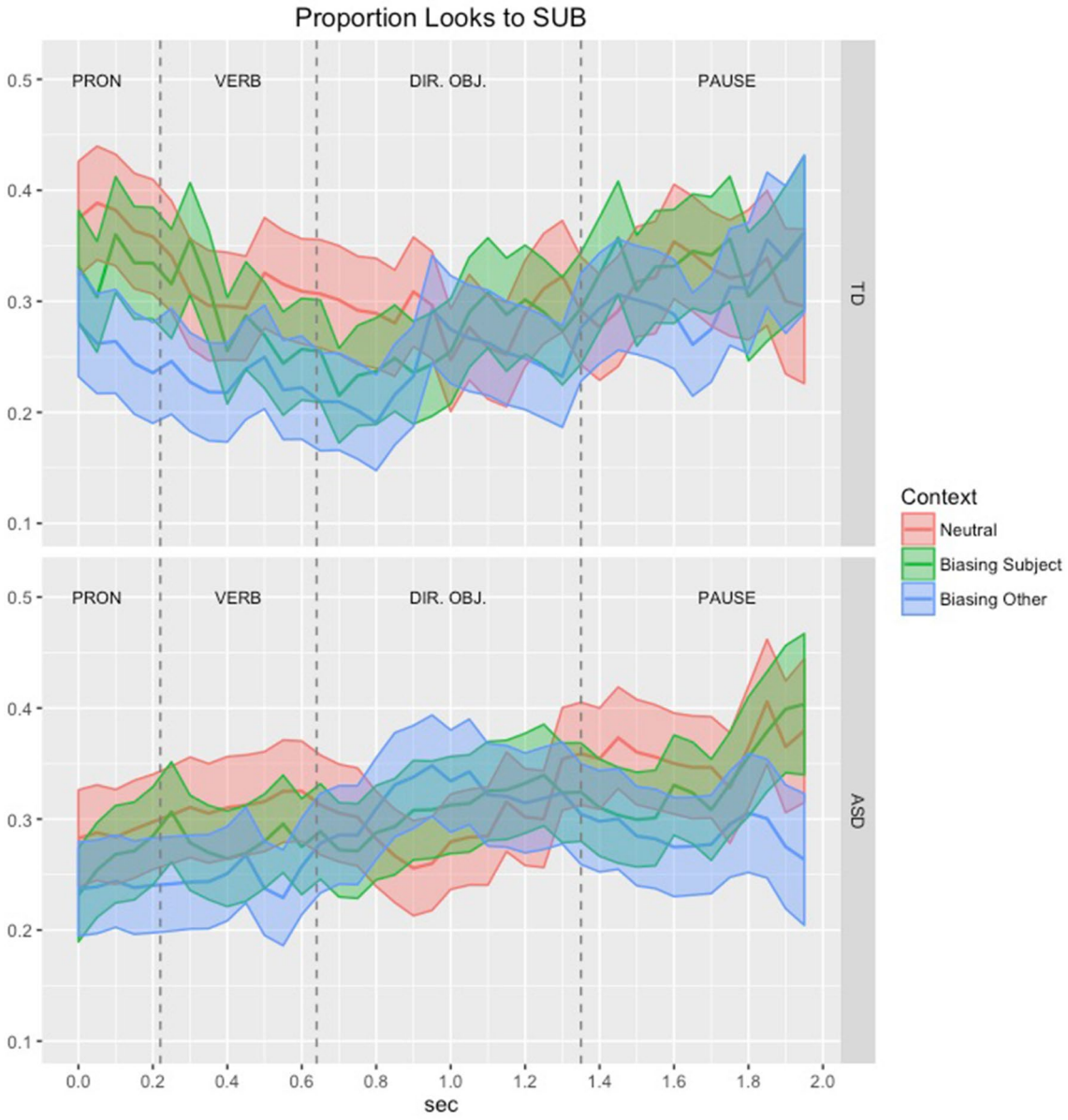

Fig. 3 Proportion of looks to the subordinate clause subject (SUB) during the matrix clause

reference level for Group was NT and the reference level for Context was Neutral.

Looks to Subject (SUB)

Time-course graphs for proportion of looks to the subject of the subordinate clause (SUB) by each group during the matrix clause are displayed in Fig. 3. The results of the full GCA model are presented in Table 4.

The model shows [1] a significant positive effect of Group on the linear term ($t = 2.00$, $p = 0.045$). This effect reflects relatively negative slopes in looks to subject for NT participants for all trial types. The NT group begins the trial by looking at the subject soon after hearing the pronoun, and then they switch their gaze to other pictures. The ASD group shows a more positive slope in looks to subject, overall, where they look more at the subject pictures at the end of the trial (2 s after the pronoun plays) than they do when the pronoun is first played. Across groups, there is a positive effect on [2] the linear term and [3] the quadratic term when the context is Biasing Other as compared to Neutral ($t = 13.95$, $p < 0.001$ and $t = 6.47$, $p < 0.001$). The linear effect reflects a pattern where the proportion of looks to subject are relatively low right after the pronoun is played for Biasing Other

Table 4 Results of GCA model for looks to Subject

Group			*Context*				*Group x Context*		
Estimate (SE)	*t*-value	*p*-value	Biasing	Estimate (SE)	*t*-value	*p*-value	Estimate (SE)	*t*-value	*p*-value
Fixed									
0.041 (0.056)	0.721	0.436	Subject	0.045 (0.060)	0.721	0.454	−0.009 (0.083)	−0.102	0.919
			Other	0.014 (0.054)	0.255	0.799	−0.030 (0.075)	−0.393	0.694
Linear									
0.213 (0.107)	2.001	0.045*	Subject	−0.073 (0.027)	−2.656	0. 008**	0.082 (0.038)	2.149	0.032*
			Other	0.387 (0.028)	13.95	0.00E+0***	−0.423 (0.039)	−10.85	0.00E+0***
Quadratic									
0.074 (0.104)	0.709	0.478	Subject	0.027 (0.028)	0.977	0.478	−0.092 (0.038)	−2.405	0.016*
			Other	0.180 (0.028)	6.468	9.96E−11***	−0.384 (0.039)	−9.304	0.00E+0***

*$p < 0.5$; **$p < 0.01$; ***$p < 0.001$

trials. Later in the trial, looks to subject increase. The positive quadratic effect suggests that proportions of looks to subject show a more U-shaped pattern across the trial (from pronoun onset to 2 s later) for items that bias the other referent. This reflects a pattern that is more evident in the NT group, as seen in Fig. 3, which suggests that this group is driving the overall effect; for the Biasing Other condition, the proportion of looks to subject start off relatively high right after the pronoun is played, decrease quite a bit (to about 15%) and then increase again (to about 35%).

That the NT group may be driving the former two effects is supported by negative interactions between Group and Context on [4] the linear term and [5] the quadratic term, when the Context was Biasing Other ($t = -10.85$, $p < 0.001$; $t = -9.30$, $p < 0.001$). The linear effect indicates a more negative slope for the ASD group for Biasing Other items as compared to Neutral items, across the trial. This effect reflects a relatively *positive* slope for Neutral items in the ASD group (beginning around 28% and ending at around 38%), so that Biasing Other items have a more negative slope in comparison. This is not true for the NT group, where Neutral items have a slightly negative slope (proportions of looks to subject start off at around 38% and end at around 30%, on average). The negative quadratic effect reflects a pattern whereby participants with ASD show a peak in proportions of looks to subject for Biasing Other trials about one second after the pronoun is played (with looks beginning to rise around verb onset), whereas the NT group shows no such peak.

Results also show [6] a negative effect on the linear term when the context is Biasing Subject ($t = -2.66$, $p = 0.008$); this likely reflects the negative slope found in the first half of the clause in the NT group. Indeed, a [7] positive interaction of Group and Context on the linear term for Biasing Subject trials ($t = 2.15$, $p = 0.032$) suggests that slopes are more positive for the ASD group in this condition compared to the Neutral condition, unlike the NT group. Similarly, there was also [8] a significant negative interaction between Group and Context on the quadratic term when the condition was Biasing Subject ($t = -2.41$, $p = 0.016$). In the ASD group, there is again a rise in looks to the subject referent culminating in a small peak at the end of the verb. Looks to subject continue to rise to about 35% at the end of the sentence and continue to rise in the pause to approximately 40% after the end of the sentence. The NT group, on the other hand, shows a downward trend from onset of the matrix clause culminating in a slight trough during the verb.

