Do individuals with autism process words in context? Evidence from language-mediated eye-movements

Jon Brock a,b,*, Courtenay Norbury a,c, Shiri Einav a, Kate Nation a

a Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford, OX1 3UD, UK
b Macquarie Centre for Cognitive Science, Macquarie University, Sydney, NSW 2109, Australia
c The Department of Psychology, Royal Holloway, University of London, Egham, Surrey, TW20 0EX, UK

ABSTRACT

It is widely argued that people with autism have difficulty processing ambiguous linguistic information in context. To investigate this claim, we recorded the eye-movements of 24 adolescents with autism spectrum disorder and 24 language-matched peers as they monitored spoken sentences for words corresponding to objects on a computer display. Following a target word, participants looked more at a competitor object sharing the same onset than at phonologically unrelated objects. This effect was, however, mediated by the sentence context such that participants looked less at the phonological competitor if it was semantically incongruous with the preceding verb. Contrary to predictions, the two groups evidenced similar effects of context on eye-movements. Instead, across both groups, the effect of sentence context was reduced in individuals with relatively poor language skills. Implications for the weak central coherence account of autism are discussed.

1. Introduction

In 1959, Polan and Spencer proposed that “the primary symptom of early infantile autism is a lack of integration pervading all behavior of the organism and manifesting itself in the distorted language, in the lack of social responsiveness, and in the lack of adaptability to environmental changes” (p. 11). The notion of some form of integration deficit remains central to contemporary cognitive theories of autism (e.g., Minshew, Goldstein, & Siegel, 1997). Most prominently, perhaps, Frith (1989) coined the term weak central coherence to describe autistic individuals’ “inability to draw together information so as to derive coherent and meaningful ideas” (p. 187). This account has proven highly influential, motivating numerous studies of autism (Happé & Frith, 2006), and being linked to apparent abnormalities of brain connectivity in autism (Belmonte et al., 2004; Brock, Brown, Boucher, & Rippon, 2002). However, ‘integration’ and ‘central coherence’ remain elusive and ill-defined concepts, whose underlying cognitive mechanisms are still not properly understood.

In the current study, we investigated central coherence in relation to language comprehension. As Frith (1989) pointed out, “The meaning of any utterance in word or gesture can only be properly understood by... placing it in context” (p. 180). As such, weak central coherence has been linked to difficulties people with autism face with pragmatic conversational skills (Noens & van Berckelaer-Onnes, 2005), particularly their understanding of jokes and figurative language (see Norbury, 2004), as well as their poor performance on various measures of reading comprehension (e.g., Frith & Snowling, 1983). More direct evidence of impaired ‘contextual integration’ comes from studies showing that individuals with autism have difficulty generating inferences about events that are implied by a combination of statements but are not explicitly stated (e.g., Losh & Capps, 2003; Norbury & Bishop, 2002). Most pertinently, it has consistently been reported that children and adults with autism make less use of sentence context when reading ambiguous words. For example, they tend to pronounce the homograph
“tear” the same regardless of whether it occurs in the context of crying or ripping (Burnette et al., 2005; Frith & Snowling, 1983; Happé, 1997; Jolliffe & Baron-Cohen, 1999; López & Leekam, 2003).

Several recent studies have, however, failed to support the predictions of the central coherence account. Saldana and Frith (2007) found that children with autism showed a normal reduction in reading times for sentences that were congruent with events implied by previous sentences. Similarly, in a study using spoken rather than written stimuli, Norbury (2005) found that children with autism performed at the same level as language-matched peers when asked to decide whether a picture matched the meaning of a spoken sentence containing a homophone (e.g., “John fished from the bank”). Instead, performance was related to children’s language ability – those with poor language scores found the task more difficult, regardless of their autism diagnosis.

Interpretation of these discrepant findings is, unfortunately, complicated by the considerable linguistic and metalinguistic demands of the tasks used across different studies. Performance depends not only on the ability to integrate contextual information, but also on the comprehension of the individual words and on the maintenance of contextual information in working memory. Some tasks also require participants to re-evaluate initial interpretations of ambiguous material or to reason off-line about previously presented material, and most depend on basic reading skills, which are rarely evaluated comprehensively. In principle, any of these confounding factors could explain poor performance among individuals with autism.

To address these issues, we adopted a novel approach, recording the eye-movements of adolescents with autism as they monitored spoken sentences for words that matched pictures on a visual display. Typically, fixation patterns are closely time-locked to the ongoing verbal input, providing a continuous online measure of comprehension that is independent of any overt spoken or manual response (Cooper, 1974; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995), making this technique particularly suitable for research into developmental disorders (Nation, 2008). Studies have shown that eye-movements are sensitive to both semantic and phonological cues: for instance, on hearing the phrase “eat the cake”, participants tend to look towards a picture of a cake even before the onset of the word “cake” (Altmann & Kamide, 1999; Nation, Marshall, & Altmann, 2003); on hearing “beetle”, they look more at a beaker than at a phonologically unrelated object (Allopenna, Magnuson, & Tanenhaus, 1998; Tanenhaus et al., 1995).

Significantly for current purposes, these semantic and phonological effects appear to interact. Dahan and Tanenhaus (2004) reported that Dutch-speaking participants looked less at a phonological competitor if the preceding verb made it an unlikely referent. For example, participants tended to look at a camel [“kameel”] on hearing the word “cake” (Altmann & Kamide, 1999; Nation, Marshall, & Altmann, 2003); on hearing “beetle”, they look more at a beaker than at a phonologically unrelated object (Allopenna, Magnuson, & Tanenhaus, 1998; Tanenhaus et al., 1995).

2. Methods

2.1. Participants

The autism group comprised 24 adolescent boys, who each had clinical diagnoses of autism spectrum disorder and whose algorithm scores on the Autism Diagnostic Observation Schedule Module 4 (ADOS; Lord, Rutter, & Di Lavore, 1997) were all above diagnostic threshold (7) for autism spectrum disorder (mean = 11.8; SD = 3.3). To address the variation in language ability within the autism group, the comparison group included 13 typically developing boys recruited from mainstream schools, together with 11 adolescents (9 boys) recruited from special education schools for children with communication difficulties. Each participant in the latter subgroup had a formal diagnosis of language impairment from a certified speech-language pathologist, but had no other clinical diagnoses and did not meet diagnostic criteria on the ADOS.

Performance on language measures was highly intercorrelated, so a composite language score was derived for initial analyses. Table 1 shows that the comparison group was well-matched to the autism group for age, non-verbal ability and language skills.

2.2. Stimuli

A sample stimulus display is shown in Fig. 1. In the target-present condition, each display comprised the target...
object (e.g., a hamster), a phonological competitor (e.g., a hammer), and two unrelated distracters, themselves a cohort pair. In the target-absent condition, each display contained a phonological competitor of the target word and three unrelated distracters. The same 16 pictures were used across all experimental trials, with each picture assuming the role of target, competitor, or distracter on different occasions, effectively controlling for stimulus characteristics such as word frequency or picture salience. The position and identity of the target object and phonological competitor were fully counterbalanced.

Sentences (see Appendix) were digitally recorded by a female native English speaker. Each test sentence was five words long, consisting of agent, verb, definite article, target noun, and adverb. In constraining sentences, the verb was strongly associated with the target (e.g., "Joe stroked the *hamster* quietly"). In the neutral condition, the verb was always "chose" (e.g., "Sam chose the *hamster* reluctantly"). Filler trials involved a non-overlapping set of stimuli. The target word was a sentence-final noun occurring after the main object of the sentence (e.g., "Alex hit the ball with the *racquet*"). The target object was again present on half the trials, but there were no phonological competitors.

### 2.3. Procedure

The experiment was controlled using e-Prime (Schneider, Eschmann, & Zuccolotto, 2002). Visual stimuli were displayed on a 34 × 27 cm monitor with sentences presented via Sennheiser HD25 headphones. Eye-movements were recorded at 50 Hz using an ASL 504 remote infrared camera situated underneath the screen. The eye-tracker was calibrated prior to testing and was recalibrated at the experimenter’s discretion.