Looks to Other (OTH)

Time-course graphs for looks to the other, non-subject human referent (OTH) by each group during the matrix clause are displayed in Fig. 4. The results of the full GCA model are presented in Table 5.

Across groups, there is a significant negative effect on both [9] the linear and [10] the quadratic term when Context is Biasing Other ($t = -12.17$, $p < 0.001$; $t = -6.76$, $p < 0.001$) and [11] the linear and [12] quadratic terms in Biasing Subject ($t = -2.01$, $p < 0.044$; $t = -9.21$, $p < 0.001$), as compared to Neutral. The linear effects reflect a more positive slope in looks to Other in the Neutral context, and a relatively negative slope in the other two contexts. The quadratic effects suggest that looks to the other referent show a relative decrease in the center of the trial epoch (a U-shaped pattern) for Neutral items and a relative peak in the center of the trial epoch (an upside-down-U-shaped pattern) for Biasing Other and Biasing Subject items.

There is [13] a significant negative interaction between Group and Context for Biasing Subject ($t = -7.35$, $p < 0.001$) on the linear term. In the ASD group, looks to the other decrease across the trial epoch for Biasing Subject items, while looks to the other increase across the trial epoch for Neutral items. Conversely, there is [14] a significant

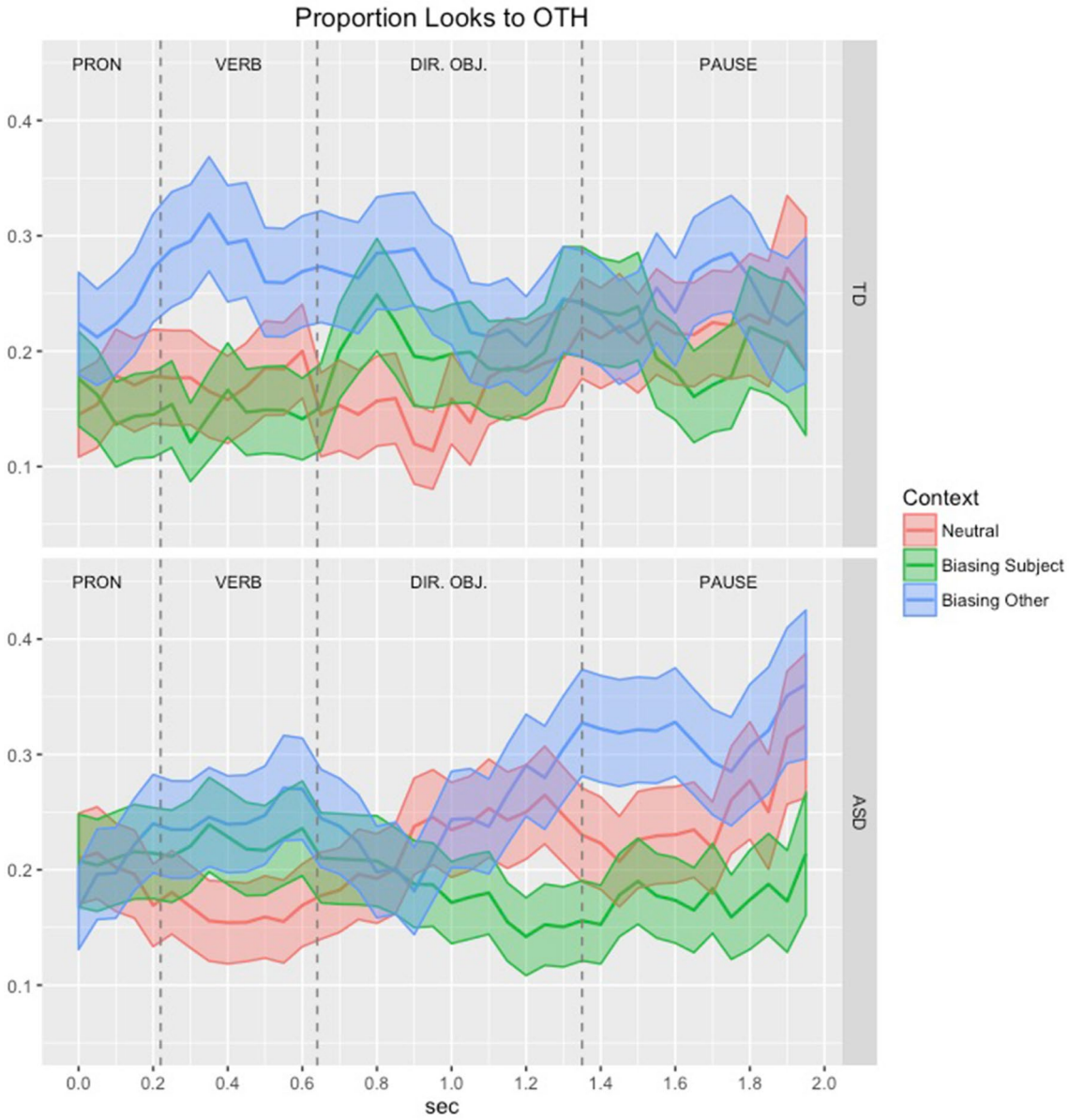

Fig. 4 Proportion of looks to the other human referent (OTH) during the matrix clause

positive interaction between Group and Context for Biasing Other ($t = 10.46$, $p < 0.001$). In the ASD group, looks to the other steadily increase across the trial epoch for Biasing Other items. The slope of this increase is steeper than it is for Neutral items, particularly towards the end of the trial. In the NT group, looks to the other show the steepest linear increase for Neutral items.

Finally, there are significant positive interactions between Group and both [15] Biasing Other and [16] Biasing Subject on the quadratic term ($t = 5.74$, $p < 0.001$; $t = 6.13$, $p < 0.001$). These interactions reflect a relatively U-shaped distribution (a positive quadratic pattern) for Neutral items in the NT group: looks to the other decrease in the center of the trial epoch and then increase at the end for these items. Biasing Subject items show the opposite distribution for the NT group (a peak in the center of the trial), and Biasing Other shows an early peak. In contrast, looks to Other in the ASD group show an upside-down-U-shaped distribution (a peak in the center of the trial) for Neutral items, and relatively low proportions of looks to Other (less than 20% on average) in the center of the trial epoch for both Biasing Subject and Biasing Other items. Unlike the NT group, the ASD group shows no peaks in these two conditions.

Table 5 Results of GCA model for looks to other

Group			*Context*				*Group x Context*		
Estimate (SE)	*t*-value	*p*-value	Biasing	Estimate (SE)	*t*-value	*p*-value	Estimate (SE)	*t*-value	*p*-value
Fixed									
0.011 (0.063)	0.175	0.861	Subject	− 0.039 (0.053)	− 0.074	0.461	− 0.034 (0.074)	− 0.472	0.962
			Other	− 0.015 (0.062)	− 0.249	0.803	0.013 (0.086)	0.150	0.881
Linear									
− 0.053 (0.092)	− 0.573	0.567	Subject	− 0.050 (0.025)	− 2.012	.044*	− 0.255 (0.035)	− 7.353	1.94E−13***
			Other	− 0.306 (0.025)	− 12.17	0.00E+00***	0.370 (0.035)	10.46	0.00E+00***
Quadratic									
− 0.128 (0.081)	− 1.681	0.114	Subject	− 0.230 (0.025)	− 9.212	0.00E+00***	0.213 (0.035)	6.127	8.97E−10***
			Other	− 0.171 (0.025)	− 6.760	1.38E−11***	0.203 (0.035)	5.736	9.68E−09***

*$p < 0.5$; **$p < 0.01$; ***$p < 0.001$

Preliminary Discussion of Eye-Tracking Results

In the Neutral condition, NT processing of the subject referent during the matrix clause shows little effect of the pronoun itself, with looks to the subject referent image (SUB) trending down throughout the duration of the clause (as seen in the effects labeled [1] and [4] above); looks to the other possible human referent (OTH) are also relatively low (as seen in effects [15] and [16]), suggesting that the NT participants are not struggling to identify a referent for the pronoun during this clause. This is not surprising, as the function of a pronoun in English is to maintain reference to the current subject referent; as such, the pronoun does not trigger a new round of processing for the subject of the matrix clause. The NT group is therefore free to look at other pictures (see the section on Follow-up Analyses below). In the ASD group, on the other hand, there is evidence of processing of the pronoun during the matrix clause, with a rise in looks to the subject referent culminating in a small peak at the end of the verb, accompanied by a similar drop in looks to the other human referent. After a brief period where participants seem to be looking at the subject and other at almost equal rates, looks to subject continue to rise to about 35% at the end of the sentence, and continue to rise during the pause to ~ 40% (effects [1, 4, 15, 16] above). There is therefore evidence that the ASD group is spending more time on processing the subject referent during the matrix clause than the NT group (including brief consideration of the other human referent).