Participants were required to press a button if any word in the sentence matched any of the pictures in the display. Each trial began with a pulsating, centrally-located dot. Once the participant was looking at the centre, the experimental display was presented and remained onscreen until the trial’s conclusion. Sentences began 1000msecs after.

---

**Table 1**

<table>
<thead>
<tr>
<th></th>
<th>Autism N = 24</th>
<th>Comparison N = 24</th>
<th>t(46)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>15.0 (1.2)</td>
<td>14.5 (1.0)</td>
<td>1.60</td>
<td>.115</td>
</tr>
<tr>
<td>Nonverbal ability1</td>
<td>98.9 (15.1)</td>
<td>100.1 (15.9)</td>
<td>0.27</td>
<td>.788</td>
</tr>
<tr>
<td>Receptive vocabulary2</td>
<td>96.0 (22.7)</td>
<td>96.8 (19.1)</td>
<td>0.13</td>
<td>.897</td>
</tr>
<tr>
<td>Receptive grammar3</td>
<td>97.7 (10.7)</td>
<td>98.0 (10.5)</td>
<td>0.11</td>
<td>.914</td>
</tr>
<tr>
<td>Sentence memory4</td>
<td>86.9 (14.9)</td>
<td>87.5 (17.2)</td>
<td>0.14</td>
<td>.893</td>
</tr>
<tr>
<td>Language composite</td>
<td>65–120</td>
<td>65–120</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard deviations are in parentheses with ranges below. Statistical comparisons are based on independent two-tailed t-tests.

1 Performance subtests of the Wechsler abbreviated scales of intelligence (Wechsler, 1999).
2 British picture vocabulary scale 2nd edition (Dunn, Dunn, & Whetton, 1997).
3 Test for reception of grammar-2 (Bishop, 2003).
4 Recalling sentences sub-test of the clinical evaluation of language fundamentals-3rd UK edition (Semel, Wiig, & Secord, 2000), transformed to have a mean of 100 and standard deviation of 15.

---

**Fig. 1.** Screen-shot of experimental stimuli for a target-present trial (actual stimuli were in full colour on a black background). Eye-movements were analyzed by defining four rectangular regions of interest, each occupying 25% of the screen height and width and centred on one of the four objects in the display. The white box indicates the region of interest around the hammer and was not visible during the experiment.
the display onset. In total there were 64 test trials and 32 filler trials, presented in a fixed pseudo-random order such that the same pictures were not repeated on consecutive trials.

2.4. Analysis of eye-movements

For each analysis, a critical time window was defined (see below). Trials (autism $M = 12\%$; control $M = 5\%$) were rejected when more than 75% of data points within that window were not on the screen (due to excessive blinking, fixation off-screen, or eyetracker malfunction). We then calculated for each participant the proportion of time spent fixating within a region of interest (see Fig. 1), converted to log-odds (cf. Barr, in press). Analysis of covariance (ANCOVA) was employed to explore individual differences in eye-movements in terms of group membership, with language ability treated as a covariate (cf. Jarrold & Brock, 2004).

3. Results

3.1. Response accuracy

Behavioural performance was close to ceiling (see Fig. 2), although non-parametric tests revealed somewhat poorer performance among children with autism, $U = 188$, $p = .034$, and a significant positive association between language scores and performance, $\tau = .237$, $p = .038$. However, the effect of sentence type was non-significant, $U = 0.56$, $p = .571$, indicating that individual differences in performance reflected response requirements (e.g., difficulty withholding a response when the target was absent), rather than sensitivity to context.

3.2. Anticipatory eye-movements to the target in constraining sentences

In the target-present condition, the extent to which participants could anticipate the target object following a constraining verb provided an index of their understanding of the verb and their ability to map that verb onto a likely referent (cf. Nation et al., 2003). Fig. 3a shows that for both groups, there was an increase in target fixations well before the onset of the target word, at around 480 ms post-verb. Individual differences in the effect of verb semantics were analyzed by calculating the proportion of time spent looking at the target between 480 ms post-verb and the mean target onset at 940 ms (see Fig. 3b). A univariate ANCOVA revealed no significant effect of group or language ability or any interactions ($Fs < 1$).