Although the Biasing Other condition was predicted to reveal divergence between the ASD and NT groups, due to the need to integrate the context sentence into the interpretation of the pronoun, the two groups behaved identically in the behavioral data, resolving the matrix clause pronoun *she* as a non-subject referent for a greater percentage of trials compared to the other two conditions. However, there are some differences between the two groups in the eye-tracking data for this condition. The "other" interpretations in the behavioral data are reflected in the matrix-clause eye-tracking data for the NT group, with elevated looks to the non-subject, context-biased other referent (effect [15]) and decreased looks to the subject of the subordinate clause (effects [2–5]). For the ASD group, although looking patterns do differ in Biasing Other condition compared to the Neutral, there is nevertheless a period during the direct object in the former condition when looks to the subject are preferred (effect [5]) while looks to the other are dispreferred (effect [15]), suggesting a stronger consideration of the subject of the subordinate clause as the antecedent of the pronoun compared to the NT group. In particular, the looks to subject in the ASD group during the Biasing Other condition begin to rise at verb offset, the exact moment that the contextual information ties into the matrix clause, suggesting that the ASD group takes strong notice of the discrepancy between the structural cue (preference for subject) and discourse contextual cue (preference for other).

Unsurprisingly, in the behavioral data, both groups overwhelmingly chose the subject of the subordinate clause as the referent for the subject pronoun in the matrix clause in the Biasing Subject conditions. Eye-tracking data, however, once again revealed divergence in processing. Somewhat unexpectedly, the NT group showed decreased looks to the subject referent (effects [6, 7]) and increased looks to the non-subject referent (effect [16]), with looks to the non-subject referent showing a peak slightly after verb offset compared to a trough around the same time period in the Neutral. This may be evidence of a Repeated Name Penalty effect lingering into the matrix clause. The Repeated Name Penalty (RNP) is a well-documented effect, in which the repetition of a name rather than a switch to a pronoun or otherwise less-informationally-heavy referring expression causes a processing penalty (Gordon et al. 1993; Almor 1999). While it would be unsurprising if the NT group showed evidence of this effect during the subordinate clause (in which the

named subject is identical to the subject of the previous sentence), our results suggest that the effect lingers even after reference has been continued with a pronoun in the matrix clause, with decreased looks to the local subject referent (effect [6]) and increased looks to the other human referent (effect [16]) compared to the Neutral condition.[3] For the ASD group, however, there is little evidence of a RNP effect in the Biasing Subject condition. In fact, compared to the Neutral condition, the ASD group spends more time looking at the subject referent (effects [7, 8]) and less time at the other human referent (effects [13, 16]), suggesting that the repetition of the name in the subordinate clause may facilitate reference resolution at the pronoun in the matrix clause.

Follow-Up Analyses

Although we set out to examine looks to the two possible human referents during the matrix clause, in order to follow up on data from the Biasing Subject condition, which suggested an RNP effect lingering into the matrix clause for the NT but not the ASD group, we created graphs of the subordinate clause to look for evidence of a processing cost at the repeated name itself. Results, seen in Figs. 5 and 6 below, support a RNP in the NT group, with looks to the subject image greatly attenuated in the Biasing Subject condition compared to the Neutral. Looks to the other human referent are also elevated compared to the Neutral throughout the clause. For the ASD group, looks to the subject are similar in both Biasing Subject and Neutral conditions, and although the ASD group shows a brief peak of looks to the other referent, these looks quickly trail off to fall below their level in the Neutral condition.

Due to the low number of looks to the two human images in the Neutral condition for the ASD group at the onset of the matrix clause, we also examined the two non-human images, as seen in Fig. 7. The graph shows the ASD group almost split between looks to the subject (SUB) and the last word of the previous clause (LW), with looks to the Setting (SET) not far behind. The NT group, on the other hand, clearly prefers the subject referent going into the matrix clause, then switches to non-human images. Graphs of the subordinate clause (seen in Fig. 8) reveal that for the ASD group, looks to the LW begin to rise right around the time this word is spoken, and overtake looks to the subject by the end of the clause; meanwhile, the NT group shows a strong increase and peak in looks to the subject (and therefore the referent of the upcoming pronoun) during the second half of the subordinate clause.

General Discussion

Implications for WCC

We hypothesized that if the Weak Central Coherence account of language processing in ASD holds, the ASD group would have difficulty integrating discourse contextual information into the pronoun interpretation process, instead relying more heavily on structural constraints. Behavioral data did not support this account, with the ASD group using context to override the structural preference for a local subject antecedent in ~ 30% of Biasing Other trials, comparable to the NT group. Furthermore, eye-tracking patterns for the ASD group differed in the two context-mitigated conditions compared to the Neutral condition, suggesting that discourse context also influences on-line interpretation. These results contrast with previous studies that suggest ASD groups are unable to use discourse context information to predict the correct reading of a homograph (Happe 1997; Jolliffe and Baron-Cohen 1999; Lopez and Leekam 2003), although they are in line with other studies in which ASD groups could use verb semantic information to predict a later part of the sentence (Brock et al 2008; Bavin et al. 2016), and with studies that suggest factors other than WCC lead to divergence between ASD and NT groups on homograph/homophone tasks (Norbury 2005; Brock et al. 2017). Furthermore, our behavioral results, which suggest comparable pronoun interpretation in the ASD and NT groups, contrast with findings from previous studies on pronoun *production* in ASD, which have reported either an increased use of ambiguous pronouns and/or infelicitous repetitions of full noun phrases in narratives (e.g., Colle et al. 2008). Our behavioral results suggest that divergence in pronoun production in ASD groups is not rooted in a fundamental deficit in English pronoun use, as the ASD group in this study showed both the ability to follow the general tendency for a pronoun to retrieve a local subject, and the ability to override this tendency using discourse context information. In short, the behavioral results provide evidence that older children with ASD (at least those who test as having at least normal IQ and language) do understand how pronouns work, although they may struggle to apply this knowledge during production.

On the other hand, our eye-tracking data complicate this story somewhat. Although the ASD group did show differences in looking patterns across conditions, suggesting that discourse context does affect processing to some extent, the data also reveal significant differences in eye-tracking patterns across groups, suggesting that ASD and NT groups

[3] It may also be possible that the repetition of the name in the subordinate clause beginning with "While" sets up some sort of expectation of a contrast in the matrix clause, resulting in surprise when a pronoun appears in subject position rather than the name of the other referent. Since we did not set out to test the RNP in this experiment, we refrain from further speculating here.