3.3. Effect of sentence context on eye-movements to the phonological competitor

In the target-absent condition, participants demonstrated a strong phonological effect during neutral sentences, with an increase in fixations on the competitor between 340 and 860 ms following target word onset (Fig. 4a). This effect was attenuated markedly during constraining sentences (Fig. 4c). Data were analyzed by calculating the proportion of time spent fixating on the competitor during the 340–860 ms window (Fig. 4b and d). Repeated measures ANCOVA revealed a significant effect of sentence type, $F(1,45) = 7.55$, $p = .009$, $\eta^2_p = .144$. The effect of group on competitor fixations was non-significant, as was the interaction between group and sentence ($Fs < 1$). There was, however, a significant effect of language ability, $F(1,45) = 8.46$, $p = .006$, $\eta^2_p = .158$, which was mediated by a significant interaction between language ability and sentence type, $F(1,45) = 6.86$, $p = .012$, $\eta^2_p = .132$.

Separate regression analyses explored this interaction. Language scores were not associated with eye-movements during neutral sentences, $R^2 = .005$, $p = .621$. By contrast, in constraining sentences, individuals with poorer language scores spent significantly more time gazing at the (contextually inappropriate) phonological competitor, $R^2 = .216$, $p = .001$. As shown in Fig. 4d, this relationship between language scores and eye-movements was similar in both groups: neither the addition of group membership, $\Delta R^2 = .000$, $p = .904$, nor an interaction term enabling non-parallel slopes, $\Delta R^2 = .000$, $p = .915$, significantly improved the fit of the regression model (cf. Wright, 1997); and the association was significant both for the autism group, $R^2 = .177$, $p = .041$, and for the control group, $R^2 = .266$, $p = .010$.

Further regression analyses showed that each of the individual language tests was a significant predictor of eye-movements in this condition, $p < .025$, although none of them accounted for significant unique variance beyond the other two measures. There was, however, no significant association between eye-movements and non-verbal IQ, $R^2 = .035$, $p = .200$, and the association with composite language scores remained strong, even after variance associated with non-verbal IQ had been accounted for, $\Delta R^2 = .183$, $p = .002$.

Similar results were obtained when we analyzed the median time to first fixation on the target after the verb onset.

The effect of group and the group by condition interaction remained non-significant when the covariate was removed from the analysis.

---

Fig. 2. Proportion of trials correct.
Finally, we re-analyzed the data from this condition, treating autism diagnosis and language impairment as orthogonal categorical variables (cf. Norbury, 2005). Language impairment was defined as a language composite score outside the range of the typically developing members of the control group. Eleven language-unimpaired adolescents with autism were then pairwise-matched to within 2.5 points on the language composite measure to 11 typically developing adolescents and 7 language-impaired adolescents with autism were matched in the same way to 7 adolescents with a clinical diagnosis of language impairment. As shown in Fig. 5, eye-movements segregated according to language impairment rather than autism diagnosis. This was confirmed by univariate ANOVA, which revealed a significant effect of language impairment, $F(1,32) = 7.64$, $p = .009$, $\eta^2_{p} = .193$, but not autism diagnosis, $F < 1$.

4. Discussion

This study provides, to our knowledge, the first investigation of language-mediated eye-movements in autism. Consistent with previous studies of non-autistic children and adults, eye-movements were affected by the semantic association between the sentence verb and the target object and by the phonological overlap between the target word and competitor objects. Moreover, these two effects interacted such that participants looked less at the phonological competitor when the preceding verb made it an unlikely referent, even when the target object was absent from the display. In accordance with the central coherence account of autism, it was predicted that this effect of sentence context would be reduced in adolescents with a diagnosis of autism; however, this prediction was not supported.

The absence of group differences in the current study is consistent with other recent studies of central coherence in language comprehension reported by Norbury (2005) and by Saldaña and Frith (2007). It is, however, clearly at odds with earlier studies cited in support of the central coherence account, particularly those showing that individuals with autism have difficulty using sentence context to disambiguate homographs. Importantly, our diagnostic criteria were stricter than in published studies of homograph reading, we tested more participants per group, and each participant completed many more trials per condition. It seems unlikely, therefore, that the null group effect in the current study reflects either misdiagnosis or a lack of experimental power. Instead, consideration of task requirements may help account for these discrepant findings. As mentioned earlier, individuals with autism could have performed poorly in previous studies for a variety of reasons unrelated to their ability to utilise contextual information. Our eye-tracking measure avoided many of these confounds: there were no metalinguistic requirements; participants were not required to abandon or update their initial interpretations; and verbal working memory demands were minimal, with only one word (“the”) occurring between the contextual information and the target word. While we are unable to determine precisely which of these factors proved critical, it does appear that their elimination has removed any effect of autism diagnosis.
We are not, however, simply reporting a null result. Irrespective of their autism diagnosis, individuals with poorer language scores evidenced reduced sensitivity to sentence context. This is again consistent with the behavioural data reported by Norbury (2005) and, indeed, with several recent studies of individuals with specific language impairment (Friedmann & Novogrodsky, 2007; Sabisch, Hahne, Glass, von Suchodoletz, & Friederici, 2006). The advantage of the current study is that we are able to rule out a number of possible explanations for this association. As noted above, our eye-tracking task avoided several confounding factors that could contribute to poor performance. More importantly, language ability was not associated with the isolated effects of verb semantics (anticipatory eye-movements towards the target in constraining sentences) or phonological similarity (eye-movements to the competitor in neutral contexts). Given that the same words and pictures were used across conditions, this implies that the relationship between language ability and our measure of context sensitivity was not mediated by familiarity with the individual stimuli. Similarly, factors such as attention or motivation can also be ruled out because these would also have affected eye-movements across conditions.

Having eliminated these potential explanations, we are left to speculate on the nature of the relationship be-

Fig. 4. Competitor fixations in the target-absent condition for neutral and constraining sentences. (A and C) The probability of fixation is plotted relative to the onset and mean offset (540 ms) of the target word. Error bars represent ±1 standard error in the mean. (B and D) data from the highlighted time window are plotted as a function of language ability. Regression lines show this relationship for the autism group (thick) and control group (thin).
between language impairment and context-processing difficulties. One intriguing possibility is suggested by a recent functional magnetic resonance imaging study of healthy adults, showing that the use of sentence context to disambiguate acoustically degraded words is associated with increased functional connectivity within the left hemisphere (Obleser, Wise, Dresner, & Scott, 2007). The implication here is that reduced contextual sensitivity in individuals with language impairment may be an indicator of reduced connectivity. Given that disconnection within the left hemisphere is seen as the cause of many cases of acquired language deficits (see e.g., Catani & ffytche, 2005), it seems reasonable to hypothesize that ‘under-connectivity’ may also be implicated in some cases of developmental language impairment. An important question for future research, therefore, concerns the extent to which variation in language ability and sensitivity to context are associated with measures of connectivity between language-relevant brain regions in individuals both with and without autism.

This brings us to our final point. We have reported that sensitivity to sentence context is associated with an individual’s degree of language impairment, but that adolescents with autism experience no additional difficulties when language ability is tightly controlled for and confounding task demands are removed. It is, nevertheless, important to recognise that autism and language impairment are not independent, orthogonal deficits. Indeed, we were only able to match our two groups for language ability by over-sampling non-autistic children with a clinical record of language difficulties. Put another way, this study has shown that insensitivity to linguistic context is directly associated with language impairment, which in turn is a relatively common symptom of autism. Our findings are clearly problematic for current formulations of the central coherence account (e.g., Happé & Frith, 2006) and, consequently, for theories linking central coherence to neural connectivity (e.g., Brock et al., 2002). However, they lead us to conclude that further consideration of the links between cognitive and neural integration mechanisms and individual symptoms of autism will reveal important insights into the overlap between autism and other developmental disorders as well as the individual variability within the autism spectrum.