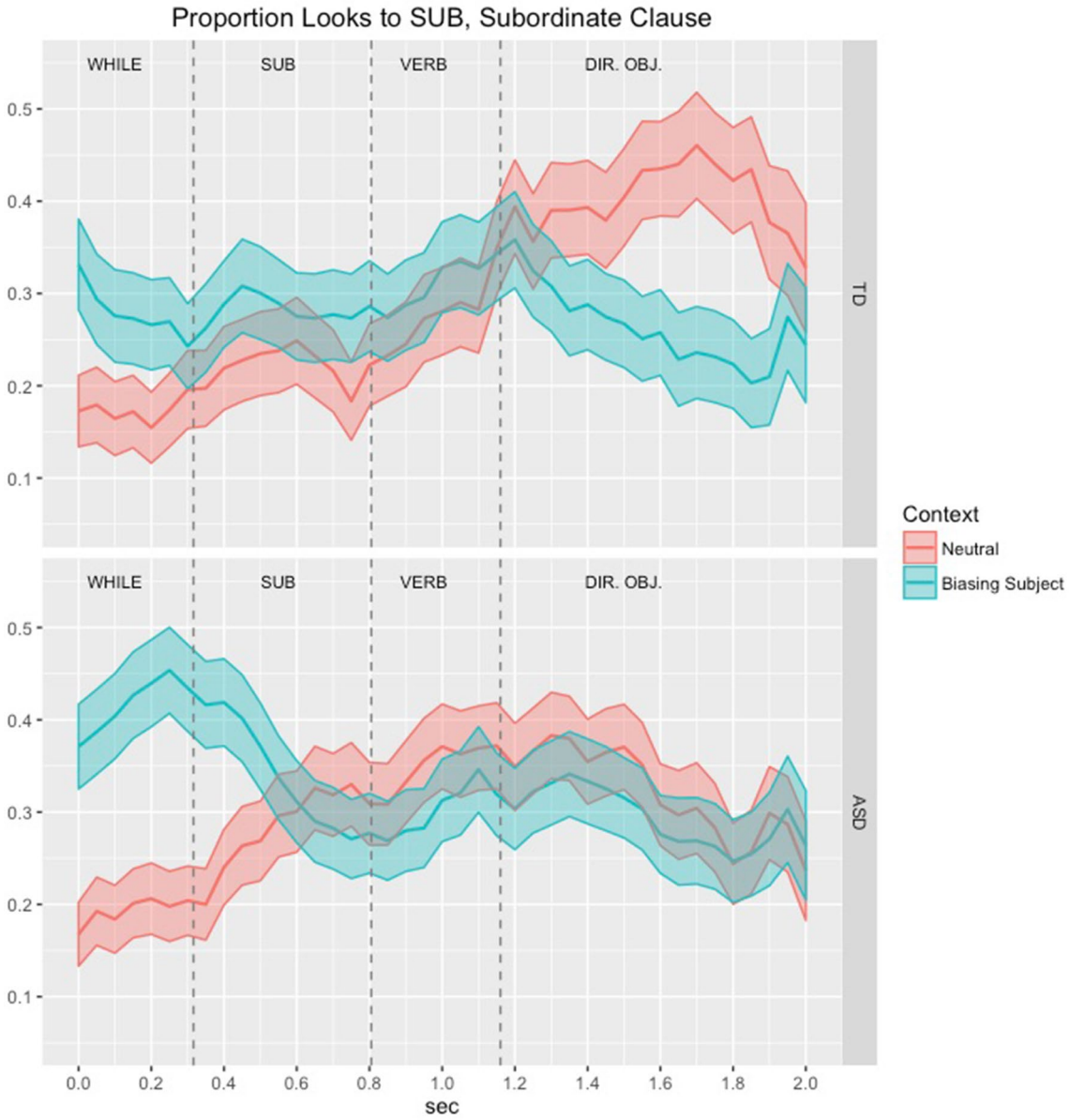

Fig. 5 Proportion of looks to the subject of the subordinate clause (SUB) during the subordinate clause

may be getting to the same interpretations through different processing paths. In the Neutral condition, increased looks to the subject referent (the ultimate choice for interpretation for both groups in over 90% of trials) was delayed in the ASD group compared to the NT group, despite the fact that this condition is most similar to natural speech. In both the Biasing Other and Biasing Subject conditions, the ASD group had a tendency to look at the subject referent to a greater extent that the NT group.

The fact that the ASD and NT groups diverged on the eye-tracking but not the behavioral data suggests that although the ASD group is capable of interpreting pronouns like the NT group, processing differences between the two groups may lead to divergences in production. The need for utterance planning during production may draw more strongly on cognitive resources than comprehension. It is possible that this differential cognitive load explains the group differences found in previous studies on narrative production. Furthermore, while processing differences did not lead to divergent interpretations in the current study, it may be the case that our items were not complicated enough for processing differences to accumulate to the point where they affect interpretation. The narratives used in our interpretation task were relatively short (four sentences each) and uncomplicated (e.g., there were only two characters, and each story contained the same sequence of predictable sentence structures).

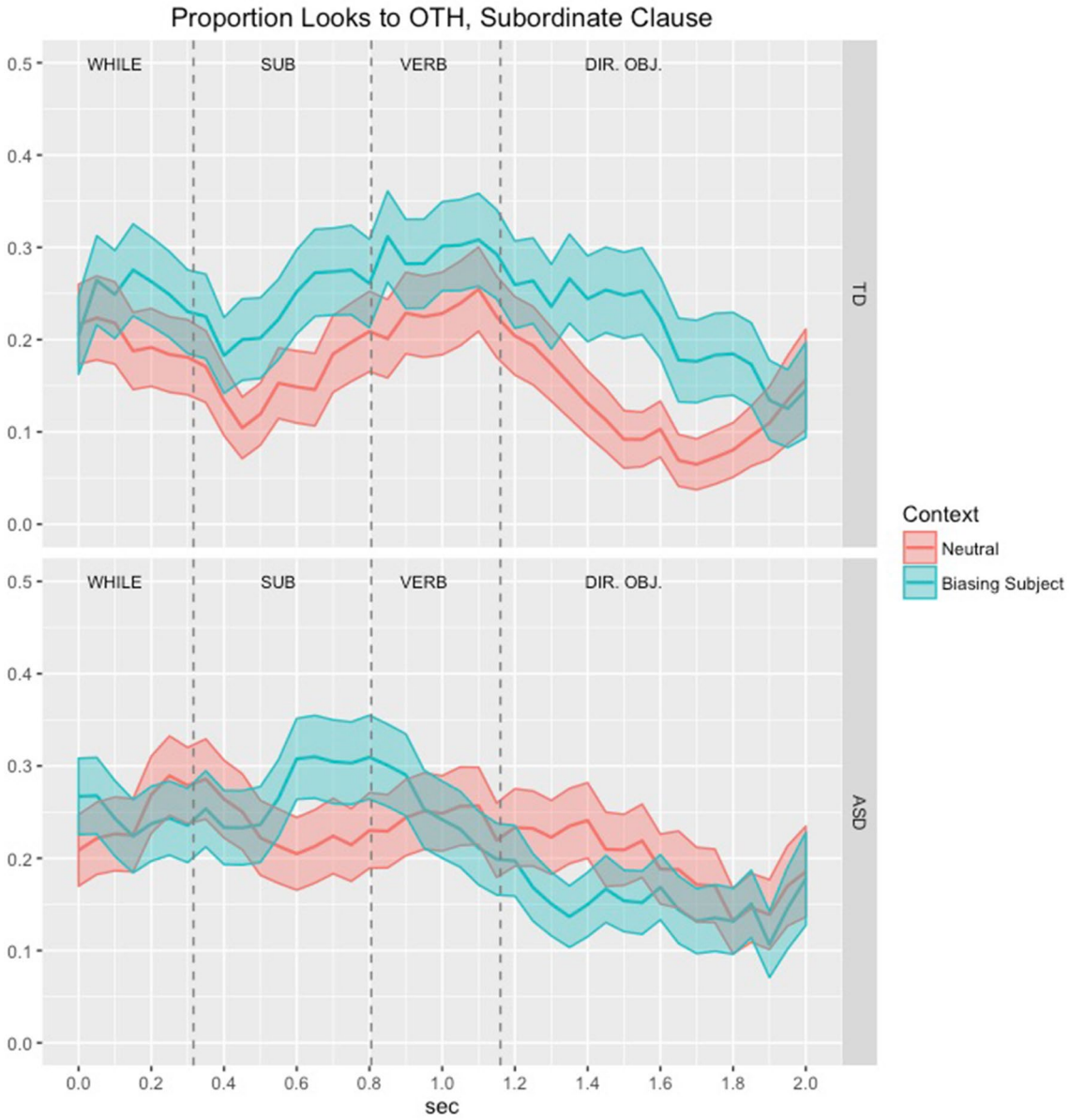

Fig. 6 Proportion of looks to the other human referent (OTH) during the subordinate clause

Interpretation itself was elicited by a closed-choice question, presented visually, after the story was completed, with unlimited time to provide an answer. This procedure is very different from what happens when listeners must interpret pronouns during real-world interactions, in which listeners must constantly update their discourse representations without a processing break during which they can think back and determine reference. While many of the features of our story stimuli—their simplicity, brevity, and predictability—can help explain why our interpretation results did not differ between groups, these same features make our eye-tracking results even more striking. The fact that there were significant processing differences for the two groups for even these simple, short narratives, is noteworthy, and further supports the possibility of divergence in interpretation as processing differences build up across more complicated narratives. Firsthand reports from people with ASD further suggest that this may be the case; in an account of using/understanding pronouns, E. J. Grace, an adult with ASD states, "When my friends tell me a story with … more than two characters, I may get lost in the pronouns" (2013).

A Preference for the Structural Cue

Overall, it is too simple to say that our processing results suggest that "pragmatics is impaired" in the ASD group.

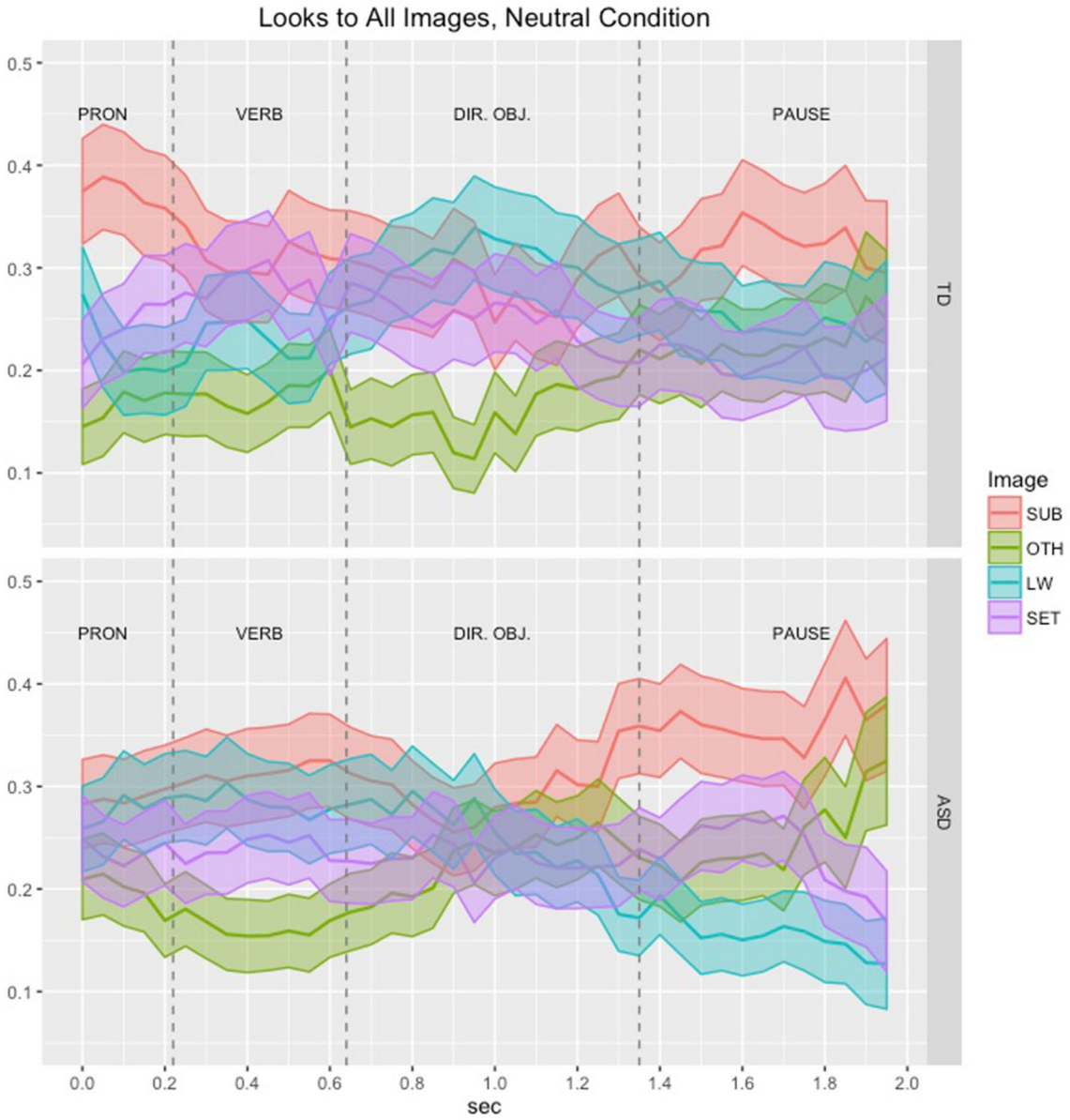

Fig. 7 Proportion of looks to each of the four images during the in the Neutral condition only. *SUB* subject of subordinate clause, *OTH* other human referent, *LW* last word of the subordinate clause, *SET* setting

Discourse context does indeed influence both the ultimate interpretation and the on-line processing of the pronoun for the ASD group, so the processing of anaphora clearly happens in context for both groups. However, the processing data does suggest a strong preference for the local subject cue in the ASD group, especially in the two context-manipulated conditions. In the condition where context biased the other, non-subject referent, the ASD group showed a window from verb offset where the subject was clearly preferred (to a greater extent than in the Neutral), whereas the NT group looked back and forth evenly between the subject and the other referent. In the condition where context biased the subject of the subordinate clause, the NT group seemed to show a Repeated Name Penalty effect (due to the same name being repeated as the subject of both the Context Sentence and the subordinate clause of the Test Sentence), which lingered into the matrix clause, resulting in decreased looks to the subject referent and increased looks to the other referent compared to the Neutral condition. The ASD group, on the other hand, once again looked at the subject referent to a greater extent than they did in the Neutral condition, suggesting that the repetition of the subject, rather than causing a processing penalty, facilitated processing. As such, in both of the context-mitigated conditions, despite the behavioral

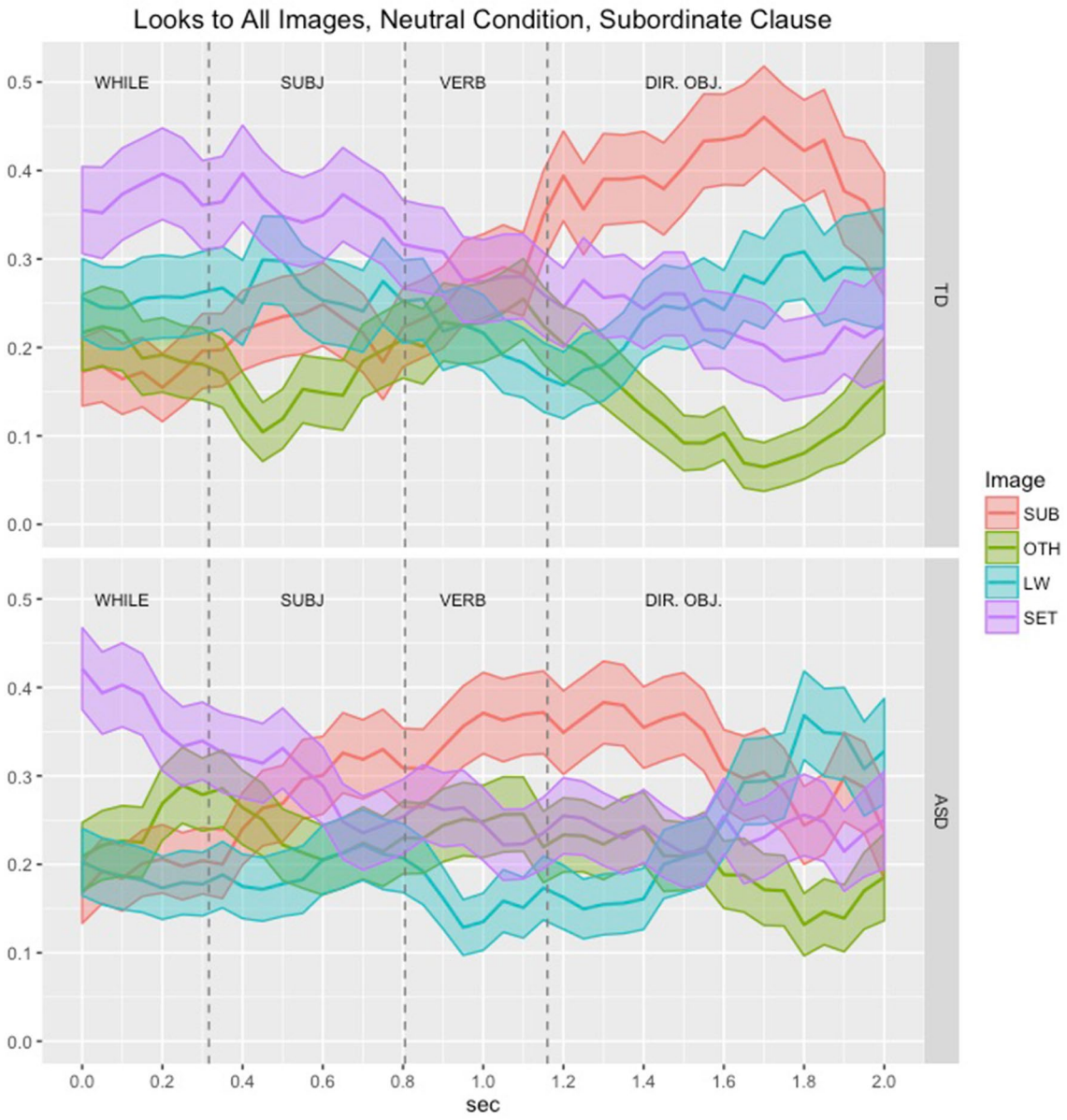

Fig. 8 Proportion of looks to each of the four images during the subordinate clause in the Neutral condition only. *SUB* subject of subordinate clause, *OTH* other human referent, *LW* last word of the subordinate clause, *SET* setting

results, the ASD group's looking patterns showed a stronger preference for the subordinate clause subject, the antecedent that is biased by the structural cue.