**Acknowledgements**

This research was conducted in the Department of Experimental Psychology at the University of Oxford. It was supported by grants from the Medical Research Council and the Wellcome Trust to Kate Nation and a Nuffield New Career Development Fellowship to Courtenay Norbury. We thank Kris Kinsey for his technical assistance, Tim Atkins for the script synchronizing the eyetracker with e-Prime, Sarah Jones for assistance in piloting the eye-tracking task, and Gerry Altmann for advice on eye-movement analyses. We are also grateful to Vic Ferreira, James Magnuson and two anonymous reviewers for their many helpful comments and suggestions. Warm thanks are extended to our participants, and to their schools and families, for their continuing support of our research. Correspondence may be addressed to Jon Brock (jbrock@macs.mq.edu.au) or Kate Nation (kate.nation@psy.ox.ac.uk).
Appendix.

Each sentence began with one of four gender-neutral names (Joe, Sam, Alex, or Ashley) and ended with one of six adverbs (rapidly, slowly, carefully, reluctantly, quietly, or sadly). 8 pairs of target words were selected based on high familiarity (467–631) and imagability ratings (529–618; Coltheart, 1981). The 16 target words are listed below, together with their corresponding verbs for the constraining condition and the make-up of the screen display for the corresponding target-present and target-absent trials. Although the same displays were presented on multiple occasions, the locations of the four objects were different each time.

<table>
<thead>
<tr>
<th>Target</th>
<th>Constraining verb</th>
<th>Target-present display</th>
<th>Target-absent display</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bucket</td>
<td>Filled</td>
<td>Bucket, buckle, money, monkey</td>
<td>Buckle, lettuce, monkey, cannon</td>
</tr>
<tr>
<td>Buckle</td>
<td>Fastened</td>
<td>Buckle, bucket, money, monkey</td>
<td>Bucket, candle, money, letter</td>
</tr>
<tr>
<td>Butter</td>
<td>Spread</td>
<td>Butter, button, candle, cannon</td>
<td>Button, coffee, hammer, medicine</td>
</tr>
<tr>
<td>Button</td>
<td>Fastened</td>
<td>Button, butter, candle, cannon</td>
<td>Butter, coffin, hamster, medal</td>
</tr>
<tr>
<td>Candle</td>
<td>Lit</td>
<td>Candle, cannon, button, button</td>
<td>Cannon, buckle, lettuce, monkey</td>
</tr>
<tr>
<td>Cannon</td>
<td>Fired</td>
<td>Cannon, candle, butter, button</td>
<td>Candle, bucket, letter, money</td>
</tr>
<tr>
<td>Coffee</td>
<td>Drank</td>
<td>Coffee, coffin, letter, lettuce</td>
<td>Coffin, butter, hammer, medical</td>
</tr>
<tr>
<td>Coffin</td>
<td>Buried</td>
<td>Coffin, coffee, letter, lettuce</td>
<td>Coffee, button, hammer, medical</td>
</tr>
<tr>
<td>Hammer</td>
<td>Banged</td>
<td>Hammer, hamster, medal, medicine</td>
<td>Hamster, butter, coffin, medical</td>
</tr>
<tr>
<td>Hamster</td>
<td>Stroked</td>
<td>Hamster, hammer, medal, medicine</td>
<td>Hamster, button, coffee, medicine</td>
</tr>
<tr>
<td>Letter</td>
<td>Wrote</td>
<td>Letter, lettuce, coffee, coffin</td>
<td>Lettuce, cannon, monkey, bucke</td>
</tr>
<tr>
<td>Lettuce</td>
<td>Ate</td>
<td>Lettuce, letter, coffee, coffin</td>
<td>Letter, bucket, candle, money</td>
</tr>
<tr>
<td>Medal</td>
<td>Wore</td>
<td>Medal, medicine, hammer, hamster</td>
<td>Medicine, button, coffee, hammer</td>
</tr>
<tr>
<td>Medicine</td>
<td>Swallowed</td>
<td>Medicine, medal, hammer, hamster</td>
<td>Medal, butter, coffin, hamster</td>
</tr>
<tr>
<td>Money</td>
<td>Paid</td>
<td>Money, monkey, bucket, buckle</td>
<td>Monkey, bucke, cannon, lettuce</td>
</tr>
<tr>
<td>Monkey</td>
<td>Fed</td>
<td>Monkey, money, bucket, buckle</td>
<td>Money, bucket, candle, letter</td>
</tr>
</tbody>
</table>

References