One possible explanation for this preference is that the ASD group represents the structural bias for a local subject as a purely syntactic rule, without the pragmatic underpinnings that are theorized to underlie anaphora interpretation. Thus far, we have referred to the tendency for pronouns in English to retrieve a local subject antecedent as a "structural cue" (since it involves grammatical position), but this tendency has been characterized as one rooted in discourse-pragmatics. As mentioned in the Introduction, discourse-pragmatic theories of anaphora stress the connection of light anaphors (like pronouns) to "highly salient" referents (such as the local subject). Levinson (2000) characterizes this connection as a kind of default conversational implicature arising from a maxim derived from the second half of Grice's Maxim of Quantity.[4] Repeating the same full noun phrase as the subject of two consecutive clauses is being "too informative"; using a semantically-blanched,

[4] "Do not make your contribution more informative than is required" (Grice 1975, p. 45).

"light" pronoun for the second mention is preferred. Hence the *she* in (6) (repeated below) is implicated by its lightness to refer back to *Julia*, the most salient referent (in English the local subject), and not to some referent previously mentioned in the discourse, or to some new, previously-unmentioned referent. While this conversational implicature holds generally, in cases where discourse context biases an antecedent other than a local subject, such as our Biasing Other condition (see (7), repeated below), the implicature may be canceled.

(6) While Julia watches some birds, she drinks water.
(7) Cassie is very thirsty. While Julia watches some birds, she drinks water.

In the NT group, this cancellation manifests itself straightforwardly in increased looks to the other, non-subject referent during pronoun processing (e.g., *Cassie* in (7)). However, looking patterns suggest that it is relatively difficult for our ASD participants to disprefer the subject referent (e.g., *Julia* in (7)). If we assume that it is more costly to break a syntactic rule than to cancel a pragmatic implicature, then differences between ASD and NT groups can be attributed to differences in how each group represents the tendency for pronouns to refer to local subjects: in the discourse or in the syntax, respectively. As such, the elevated looks to the subject by the ASD group in the Biasing Other condition may reflect their difficulty to switch from the referent that is determined by the syntax. In the Biasing Subject condition, the NT group seems to show influence of an RNP effect in the subordinate clause that lingers into matrix clause pronoun interpretation, resulting in decreased looks to the subject and increased looks to the other referent continuing into the matrix clause. The processing data for the ASD group, on the other hand, does not seem to show as strong a reaction to the repeated name. Since, in this condition, the discourse context and the syntactic rule match (and the RNP effect apparently does not arise[5]), processing of the pronoun is facilitated rather than hindered, resulting the ASD group looking more at the subject and less at the other referent in Biasing Subject compared to the Neutral condition.

Possible Executive Functioning Effects

However, there is another possibility that explains why the ASD group eye-tracking data (seemingly) does not conform to what is expected in discourse-pragmatic theories of anaphora: not a mis-representation of the local subject antecedent preference as a syntactic rule, but rather differences in the degrees of saliency of the referents, exacerbated by executive functioning differences. There is evidence that individuals with ASD show executive dysfunction (Hill 2004), including deficits in working memory (Wang et al. 2017) and problems with cognitive flexibility, including issues with "sticky attention" (Kleinhans et al. 2005). Previous studies on pronoun production (Kuijper et al. 2015) and the ability to use context to predict an upcoming word (Bavin et al 2016) have implicated working memory and attentional differences, respectively, in accounting for task performance variability in their studies. Thus, it is possible that processing differences in the current study may also be driven by executive functioning differences, rather than differences in whether underlying grammatical representation is syntactic versus pragmatic. The elevated looks to the subject referent by the ASD group in both the Biasing Other and Biasing Subject condition may be attributed to issues with working memory; attentional differences also seem to arise in the Neutral condition, explaining delays in looks to subject for the ASD group (the latter of which is difficult to explain under a syntactic representation account).

In the Biasing Subject condition, the ASD group, unlike the NT group, did not show a RNP effect in either the subordinate clause (where the repeated name occurred) or the matrix clause (where the pronoun occurred); instead, looks to the subject referent were higher than in the Neutral condition. These results recall those of Almor et al. (1999), who, comparing participants with Alzheimer's Disease (AD) to age-controlled typical peers, found that AD participants processed short narratives faster when a full subject NP was repeated compared to control items where the NP was replaced by pronouns after the first sentence. In other words, the AD group not only failed to show an RNP effect, but showed *facilitated* processing when the name was repeated. Given that this result and the results of other reference resolution tasks in the same study were correlated with working memory scores, the authors conclude that "the working memory impairment in AD leads to an overall decrease in the activation of referents, therefore enabling costly referring expressions (full NPs vs. pronouns) to attain more functionality for AD patients than for healthy participants" (p. 764). Given that our ASD participants similarly showed no RNP effect and instead showed facilitated processing for repeated-name trials, a similar working memory effect may be responsible. Similarly, in the Biasing Other condition, while the other referent is mentioned in the Context Sentence (***Cassie is very thirsty***), the subordinate clause with the subject referent (*While Julia looks at some birds…*) intercedes between the Context Sentence and the pronoun in the matrix clause (*…she drinks some water*). It is possible that the activation of the other referent degrades faster in working memory in

[5] Since the RNP itself can be described as a Maxim of Quantity violation (Almor 1999), the lack of an RNP effect in the ASD group could be further evidence that the ASD group prioritizes syntax (a subject is a subject) over pragmatics (different anaphors tend to be used with referents of different saliency) in anaphora interpretation.

the ASD group compared the NT group, exacerbated by both the linear distance from the pronoun and interference from the intervening subordinate clause subject, resulting in greater looks to the subject referent during the matrix clause. (Note that the Neutral results discussed below also suggest attentional interference, in that case from an intervening non-human referent). In other words, the ASD group may not represent the structural cue any differently than the NT group as speculated in the previous section; it may simply be the case that working memory differences interfere with the activation levels of each referent, resulting in divergent processing data. In situations where increased processing strain taxes working memory even further (including the interpretation/production of longer, more complex narratives than those used in the present study), differences in processing may accumulate and lead to divergent pronoun interpretation/production, as has been found in previous studies. That the ASD group performs similarly to the NT group in our behavioral task, with its simple narratives, suggests that ultimately it is processing differences—rather than a fundamental mis-representation of the rules governing pronouns—that lead to divergent eye-tracking data.

Eye-tracking patterns in the Neutral condition provide even more evidence that group differences come down to executive functioning differences, rather than in the way that the two groups represent the relationship between a pronoun and its referent. If the ASD group were to represent the structural preference for the local subject antecedent as a simple syntactic rule, they should show attentional preference to the subject referent in all conditions. However, in the Neutral condition, the ASD group was actually *slower* to visually attend to the subject referent, even though this condition represents discourse that is closest to natural speech (the pronominal antecedent is the local subject and there is no competing contextual information). We argue that their relatively slow fixation to the subject referent can be attributed to differences in attentional allocation. In this condition, the ASD group seems to struggle to reallocate attention to the characters/objects mentioned in the current, matrix clause because they are still focused on information mentioned previously, specifically the last word of the subordinate clause. That is, they start looking at the picture this last word refers to (e.g., *bird* in our example item) when the word is said, and they continue fixating on this picture even while the pronoun and following verb is played, as seen in Fig. 8. This is an interesting finding for two reasons related to attentional control: (1) the ASD group seems to get "stuck" processing previous linguistic information; (2) the ASD group spends time processing linguistic information that is rarely relevant for the task at hand. With regards to the former point, previous work has postulated that "sticky attention" is a manifestation of executive dysfunction in ASD (Kleinhans et al. 2005), but this is perhaps the first study to show that attentional stickiness can get in the way of online, linguistic processing. With regards to the latter point, for 80% of items, participants are asked to identify a pronominal referent. For only 20% of items (i.e., Filler items) are they asked about anything else, and these items were split between asking about the setting and the object referred to by the last word. Thus, the most effective strategy for successfully completing the experimental task would be to focus on the pronoun and its possible referents, rather than the object referred to by the last word. The NT group seems to adopt this strategy, focusing their visual attention on possible pronominal referents. The ASD group, on the other hand, starts looking at the object referred to by the last word as soon as they hear it, and they do not return their attention to possible pronominal referents (pictures of humans) until the end of the matrix clause. Thus, they seem to be distracted by task-irrelevant information which leads to pronoun-processing costs. Once again, the underlying knowledge of how pronouns work is not impaired or incomplete, but attentional issues not directly related to pronoun resolution itself interfere with processing the pronoun.

Limitations and Future Directions

In our original design of this experiment, our focus was on pronoun processing; as such, we did not expect to find so many differences during processing of the subordinate clause that precedes the pronoun. However, our results strongly suggest that group divergences in pronoun interpretation may be related to differences in the processing of reference generally, including proper names and full NPs. For example, we found evidence in this study that the ASD group may not show the Repeated Name Penalty; however, since we did not directly set out to test the RNP, this conclusion must remain tentative until a future study specifically designed to test it. Similarly, we found evidence that the ASD group processed the non-human last word of the subordinate clause earlier than the NT group, and showed more fixed attention to it, even to the point where it may have interfered with pronoun reference resolution in the following clause. A follow-up study explicitly designed to examine reference resolution of nominals (not just pronouns) throughout the sentence would clarify these trends. Differences in executive function (including both working memory and attention) also seem relate the tracking of reference in both groups; a follow-up that includes executive function data (measures of working memory and attentional reallocation abilities) could help clarify the extent to which these factors can explain processing differences in this study. Similarly, a study involving longer and/more complex discourse might reveal behavioral differences in pronominal interpretation, especially given the possibility of working memory and/or attentional differences between groups. Finally, given

current psycholinguistic theories that processing constraints on pronouns and other referring expressions may vary from language to language (Kaiser and Trueswell 2008; Filiaci et al. 2014), follow-up studies in languages other than English may also provide interesting comparisons.

Conclusion

Our behavioral and eye-tracking results suggest that children with ASD use both discourse context cues and structural cues for resolving pronominal reference. This finding calls into question the WCC theory for ASD, since WCC should lessen the involvement of discourse contextual information in determining a pronoun's referent. These findings are also interesting since they contrast so greatly with previous production literature that has overwhelmingly found that individuals with ASD use pronouns inappropriately. That literature finds that ASD participants use pronouns when they should use full NPs, resulting in ambiguous reference, and, vice versa, that they use full NPs when they should use pronouns, resulting in the violation of the pragmatic principles underpinning anaphor choice. Our results provide some evidence that differences in pronoun production in ASD are not due to a lack of knowledge about how pronouns work. That is, the ASD groups understand the principles that determine who pronouns refer to when other people use them, but they may struggle to follow those principles in production.

Instead, divergences in pronoun production may stem from processing differences, as our eye-tracking results suggest subtle processing differences between the two groups, including increased looks to the subject referent in the two context-mitigated conditions, and a delay in looking at the subject referent in the Neutral condition compared to the NT group. While the eye-tracking data are not inconsistent with the ASD group representing the structural cue for pronoun resolution as purely syntactic rather than pragmatically-based, our results as a whole are better explained by possible group differences in working memory and attention. In this case, the ASD group diverges from the NT group not in underlying knowledge of how pronouns work, but in the levels of activation of the various potential referents, exacerbated by both working memory and attention-shifting differences. This latter explanation is more consistent with both our behavioral data and the delayed looks to subject in the Neutral condition (which would be suspect in the case where a strong syntax-based preference for a local subject should hold). Because children with ASD show processing differences (issues maintaining referents activated in working memory; atypical attention to pronoun referents during processing) in the current experiment, when narratives were simple and short, there is reason to suspect that children with ASD may show more consequential problems processing referential information in natural contexts, when narratives are longer, less predictable, and more complicated.

Acknowledgment This research was supported by grant from the National Institutes of Health (NIH-NIDCD R01 DC012774-01; Grossman, PI).

Author Contributions Drs. MN and EZ developed the original idea for this paper. Dr. MN developed the experimental design and items, helped edit the audio recordings and visual stimuli, conducted data analysis on the interpretation data, and drafted the manuscript. Dr. EZ created the audio recordings and visual stimuli, conducted data collection, was responsible for the primary eye-tracking data analysis, and provided feedback on the manuscript. Dr. RBG oversaw data collection in her laboratory and provided feedback on the manuscript.

Appendix

See Table 6.

Table 6 Descriptive statistics for proportion of trials in which participants chose the subject of the subordinate clause as the matrix subject pronoun referent.

Context	N	Mean	SD	SE	CI
ASD					
Neutral	18	0.920	0.125	0.030	0.062
Biasing Other	18	0.630	0.277	0.065	0.138
Biasing Subject	18	0.926	0.101	0.024	0.050
NT					
Neutral	18	0.975	0.061	0.014	0.030
Biasing Other	18	0.667	0.212	0.050	0.106
Biasing Subject	18	0.975	0.061	0.014	0.030

References

Almor, A. (1999). Noun-phrase anaphora and focus: The informational load hypothesis. *Psychological Review, 106*(4), 748–765. https://doi.org/10.1037/0033-295X.106.4.748.

Almor, A., Kempler, D., MacDonald, M. C., Andersen, E. S., & Tyler, L. K. (1999). Why do Alzheimer patients have difficulty with pronouns? Working memory, semantics, and reference in comprehension and production in Alzheimer's disease. *Brain and Language, 67*(3), 202–227. https://doi.org/10.1006/brln.1999.2055.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Ariel, M. (1990). *Accessing noun-phrase antecedents*. New York: Routledge.

Arnold, J. E., Bennetto, L., & Diehl, J. J. (2009). Reference production in young speakers with and without autism: Effects of discourse status and processing constraints. *Cognition, 110*(2), 131–146. https://doi.org/10.1016/j.cognition.2008.10.016.

Banney, R. M., Harper-Hill, K., & Arnott, W. L. (2015). The Autism Diagnostic Observation Schedule and narrative assessment: Evidence for specific narrative impairments in autism spectrum disorders. *International Journal of Speech-Language Pathology, 17*(2), 159–171. https://doi.org/10.3109/17549507.2014.977348.

Bavin, E. L., Kidd, E., Prendergast, L. A., & Baker, E. K. (2016). Young children with ASD use lexical and referential information during on-line sentence processing. *Frontiers in Psychology, 7*, 1–12. https://doi.org/10.3389/fpsyg.2016.00171.

Bates, D.M., Machler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes. R package version 1.1–7.

Brock, J., Norbury, C., Einav, S., & Nation, K. (2008). Do individuals with autism process words in context? Evidence from language-mediated eye-movements. *Cognition, 108*(3), 896–904. https://doi.org/10.1016/j.cognition.2008.06.007.

Brock, J., Sukenik, N., & Friedmann, N. (2017). Individual differences in autistic children's homograph reading: Evidence from Hebrew. *Autism & Developmental Language Impairments, 2*, 1–10. https://doi.org/10.1177/2396941517714945.

Colle, L., Baron-Cohen, S., Wheelwright, S., & Van Der Lely, H. K. J. (2008). Narrative discourse in adults with high-functioning autism or Asperger syndrome. *Journal of Autism and Developmental Disorders, 38*, 28–40. https://doi.org/10.1007/s10803-007-0357-5.

Fay, W. H. (1979). Personal pronouns and the autistic child. *Journal of Autism and Developmental Disorders, 9*(3), 247–260. https://doi.org/10.1007/BF01531739.

Filiaci, F., Sorace, A., & Carreiras, M. (2014). Anaphoric biases of null and overt subjects in Italian and Spanish: A cross-linguistic comparison. *Language, Cognition and Neuroscience, 29*(7), 825–843. https://doi.org/10.1080/01690965.2013.801502.

Frith, U., & Happé, F. (1994). Autism: Beyond "theory of mind". *Cognition, 50*, 115–132.

Frith, U., & Snowling, M. (1983). Reading for meaning and reading for sound in autistic and dyslexic children. *British Journal of Developmental Psychology, 1*(4), 329–342.

Fukumura, K., & van Gompel, R. (2014). Effects of order of mention and grammatical role on anaphor resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(2), 501–525. https://doi.org/10.1016/j.jml.2009.09.001.

Grace, E. J. (2013). Pronoun Problems (Asides on Self-Loathing and Gender). Retrieved July 9, 2019, from https://tinygracenotes.blogspot.com/2013/05/pronoun-problems.html

Givón, T. (1983). Topic continuity in discourse: An introduction. In T. Givón (Ed.), *Topic continuity in discourse*. Amsterdam: John Benjamins Publishing Co.

Grice, P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics 3* (pp. 41–58). New York: Academic Press.

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science, 17*(3), 311–347.

Gundel, J., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language, 69*(2), 274–307. https://doi.org/10.2307/416535.

Happé, F. (1997). Central coherence and theory of mind in autism: Reading homographs in context. *British Journal of Developmental Psychology, 15*(1), 1–1.

Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 36*(1), 5–25. https://doi.org/10.1007/s10803-005-0039-0.

Hill, E. L. (2004). Evaluating the theory of executive dysfunction in autism. *Developmental Review, 24*(2), 189–233.

Jolliffe, T., & Baron-Cohen, S. (1999). A test of central coherence theory: Linguistic processing in high-functioning adults with autism or Asperger syndrome: Is local coherence impaired? *Cognition, 71*(2), 149–185. https://doi.org/10.1016/S0010-0277(99)00022-0.

Jordan, R. (1989). An experimental comparison of the understanding and use of speaker-addressee personal pronouns in autistic children. *International Journal of Language & Communication Disorders, 24*(2), 169–179.

Kaiser, E., & Trueswell, J. (2008). Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes, 23*(5), 709–748. https://doi.org/10.1080/01690960701771220.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2*, 217–250. https://doi.org/10.1105/tpc.11.5.949.

Karimi, H., & Ferreira, F. (2016). Informativity renders a referent more accessible: Evidence from eyetracking. *Psychonomic Bulletin & Review, 23*(2), 507–525. https://doi.org/10.3758/s13423-015-0917-1.

Kaufman, A., & Kaufman, N. (2004). *Kaufman brief intelligence test* (2nd ed.). Circle Pines, MN: American Guidance Service.

Kleinhans, N., Akshoomoff, N., & Delis, D. (2005). Executive functions in Autism and Asperger's Disorder: Flexibility, fluency, and inhibition. *Developmental Neuropsychology, 27*(3), 379–401. https://doi.org/10.1207/s15326942dn2703_5.

Kuijper, S. J., Hartman, C. A., & Hendriks, P. (2015). Who is he? Children with ASD and ADHD take the listener into account in their production of ambiguous pronouns. *PLoS ONE, 10*(7), e0132408. https://doi.org/10.1371/journal.pone.0132408.

Lee, A., & Hobson, R. P. (1994). I, You, Me, and Autism : An experimental study. *Journal of Autism and Developmental Disorders, 24*, 155–176.

Levinson, S. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT Press.

Lopez, B., & Leekam, S. R. (2003). Do children with autism fail to process information in context? *Journal of Child Psychology and Psychiatry, 44*(2), 285–300. https://doi.org/10.1111/1469-7610.00121.

Lord, C., DiLavore, P. C., Gotham, K., Guthrie, W., Luyster, R. J., Risi, S., et al. (2012). *Autism Diagnostic Observation Schedule: ADOS-2*. Los Angeles, CA: Western Psychological Services.

Mäkinen, L., Loukusa, S., Leinonen, E., Moilanen, I., Ebeling, H., & Kunnari, S. (2014). Characteristics of narrative language in autism spectrum disorder: Evidence from the Finnish. *Research in Autism Spectrum Disorders, 8*(8), 987–996. https://doi.org/10.1016/j.rasd.2014.05.001.

Malkin, L., Abbot-Smith, K., & Williams, D. (2018). Is verbal reference impaired in autism spectrum disorder? A systematic review. *Autism & Developmental Language Impairments, 3*, 1–24. https://doi.org/10.1177/2396941518763166.

Micai, M., Joseph, H., Vulchanova, M., & Saldaña, D. (2017). Strategies of readers with autism when responding to inferential questions: An eye-movement study. *Autism Research*, *10*(5), 888–900.

Mirman, D. (2014). Growth curve analysis and visualization using R, 168.

Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language, 59*(4), 475–494. https://doi.org/10.1016/j.jml.2007.11.006.

Naigles, L. R., & Tek, S. (2017). 'Form is easy, meaning is hard' revisited: (re) characterizing the strengths and weaknesses of language in children with autism spectrum disorder. *Wiley Interdisciplinary Reviews: Cognitive Science, 8*(4), 1–12. https://doi.org/10.1002/wcs.1438.

Norbury, C. F. (2005). Barking up the wrong tree? Lexical ambiguity resolution in children with language impairments and autistic spectrum disorders. *Journal of Experimental Child Psychology, 90*(2), 142–171. https://doi.org/10.1016/j.jecp.2004.11.003.

Norbury, C. F., & Bishop, D. V. (2003). Narrative skills of children with communication impairments. *International Journal of Language & Communication Disorders, 38*(3), 287–313. https://doi.org/10.1080/1368203100001081 33.

Norbury, C. F., Gemmell, T., & Paul, R. (2014). Pragmatics abilities in narrative production: A cross-disorder comparison. *Journal of Child Language, 41*(3), 485–510. https://doi.org/10.1017/S0305000091300007X.

Novogrodsky, R. (2013). Subject pronoun use by children with autism spectrum disorders (ASD). *Clinical Linguistics & Phonetics*. https://doi.org/10.3109/02699206.2012.742567.

Novogrodsky, R., & Edelson, L. R. (2016). Ambiguous pronoun use in narratives of children with Autism Spectrum Disorders. *Child Language Teaching and Therapy*. https://doi.org/10.1177/0265659015602935.

Perovic, A., Modyanova, N., & Wexler, K. (2013). Comparison of grammar in neurodevelopmental disorders: The case of binding in Williams syndrome and autism with and without language impairment. *Language Acquisition, 20*(2), 133–154. https://doi.org/10.1080/10489223.2013.766742.

R Core Team. (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rooryck, J., & Wyngaerd, G. V. (2015). Morphological transparency and the Delay of Principle B Effect. *Lingua, 155*, 121–139.

Rutter, M., Bailey, A., & Lord, C. (2003). *Social Communication Questionnaire (SCQ)*. Los Angeles, CA: Western Psychological Services.

Rumpf, A. L., Kamp-Becker, I., Becker, K., & Kauschke, C. (2012). Narrative competence and internal state language of children with Asperger Syndrome and ADHD. *Research in Developmental Disabilities, 33*(5), 1395–1407. https://doi.org/10.1016/j.ridd.2012.03.007.

Saldaña, D., & Frith, U. (2007). Do readers with autism make bridging inferences from world knowledge? *Journal of Experimental Child Psychology*, *96*(4), 310–319.

Sansosti, F. J., Was, C., Rawson, K. A., & Remaklus, B. L. (2013). Eye movements during processing of text requiring bridging inferences in adolescents with higher functioning autism spectrum disorders: A preliminary investigation. *Research in Autism Spectrum Disorders*, *7*(12), 1535–1542.

Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical evaluation of language fundamentals, fourth edition (CELF-4)* (4th ed.). Toronto, Canada: The Psychological Corporation/A Harcourt Assessment Company.

Snowling, M., & Frith, U. (1986). Comprehension in "hyperlexic" readers. *Journal of Experimental Child Psychology, 42*(3), 392–415. https://doi.org/10.1016/0022-0965(86)90033-0.

Suh, J., Eigsti, I. M., Naigles, L., Barton, M., Kelley, E., & Fein, D. (2014). Narrative performance of optimal outcome children and adolescents with a history of an autism spectrum disorder (ASD). *Journal of Autism and Developmental Disorders, 44*(7), 1681–1694. https://doi.org/10.1007/s10803-014-2042-9.

Tager-Flusberg, H. (1995). 'Once upon a ribbit': Stories narrated by autistic children. *British Journal of Developmental Psychology, 13*(1), 45–59.

Thornton, R., & Wexler, K. (1999). *Principle B, VP ellipsis, and interpretation in child grammar*. Cambridge, MA: MIT Press.

Wang, Y., Zhang, Y. B., Liu, L. L., Cui, J. F., Wang, J., Shum, D. H., et al. (2017). A meta-analysis of working memory impairments in autism spectrum disorders. *Neuropsychology Review, 27*(1), 46–61.